**From Psychiatric Kinds to Harmful Symptoms**

GAULD Christophe1,2,3

1 Department of Child Psychiatry, University of Lyon, 59 Bd Pinel, 69 000 Lyon, France.

2 UMR CNRS 8590 IHPST, Sorbonne University, Paris 1, France

3 Institut de Recherches Philosophiques de Lyon (EA 4187), 1 rue de l’Université, 69007 Lyon, France

*\* Corresponding author:*

Christophe GAULD

Department of Child Psychiatry, University of Lyon, 59 Bd Pinel, 69 000 Lyon, France

Phone: 00 33 785 516 497

Mail: gauldchristophe@gmail.com

**Abstract**

Much research in the philosophy of psychiatry has been devoted to the characterization of the normal and the pathological. In this article, we identify and deconstruct two postulates that have held sway in the philosophy of psychiatry. The first postulate concerns the belief that clinicians would benefit from conceiving of psychiatric disorders as stable entities with clear boundaries. By relying on a symptom-based approach, we support a conception of psychiatric disorders whose symptoms are the products of many activated mechanisms in mutual interactions. The second postulate concerns the way in which the philosophy of psychiatry has approached the question of harm. We posit that clinicians perceive harms primarily through networks of clinical manifestations. The identification and deconstruction of these two postulates leads us to propose a *practical* definition of a psychiatric disorder that is useful for the clinician, while adopting a principle of prudential conservatism that does not exclude other *theoretical* definitions of psychiatric disorder that are useful as an epistemic hub. Thus, we propose the following definition of a psychiatric disorder that is relevant to clinical practice: “a set of clinical manifestations belonging to the prototypical central core of psychiatry, organized in a causal network, involving at least one harm, and whose set of harms exceeds the traditional threshold of significance for a given psychiatric disorder, directly related to an activated mechanism. Such a network composed of clinical manifestations is frozen in a state that is unable to allow the deactivation of these manifestations spontaneously without a therapeutic intervention”.

**Keywords:** symptom;psychiatric kinds; harmful dysfunction analysis; clinical practice; pluralism; symptom approach.

**Introduction**

In recent decades, much research in the philosophy of psychiatry has been devoted to characterizing the normal and the pathological. Attention has shifted between psychiatric kinds (Haslam 2002; Kendler et al. 2011; Zachar 2000, 2015), the conceptual analysis of demarcation criteria (Boorse 1977; Cummins 1975; Engelhardt 1975; Nordenfelt 2006), pragmatic (De Vreese 2017), gradualist (Hauswald 2016) and anticonceptual approaches (Hesslow 1993; Kincaid 2008). However, as pointed out by Kathryn Tabb (Tabb 2019), most of these approaches have focused on diagnostic kinds (Cooper 2013a, 2013b), i.e., on categories of psychiatric disorders (e.g. “schizophrenia spectrum”), as codified in taxonomies such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association 2013) or the International Classification of Diseases (ICD) (Cooper 2005; World Health Organization 2017). This initial interest in validating diagnostic categories (i.e., diagnostic kinds) can be explained by the desire to confirm the “scientific image” of psychiatry (Murphy 2006).

Over the last decade, the emergence of several original taxonomies has led to discussing psychiatric kinds in a more pluralistic framework (Tabb 2019), while maintaining the concern for the scientific legitimacy of the discipline (Ereshefsky 2009; Tekin 2016). In a very practical view of psychiatric kinds (Dupré 2002; Zachar 2014), examples of such pluralism are the following: i) the neuroscientific project of the National Institute of Mental Health of Research Domain Criteria (Cuthbert 2014) and its recent turn towards computational psychiatry (Ferrante et al. 2019), whose dimensional constructs may be considered as natural kinds; ii) the Hierarchical Taxonomy Of Psychopathology (HiTOP), whose hierarchical clusters may be considered as statistical kinds because they are grouped according to the latent variables of factorial analyses (Kotov et al. 2017); iii) and staging models, which support the existence of dynamical kinds (Fusar-Poli et al. 2018; McGorry et al. 2014; Shah et al. 2020).

However, despite this growing interest in a plurality of psychiatric kinds (Haslam 2002; Zachar 2000), the main focus has been on what Berrios (Marková and Berrios 2012) called a philosophical “syndrome-based approach”. This approach mainly targets psychiatric kinds as preconstructed entities whose (construct) validity should be analyzed (Cronbach and Meehl 1955; Messick 1989). This misses out on the important shift that research in psychiatry has undergone in the last decade which, in the name of personalized, stratified and precision psychiatry, i.e. personal psychology, has increasingly become interested in what may be called a philosophical “symptom-based approach”. This approach places value on big data from digital psychiatry, observations from everyday life (Cohen et al. 2020; Topol 2019), environmental, social and dynamic data from models of computational psychiatry (Kirmayer et al. 2015), the operationalization of phenomenology (Nelson et al. 2020; Sass et al. 2017) and on variables of symptom network approaches (Borsboom 2017; Borsboom et al. 2021). The common point between these psychiatric perspectives lies in two tendencies: their focus on explanatory pluralism and their interest in clinical phenomenology. The latter comprises signs, symptoms, and risk factors, e.g. thoughts, emotions, behaviors, somatic experiences, and social and contextual elements. For the sake of convenience, we shall group these elements together under the term of *symptoms* or clinical manifestations) (Brülde and Radovic 2006).

Our conviction is that in order to approach the issue of psychiatric kinds, the primary need is for a symptom-based approach that considers both the pluralist and clinical perspectives. This conviction can be justified: i) by the *relative* neglect of these trends in the literature on psychiatry in general and on the philosophy of psychiatry; ii) by the role that these trends should play in understanding patients seen in clinical practice; iii) by the need to re-examine how conceptual analysis can support the clinician in her/his practice (De Vreese 2021; Hesslow 1993), where there is not necessarily the need for a sharp distinction between psychiatric disorders despite their undeniable fuzzy boundaries; iv) by the apparent disharmony between the current proposals for establishing the boundaries between the normal and the pathological, and the perception of the clinician, who primarily seeks to identify the clinical manifestations of a patient. As an example of this symptom-based approach relying on Wakefield’s Harmful Dysfunction Analysis (HDA) (Wakefield 1992), the clinician may be seen as primarily and empirically collecting types of harm. Specific symptoms will allow him/her to identify low mood, acoustic-verbal hallucinations, painful jealousy, megalomaniacal delirium, etc. Indeed, what empirical evidence suggests that “clinicians are in effect assuming there is an evolutionary dysfunction” (De Vreese 2021)? Likewise, has any clinician ever seen a patient in consultation who was complaining of particular clinical manifestations, e.g., “harm related to a suspected dysfunction”? We argue throughout this article that the clinician rather assumes the existence of (activated) mechanisms that are related to each of the manifestations that she/he collects. Such a development takes place in the deep conviction that construct validity can only be questioned if content validity has first been analyzed – such contents being symptoms testifying to harm. We contend that the clinician’s conception of the patient’s disorder thus refers to modifications of mechanisms in networks built of interacting harmful symptoms (Gross 2011; Kim and Ahn 2002).

We propose to differentiate two definitions of “psychiatric disorder”. The first is a *theoretical* definition of psychiatric disorder. This is used to determine which conditions must be part of the pathological view, with all the administrative, financial, political, nosological consequences, the research, social organization of the health system, care policies and recommendations for good practice and therapies that such a definition may imply (Frances 2013; Spitzer 1983). Wakefield’s harmful dysfunction analysis provides a definition that has strong epistemic and pragmatic potential to serve as the *epistemic hub* for such a definition (Keuck 2011; Wakefield 1992). An epistemic hub corresponds to a scientific tool (here, a definition) that centralizes knowledge perspectives from multiple sources. By serving as a “concentrator” for different types of information, such hubs can facilitate communication exchanges between different institutions (e.g., medical institutions and patient communities) or different disciplines (e.g. medicine and theoretical neuroscience). The second analysis relates to a *practical* definition that is useful for clinical psychiatric practice (Engelhardt 1981, 1984; Krueger 2015). It considers psychiatric disorders through the prism of the clinician, who questions the care, clinical decisions, values, and norms concerning a given patient (Cooper 2020). For diagnostic and therapeutic management (e.g., reorientation, advice, support, risk prevention, and elimination of differential diagnosis), the clinician needs a specific definition that can be used intuitively. Subsequently, we provide an original definition of a psychiatric kind.

Of note, in the debate on psychiatric kinds, it seems necessary to clearly differentiate the question of the normal and the pathological (for instance: “Do the symptoms belonging to the category of depression together constitute a psychiatric disorder?”) from the question of dimensionality (for instance: “How much sadness is present in this patient with a clinical picture referring to the category of depression?”) and of the differential diagnosis (for instance: “How are these symptoms belonging to the category of depression distinct from those of generalized anxiety?”). In this paper, based on the notion of harmful symptoms, we seek to answer the first of these questions, aka the differentiation between normal and pathological.

Our starting point is therefore the need for a pluralist and phenomenological approach based on the symptom. To achieve this, we need two definitions: a first analytical, theoretical one (e.g. HDA), and a second that is more practical (which we provide below). To distinguish these two definitions methodologically, we propose to discuss two postulates found in the philosophy of psychiatry, which constitute the two main parts of this work. We believe that these assumptions raise a potential objection to our definition that we intend to be useful for clinical practice (second analysis). The first postulate concerns the belief that clinicians perceive psychiatric disorders as entities with clear boundaries, underpinned by a single dysfunction. This first postulate opens the way to thinking in terms of a symptom-based approach, i.e. symptoms interact with each other and are self-influencing (Borsboom 2017; Zachar 2021). The second postulate concerns the way in which the philosophy of psychiatry has approached the question of harm. We assume that clinicians perceive harms primarily through clinical manifestations. The analysis of these two postulates leads us to propose a definition of a psychiatric disorder that is easily intuitive and does not exclude the possibility of a second definition, which may act as an epistemic hub.

**Postulate I – Dysfunction of a whole psychiatric disorder and local mechanisms**

In this section, we show that the clinician’s conception of a psychiatric disorder is best informed by an network that activates local mechanisms, rather than by the analysis of a single dysfunction underlying a unique disorder.

1. **Symptom-based approach**

*Mutual causal influence.* Clinical experience, supported by empirical research such as Kim and Ahn (Kim and Ahn 2002), has shown that clinicians tend to think intuitively about psychiatric pictures in terms of mutual causal influences between clinical manifestations. This intuition supports cognitive-behavioral therapies. For example, the five columns of Beck, which underpin these therapies, are based on the identification of a situation, which itself influences an emotion via automatic thoughts. This emotion then leads to behavioral responses and consequences which in themselves reinforce the same set of automatic thoughts when similar situations arise (Beck 1993). In these models and associated therapies, symptoms influence each other.

Such a network of mutually interacting symptoms is made up of nodes (symptoms, or other environmental variables, or biological markers) and connections between these nodes (Borsboom 2008). Symptoms have a particular causal role in the network approach since each one contains an element of causality within the network. Psychiatric disorders can then be considered as vicious circles of causal problems (Stein 2021). Hence, on an ontological level, psychiatric disorders can be understood as homeostatic property clusters (Boyd 1991, 1999; Casali 2021; Kendler et al. 2011; Kendler 2008; Wakefield 2012; Westen 2012) whose components are phenotypic elements (and possibly other biomarkers).

*Release of the latent variable*. This mutualist conception implies abandoning the view of the psychiatric disorder as the product of a single underlying cause, i.e. a cause qualified more neutrally as a “latent variable” (Borsboom 2017; Jones and Robinaugh 2021). A variable latent is broadly defined as whatever factor underlying a set of expressed elements. Thus, instead of considering that a latent variable aggregates symptoms together, a psychiatric kind – or even a psychometrical construct – would be composed of a set of symptoms in mutual interaction (Zachar 2014)[[1]](#footnote-1). For example, anhedonia would lead to anxiety, which would lead to arousal insomnia, which in turn would make the initial anhedonia worse. Low mood and anhedonia are considered to be the visible consequences of a latent variable called ‘depression’. In this view, ‘depression’ would only have a literal (or ostensive – (Sadegh-Zadeh 2008)) or stipulative value. The latent variable can be statistical, e.g., when a factorial dimension aggregates symptoms bottom-up in the HiTOP, or when the Hamilton Rating Scale for Depression accounts for a “depression” dimension (Hamilton 1960; Kotov et al. 2017). It can also be physiological, e.g., when a set of genetic mutations is invoked to account for a disorder. Finally, it can be conceptual, e.g., the function conceived according to “black box essentialism” (Putnam 1975; Wakefield 1999). For example, dysfunction as detailed in the HDA corresponds to such a latent variable.

In clinical practice, sets of symptoms are not considered to be indicators of supposed latent variables. Not all have the same inferential weight in diagnostic, prognostic, predictive or therapeutic terms. We argue that it is the set of relationships between symptoms that may best reflect their presence in a single clinical picture, and not a supposed latent variable. Many diagnostic difficulties in psychiatry, such as the heterogeneity of disorders or comorbidities, can be partially overcome by ceasing to take the existence of latent variables for granted and by considering instead interactions between symptoms (Borsboom et al. 2017; Fried et al. 2017a; McNally 2016). Of course, the rationale that latent variables are not useful to clinicians does not mean that such variables should be completely excluded from all consideration in psychiatry. For example, some symptom network studies have used latent variables that influence specific symptoms. The body mass index (BMI) constitutes a well-identified latent variable, explaining the expression of several symptoms (Fried et al. 2020). Conversely, no latent variable as a dysfunction can explain all symptoms of the DSM category of major depressive disorder. Here, we are not trying to define the nature of a dysfunction. Rather, we view dysfunction as a latent variable, and we support the idea that no latent variable underlies psychiatric disorders.

Renouncing the idea of latent variables requires reconsidering the medical model, which conceptualizes symptoms as consequences of an underlying pathological entity. Based on the notion of latent variables, the medical model accounts for inter-aggregation of symptoms. In this medical model, all the symptoms of a disorder result from a single latent variable. Thus, from a theoretical and diagnostic point of view, all symptoms are situated on an equivalent and interchangeable level (Lux and Kendler 2010; Schmittmann et al. 2013). Based on the need to prioritize certain symptoms over others in clinical practice and on the impossibility of a single latent variable explaining the aggregation of a set of disparate symptoms (e.g., anhedonia and insomnia in a same patient), we argue that a unique dysfunction cannot account for the aggregation of the symptom set of all psychiatric disorders.

1. **Concept of central core**

*Central core*. If the clinician does not necessarily seek to identify a latent variable and does not rely on a syndrome-based approach, how can she/he find unity in the clinical picture of an individual presenting a set of interacting symptoms? We propose to consider the existence of a central symptom core within the network (or homeostatic property cluster). It is indeed possible to determine which symptoms are the most central within a network (Bringmann et al. 2019; Jones et al. 2019). Depending on the context (Tsou 2007), such a core of central symptoms is a set of symptoms that is stable over time between patients and is independent of the epistemic interests of researchers (Amundson and Lauder 1994). This central core defines the psychiatric kinds useful to the clinician (Zachar 2000). The network also contains peripheral symptoms that can influence the central core without necessarily being part of the definition of a psychiatric kind (Sadegh-Zadeh 2008). For example, a social rank factor or inflammation biomarker may influence the clinical picture but will not match the central core that defines the psychiatric kind. Many diagnostic difficulties in psychiatry, such as the heterogeneity of disorders or comorbidities, can be partially overcome by ceasing to take the existence of latent variables for granted and by considering instead interactions between symptoms (Borsboom et al. 2017; Fried et al. 2017b; McNally 2016). Of course, the rationale that latent variables are not useful to clinicians does not mean that such variables should be completely excluded from all consideration in psychiatry. For example, some symptom network studies use latent variables that influence specific network symptoms, such as body mass index (BMI), which influences several symptoms, themselves belonging to major depressive disorder described in the DSM. However, such a latent variable does not fall within the definition of psychiatric disorders.

As we will develop, support for the notion of a core structure can be found in the literature on psychiatry and in the philosophy of psychiatry, where a core structure is implicit in the use of various terms. First, central cores are found in the science of complex systems (Cucuringu et al. 2016; Ma and Mondragón 2015; Rombach et al. 2014). A system is defined by a set of elements linked together by relationships (Bertalanffy 2015). One of the main objectives of the science of complex systems is to isolate and analyze the properties of the elements belonging to these systems, assuming that a combination of some of the properties of the system can be reconstructed to enable an understanding of the behavior of the system as a whole. Therefore, elements that can be considered central in the system need to be identified because these elements form the essence or core of the network.

Second, symptom network theory in psychopathology uses the notion of a core structure to refer to the (stabilizing) conservatism of Principle 2 (component/symptom correspondence), according to which “the components in the psychopathology network correspond to the problems that have been codified as symptoms in the past century and appear as such in current diagnostic manuals” (Borsboom 2017). According to a semantic conception, symptom network theory is based on a set of principles, which are used in network modeling. The four principles that constitute symptom network theory refer to the characterization of a disorder as a network of interacting symptoms (Principle 1), to correspondence between symptoms and problems codified by the history of psychiatry (Principle 2), to the necessity of direct causal relationships between symptoms (Principle 3), and to the nontriviality of network topology, given that some symptoms are more closely related than others (Principle 4). A common conclusion of empirical studies on symptom networks is that the most commonly found central symptoms of a network (e.g., monothetic symptoms, such as low mood or anhedonia) are also the most relevant in terms of symptom stability, manipulability, and prevention, in addition to self-reported harm and external validity (Beard et al. 2016; Bringmann 2021; Fried 2015; Fried et al. 2016; van Borkulo et al. 2015).

Third, the notion of a central core has been attested to culturally. Some symptoms from the depression category are observed in all cultures (Kleinman 1988; Lahey et al. 1996). Furthermore, the World Health Organization has explicitly stated in several reports the need to base the clinical assessments of patients on a core set of items to which others, including culture-specific open-ended items, can be added (for instance, within questionnaires such as the Present State Examination) (Sartorius and World Health Organization 1983). A WHO intercultural study identified a “common core” of depressive symptoms in 75% of patients diagnosed with depression worldwide: sadness, loss of joy of life, anxiety, tension, lack of energy, decreased interest and focus, and feelings of worthlessness (Sartorius and World Health Organization 1983).

Fourth, the notion of a central core is supported by the naturalistic attempts of the philosophy of psychiatry to explore the possibility of what Cooper called “repeatables” (i.e., any entity, such as a dimension or a category) in which “certain entities in the world are alike and will behave similarly” (Cooper 2012) – a term that avoids ruling on the metaphysical nature of a disorder. Cooper insists on a purely clinical notion of repeatables, constituted of symptoms, found repeatedly during the formulation of clinical cases (Cooper 2012). This notion of repeatables can be partly understood by drawing on the notion of a prototype.

*Components of the central core*. Clinically, however, what are the components of such a central core? In practice, these are mainly the monothetic symptoms of DSM categories, such as anhedonia, low mood, acoustic-verbal hallucination or ideo-affective disorganization. While the entire network can be compared to a homeostatic property cluster, the symptoms constituting the central core of the network could be considered as a *prototype* of resemblance (Sadegh-Zadeh 2008). This view is justified by its fuzzy limits (Lilienfeld and Marino 1995; Sadegh-Zadeh 2001). The prototype (or central core) induces actions in members of the psychiatric community to help, support, monitor, reorient or improve the condition of the person presenting a core set of symptoms in mutual interactions. Thus, one of the main characteristics of such a central core is to stimulate pragmatic action (discussed in last part), insofar as the members of the community share fundamental values ​​and attitudes on the notion of harm. One of the main goals of a central core is to stimulate pragmatic action, insofar as the members of a community share fundamental values ​​and attitudes on the notion of harm. In this paper, our goal is to provide a general definition focused on presupposed central symptoms.

In an idiographic way, an individual may have some symptoms of a prototypical core but not others. In other words, only certain manifestations of a prototypical core will be present in a specific patient. For example, it is rare for a patient to present with compulsions, delirium, hallucinations, anhedonia, anxiety *and* a pragmatic communication disorder. Only the presence of some of the symptoms in the central core qualifies the patient as belonging to a prototype. We show below that this belonging also depends on the harm threshold (Sadegh-Zadeh 2008; Westen 2012).

*Relationships within the central core*. A symptom network may therefore be considered to consist of a stable central core. A question that naturally arises when describing a set of interconnected symptoms is where these connections come from, and what their nature is (Cramer et al. 2016). Indeed, being able to claim that one symptom interacts with another requires specifying whether this interaction acts directly at the level of clinical manifestations, e.g., “anxiety” activating “early awakening”, or at the level of local mechanisms accounting for each symptom (assuming that every symptom is underpinned by a mechanistic substrate). As we will see, activating/modifying a local mechanism that accounts for anxiety may activate/modify a local mechanism that accounts for early awakenings.

1. **Notion of mechanism**

*Activations of mechanisms.* A mechanism should be activated if one clinical manifestation is to be considered to influence another (Craver 2001; Kendler et al. 2011; Tsou 2015). We return to the notion of activation in the fourth part. For instance, activating a mechanism that causes low mood can lead to sadness or melancholy. It is important to distinguish between the clinical manifestation (state) (or symptom) and the mechanism that justifies the existence of this manifestation (Machamer et al. 2000; Wakefield 2006). It is essential to defend the existence of changes in the mechanisms that account for symptoms in order to prevent psychiatric disorders from being considered only as aggregates referring to artificial kinds, especially social constructions (Szasz 1960). This makes it possible to propose that a naturalistic (stable) central core underlies psychiatric kinds conceived as a central core of symptoms.

The relationship between a physiopathological mechanism and a clinical manifestation may be particularly difficult to establish due to neurobiological heterogeneity (Clementz et al. 2016, 2020; Ivleva et al. 2017). For instance, across the psychosis spectrum, it is difficult to derive etiological and neurobiological elements from clinical phenomenology, even when supported by multivariate constructs and laboratory tests (e.g., magnetic resonance imaging, scales, and experimental protocols) (Clementz et al. 2016). However, this observation does not challenge the idea that clinical manifestations are potentially related to cerebral abnormalities. In other words, being aware of such mechanisms but not necessarily having actual knowledge of them allows clinicians to better understand the situation of their patients. Indeed, as supported by the black-box essentialism conception (Medin and Ortony 1989; Wakefield 1999), the mechanism involved can be presumed, i.e., not directly identified empirically. However, by considering the potential existence of these manifestations and their interacting mechanisms that clinicians obtain a clinical overview of the evolution and natural history of disorders. For example, the intuition that impulsive behavior in a child accompanied by symptoms of attentional blinking and hyperactivity may have a biological basis (e.g., pro-dopaminergic), rather than a psychological basis, allows the clinician to guide the parents and child toward appropriate care, such as treatment with methylphenidate, which acts on dopaminergic pathways. By identifying a variety of interacting mechanisms in a psychiatric patient, the clinician can identify a specific treatment plan tailored to each patient. Such plans may involve psychoeducation, drug treatment, psychotherapies, and clinical consultations. They may also include short- and long-term monitoring, person-centered advice, and subjects of interest that will be discussed with the patient.

*Mechanisms at work within networks*. But what kind of mechanism are we talking about? Interest in the mechanism for characterizing diseases has grown in recent years. Two conceptions of the mechanism for defining a disorder have been proposed. The first posits that the disorder itself is conceived of as a dysfunctional mechanism: the term ‘mechanism’ is used here in reference to the typical causal processes in a healthy, normally functioning organism, while the disease is akin to the dysfunction of some of these internal mechanisms (Darden 2008; Moghaddam-Taaheri 2011). The second conception proposes that the mechanism exhibits variations as a system, independent of the disorder: the mechanisms underlying diseases are thus autonomous and organized independently of what defines a disorder (Illari and Williamson 2011; Nervi 2010). Like Gross (Gross 2011), we accept these two views at the same time by adopting a hybrid position, i.e., the disorder may itself be seen as a mechanism composed of interrelated symptoms, a practical conception useful for clinical practice. These symptoms need to be caused by the disruption of mechanisms aggregated into a network composed essentially of physiological parameters but also potentially of cultural, social and environmental parameters[[2]](#footnote-2).

1. **Summary**

We argue that there is no latent variable underlying a psychiatric disorder, i.e. no single dysfunction corresponding to a disorder. Instead, we propose the existence of a symptom network conceived as a homeostatic property cluster and underpinned by the activation of local mechanisms. Each of these symptoms is subsumed by a set of mechanisms that are mostly physiological but may also be cultural, social and environmental. A psychiatric kind is built on all the symptoms of the central core of this network, i.e., of a prototype composed of symptoms in mutual interactions. We argue that the clinical manifestations in mutual interactions are necessarily caused by many local mechanisms activated beyond a threshold.

Beyond Borsboom’s nominalist proposals (Borsboom 2017), we propose that all psychiatric symptoms are underpinned by the activation of a local mechanism. Moreover, we propose the novel idea of a central core. The two proposals described in this part are developed and discussed in the next section in relation to harm, the notion of activation, threshold, and the definition of a psychiatric disorder. We examine the nature of symptoms. We show that by proposing a set of symptoms with mutual interactions, we can reconsider the status of harm as expressed by the patient in clinical practice.

**Postulate II – Harm caused by a dysfunction, and the analysis of harmful symptom**

1. **Role of harm in philosophy of psychiatry**

To our knowledge, harm has mainly been approached in two ways in the literature on the philosophy of psychiatry: i) analyzed as a consequence of a dysfunction (Wakefield 2013; Wakefield and Conrad 2019; Wakefield and First 2013); ii) analyzed as a relative independent element of dysfunctions, in its relation to norms, values ​​or context (Cooper 2005; Gagné-Julien 2020).

This apparent dichotomy was already reflected in the 1992 Wakefield article, which described a disorder as a dysfunction that directly results in harm. Indeed, in the first view (harm as testifying to a dysfunction), Wakefield wrote: “Any kind of harm due to a dysfunction of an internal mechanism could be called a disorder”. However, he advised the following: “Note that the link, *via symptoms,* to the concept of harm is essential to the claim that the failure of repression [dysfunction] is a disorder” (emphasis added). In this statement, the harm is “connected” to the dysfunction *via* the symptoms. But if the harm is the consequence of a (natural) dysfunction, itself a consequence of the alteration of a supposed mechanism (Wakefield 1999), the question of the status of the symptoms may be posed as follows: at what point of the analysis should the symptoms be integrated, since they are the only manifestations expressed by the patient and / or noted by the clinician (Wakefield 2013; Wakefield and Conrad 2019; Wakefield and First 2013)?

In Boorse (Boorse 1977) and Wakefield (Wakefield 1992), the aggregation of a set of symptoms indicates a dysfunction (in Wakefield, this dysfunction causes harm). However, except in the case of monothetic categories, patients (or society) never complain of a dysfunction involving its survival or reproduction (and they complain even less in clinical practice of a dysfunction in relation to its survival or reproduction for the evolution of their species). They do not identify a prejudice related to a supposed dysfunction. In other words, patients never say, “I am suffering because I have a dysfunction indicating depression”, but rather, “I am suffering from my anhedonia, my fatigue and my abulia”. Likewise, nobody in society will say: “She / he has a disorder belonging to the spectrum of schizophrenia, testifying to a dysfunction, and resulting in harm”, but rather: “Her/his delirium of megalomania and her/his disorganization is harmful”. It could be argued that the opinion of clinicians or patients does not matter when defining a disorder, that the complaint is a semantic issue, and that a patient with cancer, who knows nothing about her/his bone metastasis, will complain of her/his pain without the complaint invalidating medical knowledge about the primary tumor. Yet it is precisely these arguments that we seek to criticize by proposing a practical definition, which is supposed to be intuitive for the clinician. In other words, we propose that what the clinician encounters should primarily constitute the basis for defining a disorder.

1. **Concept of harmful symptom**

Thus, what is harmful for a patient is not the consequence of a dysfunction aggregating all their symptoms. Rather, it is an intrinsic component of one of the symptoms (we define below what we mean by “component”). Not all the symptoms of the central core of a psychiatric condition have the same value all the time (Tabb 2019; Zachar 2021). Some of them, at certain times, in certain settings and in certain patients, are harmful. There may indeed be symptoms which do not cause any harm. For example, people who hear voices may say that their acoustic-verbal hallucinations contribute to the well-being, and or people with autism spectrum condition or their relatives may say that they have a form of neuro-atypicity that simply modifies but does not damage the type of communication they have with others. [[3]](#footnote-3)

We argue that a philosophical conception based on *harmful symptoms* is more consistent with the implicit position of the clinician than a conception based on a single harmful dysfunction. This conception is reinforced by the fact that the presence, in the DSM, of the criterion of clinical significance (stipulating that certain symptoms “must cause suffering or significant repercussions”): i) is present in the majority of disorders in which the symptoms themselves cause no obvious harm (Wakefield 2013; Wakefield and Conrad 2019; Wakefield and First 2013); ii) is used by the clinician when observing interacting symptoms. Even more empirically, the idea that any symptom can be considered harmful has been used in epidemiological surveys, and notably in the Epidemiological Catchment Area study, the first major psychiatric epidemiological survey based on the DSM in the United States (Eaton et al. 1981; Shapiro et al. 1984). In that study, when a symptom was reported by a patient, the interviewer asked a series of questions about its severity, and the symptom was not recorded until it reached a certain threshold.

Unlike deficiency, which refers to a biological difference (e.g., the congenital absence of limbs) and impairment, which refers to an inability to perform daily activities of living (Oliver 1996), harm has been described as the distress or disability experienced by a person or reflected in the values ​of a society (Wakefield 1992). Harms arise with regard to the cultural norms accepted in a given time frame and culture. In the present paper, we draw mainly on the description of harms put forward by Wakefield “in accordance with social values” (Wakefield 1992). We define harm as “compromising the capacity of a person to lead a normal or unhindered life according to cultural norms” (Cooper 2021). We justify our use of these definitions based on the fact that the notion of harm has been debated in the philosophy of psychiatry for more than 20 years and that despite its limits, pitfalls, and critics (De Block and Sholl 2021; Powell and Scarffe 2019a, 2019b), Wakefield’s definition of harm relating to threshold, social values, and alteration of life according to social values is fundamental for our argument.

Of course, we could seek to demonstrate again why the threshold of harm described by Wakefield (Spitzer and Wakefield 1999; Wakefield and Spitzer 2002) or by Spitzer (Spitzer 1999) seems interesting, but these discussions have already been widely discussed in the literature (Faucher and Forest 2021).

1. **Harms and psychiatric kind**

What about this notion of using harmful symptoms to define clinically useful psychiatric kinds? Symptoms may be present in a patient but not harmful, such as slightly low mood or slightly marked anhedonia. A patient might even report the presence of all nine symptoms belonging to the construct of depression, without any of them exceeding the threshold of harm historically considered significant. To throw light on this issue, we argue that, in order to belong to the field of psychiatry, a psychiatric kind should exhibit: i) at least one harm directly “integrated” in a symptom; ii) *and* that this or these symptoms should belong to the prototypical central core of the conditions considered by psychiatry; iii) *and* that this harm or the sum of the harms should exceed the threshold of clinical significance commonly accepted in the history of the discipline; iv) *and* at least one of the harms should activate a (biological) mechanism.

Regarding (ii), a clinically useful psychiatric kind should contain at least one psychiatric symptom belonging to the prototypical central core of psychiatry, directly harmful to the person, and at a certain threshold. For example, one of the central manifestations of the diagnosis of bone fracture could be bone rupture. The harm may be related to pain, which is indirectly harmful since it is secondary to the central clinical manifestation, i.e. bone rupture. The same is true for harm related to functional disability. Conversely, having a low mood is directly detrimental to the patient, e.g. low mood or anhedonia.

Regarding (iii), it is the total (weighted[[4]](#footnote-4)) harm of all clinical manifestations which provides a threshold of significance for the psychiatric disorder.

Concerning (iv), if the symptoms are diagnostic indicators and are part of what we might call with (Demazeux 2019) an “appareil sémiologique”, which is never neutral, what does it mean to say that the harm is “integrated” in a symptom (Cooper 2005, 2013a, 2014, 2020)? The answer can be formulated in two ways. First, on the level of phenomenology, it would be appropriate to study *what it is like to feel or experience* suffering through a symptom. Secondly, the harm can be explained by the presence of an altered mechanism, accounting for the symptom, as described in the previous section. The loss of a loved one, the discovery of cancer, the failure of an examination, or all other kinds of reasons, sometimes temporary, during which psychological suffering is well identified but for which there is no reason to think about the disease so far. Should these more or less transitory conditions of anguish, unhappiness, worry, poor sleep be considered as harmful symptoms? The answer is certainly in the affirmative: these conditions undoubtedly cause harm to the person.

However, such a finding is of little help in defining a psychiatric kind. In the absence of deterioration of at least one physiological mechanism, any non-medical (non-psychiatric) harm would fall within the scope of care (for example, harm related to precariousness or to tax evasion) (Tsou 2021a). Therefore, we propose that at least one harm (e.g., social exclusion) of the central core of symptoms should be the *direct* consequence of a clinical manifestation (e.g., paranoid delirium), which is itself the direct consequence of the deterioration of a (physiological) mechanism. Note that this account does not preclude the existence of harms that are consequences of socially harmful environments (e.g., discrimination against individuals with paranoid delusions).

In the next section, we summarize the implications of identifying and deconstructing these two postulates and propose a potentially workable definition of a psychiatric disorder for the field of clinical practice.

**A potentially workable definition of a psychiatric disorder for the field of clinical practice**

In the previous sections, rather than adopting a syndrome-based approach in which psychiatric kinds are defined according to an assumed latent variable (Cooper 2002; Wachbroit 1994) and mainly on the basis of diagnostic kinds derived from the DSM, we have argued that symptoms in mutual interactions (Borsboom 2017) are organized into causal networks (Kendler et al. 2011; Tsou 2016). We have seen that the symptoms belonging to a prototypical central core are associated in a network. Some of the symptoms of this core convey harms directly, as testified by the patient and collected by the clinician (Cooper 2005). It is these central symptoms that define the clinically useful psychiatric kind. Moreover, symptoms are underpinned by the activation of a local (often physiological) mechanism. In other words, each activated mechanism causes harm that is expressed in a symptom.

1. **Advantage of having useful psychiatric kinds for the clinician**

A clinically relevant symptom-based approach has clinical, social / political, and scientific advantages.

Clinically, this conception is intuitive for the clinician, i.e., it allows her/him to access relevant and rapid diagnostic, predictive, prognostic and therapeutic heuristics. A growing body of evidence shows that clinical reasoning is largely guided by network designs built on interacting variables (Bhugra et al. 2011; Kim and Ahn 2002; Meehl 1967). In addition, the notion of harmful symptoms diverts attention away from the question of comorbidities and the ambition of setting a clear limit between disorders, since its aim is to explain the aggregations of symptoms relating to each individual in her/his contextual environment.

On a social and institutional level, this conception involves a diagnostic threshold identical to that traditionally provided by the HDA and DSM. Indeed, for a given patient, the same threshold will be set whether the theoretical definition of HDA or a clinically useful one is used. With both definitions (HDA and clinical), harm is related to a supposed mechanism, thereby avoiding a purely normative approach, which can lead to overmedicalization. Thus, the idea of a symptom threshold is to prevent overmedicalization of certain conditions (e.g., undesirable states, social judgments, evaluative norms) (Wakefield 1992). On a scientific level, this conception echoes the challenge of precision medicine, which aims to identify mechanisms more reliably at the level of manifestations than at the level of disorders in general (Fernandes et al. 2017; Gómez-Carrillo et al. 2018). Indeed, from the point of view of data and models, scientific programs (especially those in computational psychiatry) are moving towards the identification of mechanisms at the level of symptoms (e.g., mechanism of hallucination) rather than at that of disorders (e.g., mechanism of schizophrenia).

Finally, from a psychometrical perspective, the clinically useful conception fits with the refinement of content validity, which has been largely set aside since the 1955 article by Cronbach and Meehl (Cronbach and Meehl 1955) or Messick’s article in 1968 (Messick 1989). As such, it is an answer to Sadowsky who wrote that “the DSM symptom menu forms a constellation of signs of suffering that have been grouped together in disparate contexts” (Sadowsky 2021).

1. **Activation of psychiatric kinds**

Psychiatric disorders may be seen as the central core of a cluster of stable and strongly connected harmful symptoms, underpinned by a set of mechanistic activations (Casali 2021). Psychiatric kinds consist of symptoms *strongly* set in a stable state and requiring a therapeutic intervention to restore the previous state (in reference to the notion of hysteresis) (Borsboom et al. 2019). In this conception, a symptom (which is not systematically harmful) is considered pathological when it results from an activated mechanism, which in turn leads the network (along with others – (Mackie 1965)) to freeze. Take the example of a persecution delirium strongly connected to an acoustic-verbal hallucination. Aggravation of the former will have a strong influence on the latter, because the two symptoms are strongly connected. Such a connection can be described in three ways: a) in semiologic terms, i.e. the most literal, the persecution delirium will be considered to increase the intuitive or interpretative mechanism of its strongly connected hallucination, for instance due to erroneous interpretations and beliefs; b) in axiological terms, the harm related to delirium will greatly aggravate the harm related to hallucination; c) in a mechanistic view, the mechanism related to the delirium will strongly reinforce (“activate”) the mechanism causing the hallucination.

Conversely, the absence of a psychiatric disorder is defined as the stable state of a *weakly* connected network. Weak connections allow network nodes to adapt to changes. Nodes in the network testify to mechanisms that are inactivated *per se* or are activated below a threshold. Thus, networks can adapt and change flexibly in response to environmental disturbances (Herrman et al. 2011). In other words, a weakly connected network is sufficiently robust in the sense that the robustness allows adaptation to disturbances (Kitano 2004; Lesne 2008, p. 200; Stelling et al. 2004).

A psychiatric disorder therefore corresponds to a difficulty of adaptation or control during which the frozen network no longer manages to adapt to the disturbances without a therapeutic intervention. For example, in the context of pedophilia or antisocial disorders, individuals experience harm owing to a lack of control of their symptoms, especially because their feelings strongly influence each other (e.g., (Gerin et al. 2017)). Conversely, the flexible common adaptation of symptoms can facilitate the individual’s social adaptation, even if she/he suffers from harms related to activated mechanisms of cognitive and emotional control that are clinically expressed. As all these events remain below the activation threshold, networks can adapt.

1. **Proposed definition of a psychiatric kind**

On the basis of the deconstruction of these two postulates and of this last analysis, we propose that a clinically useful psychiatric kind may be seen as “a set of clinical manifestations belonging to the prototypical central core of psychiatry, organized in a causal network, involving at least one harm, and whose set of harms exceeds the traditional threshold of significance for a given psychiatric disorder, directly related to an activated (often physiological) mechanism. Such a network composed of clinical manifestations is frozen in a state that is unable to allow the deactivation of these manifestations spontaneously without a therapeutic intervention”. In brief, to summarize this issue pedagogically, in a way useful to clinical practice, psychiatric disorders can be viewed as sets of interconnected symptoms that are harmful to the patient. Hence, they can only be modified by treatment. Indeed, these interconnected symptoms can be deactivated only by external forces, which refers to the notion of hysteresis (Borsboom et al. 2019).

We provide five examples below of the use of this definition for the clinician. First, a patient presents with a pattern of restrictive eating behaviors (fasting and physical hyperactivity). This leads to undernutrition, which results in hyperphagia (excessive food intake), leading to feelings of guilt because of fear of gaining weight, which leads to a return to restrictive eating habits. In this symptom network, the nodes are constituted by the symptoms, while the connections between the nodes represent the causal interactions between the symptoms. This first example illustrates the notion of causal dependence between the symptoms proposed in the definition above.

Second, a seemingly healthy female patient presents to the psychiatrist upon the advice of her family to seek help for acoustic-verbal hallucinations, which result in feelings of being watched. However, this feeling does not harm the patient and does not affect strangers or people close to her. She also considers this condition to be a form of neuroatypicality. The patient has no psychiatric disorder according to the previous practical definition, because the threshold of harm remains low, despite the mutual influence of symptoms she experiences, and the possibility of treating this set of symptoms as an incipient psychosis. This second example illustrates the notion of a harm threshold described in the definition above.

Third, a male patient describes a job-related issue that has no financial impact on his daily life or that of his family, given his high resources and savings. However, he describes significant professional harm due to perceived injustice related to a job position that he felt he deserved but that was given to his colleague. In this instance, what the patient reports has no impact on either his mood (anxiety levels) or his daily life. This third example illustrates the notion of harm directly related to clinical manifestations proposed in the definition above.

Fourth, a patient with a craving, symptom which constitutes an important clinical manifestation among other symptoms of substance use (Drummond et al. 2000; Sayette 2016; Serre et al. 2018; Vafaie and Kober 2022), can be considered a clinical manifestation. Craving is underpinned by a known physiopathological mechanism and is related to dysfunction of the dopamine and serotonin circuits at the level of the para-hippocampal and fusiform gyri, putamen, anterior cingulate cortex, amygdala, and orbitofrontal cortex (Robinson and Berridge 1993). If craving was the sole manifestation of behaviors in daily life, without being linked to cerebral mechanisms, it could not be integrated into the previous definition. This fourth example illustrates the importance of considering clinical manifestations as directly related to an activated physiological mechanism.

Fifth, in the context of a patient with post-traumatic stress symptoms, β-adrenergic blockade treatment during sessions of eye movement desensitization and reprocessing (EMDR), the preferred treatment for these symptoms, may prove effective in combating unrelenting nightmares that the patient experiences as a result of intense trauma (Argolo et al. 2015). This fifth example illustrates the importance of treating on frozen states of activated symptoms.

Finally, to better understand the full definition, we consider a set of symptoms referring to major depressive disorders. In this example, a young adolescent patient seeks help for a loss of energy that has evolved for six months and has led to an anhedonia, which has resulted in appetite loss and significant weight loss. These symptoms have aggravated her insomnia and participated in the appearance of recurrent thoughts of death for three weeks. Notably, no actions taken by herself, or her parents have ameliorated the harm that underlies this set of symptoms. Each symptom is potentially correlated with an underlying physiological mechanism, and they influence each other. These mutual effects exceed the threshold of tolerance for this adolescent, who thus requires a therapeutic force to be treated.

**Conclusion**

To provide a definition of ‘useful psychiatric kind’ for the clinician, we first examined two definitions that are relevant for different purposes: a theoretical definition serving as an epistemic hub for different non-clinical functions (conservative with regard to the HDA); and a practical definition useful for diagnostic, prognostic, preventive and therapeutic management, which is specifically analyzed in this article.

The existence of a network of harmful symptoms, underpinned by activated mechanisms, seems relevant and intuitive to the clinician. We have shown that an intuitive definition that indirectly makes it possible to predict the evolution (or the non-evolution) of a disorder and to reorient, advise and support patients, prevent high risk, and rule out differential diagnoses is more useful than the presumption of a latent variable, evolutionary assumptions or clinically invisible conceptual harms.

To define such a clinically useful psychiatric kind, we have identified two postulates relating to dysfunction and harm. We argue that the identification of a central core makes it possible to target the field of psychiatry, and we specify the criteria (mechanistic and related to harms) that the symptoms of such a core should include. These developments then served to propose a general value-laden definition of psychiatric disorder that is useful to the clinician (Cooper 2005; De Vreese 2017; Reznek 1987), and which allows a non-hierarchical pluralism to be invoked (unlike the duality of Boorse (Boorse 1977), in which the practical conception is subordinate to the theoretical conception).

As amply demonstrated thus far in this article, we support the idea that a first definition based on HDA acts as an epistemic hub. We also believe that a second definition, such as the one presented above, is specifically needed for the clinician. Our plea here is for a bipartite definitional pluralism, with an intuitive pragmatic definition that can be used in clinical practice, and a theoretical one based on the HDA definition. The notion of harm connects these two definitions, providing a single threshold (although it is present at different levels of understanding in the two analyses).

Therefore, we defend the normative harm-based definition of psychiatry that requires a naturalistic component due to the existence of activated mechanisms that justify the existence of a central core of harmful symptoms. This is tantamount to a hybrid approach, but it differs from the Wakefield analysis approach and from the view of Jonathan Tsou (Tsou 2021b, 2022) for a number of reasons. First, we propose an explicit and original conception based on harmful symptoms. Second, the starting point of our account is based on the abandonment of latent variables (Haslam 2002; Kendler et al. 2011; Tabb 2019; Zachar 2015). Third, we distinguished a potentially useful definition for non-clinical issues, that fits the DSM categories (Tabb 2019), from a definition that is useful for clinical practice. Fourth, following the new mechanical philosophy (Glennan 2017), we propose an analysis of psychiatric disorders through a mechanistic prism and apply the notion of activation (and not dysfunction). Moreover, following Wakefield, we suggest that this mechanism does not need to be empirically identified in a stable way. Fifth, the description of a central core allows us to coalesce scattered philosophical notions. Sixth, we are more interested in a set of interacting symptoms than in an operationalizable construct, interest which shifts the debate of validity to the debate of content. Seventh, we offer a new operational definition of psychiatric disorders for clinicians.

Thus, we contend that a psychiatric disorder may be considered present in clinical practice when a patient’s network of harmful symptoms is frozen in an unchanging conformation, is inflexible and unresponsive to changes in the environment, and requires therapeutic intervention to be modified. Psychiatric disorders thus correspond to networks of strongly connected symptoms whose sets of harms reach a threshold of significance, and which are underpinned by networks of activated mechanisms that drive these harmful symptoms.

**Acknowledgments**

I am grateful to the three anonymous referees for helpful comments and discussion. I also warmly thank Prof. Elodie Giroux and Prof. Denis Forest for their inspiration, Prof. Jean-Arthur Micoulaud-Franchi for his wise advice, and Dr. Anne-Marie Gagné-Julien for her trust and support.

**References**

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Amundson, R., & Lauder, G. V. (1994). Function without purpose. *Biology and Philosophy*, *9*(4), 443–469. https://doi.org/10.1007/BF00850375

Argolo, F. C., Cavalcanti-Ribeiro, P., Netto, L. R., & Quarantini, L. C. (2015). Prevention of posttraumatic stress disorder with propranolol: A meta-analytic review. *Journal of Psychosomatic Research*, *79*(2), 89–93. https://doi.org/10.1016/j.jpsychores.2015.04.006

Beard, C., Millner, A. J., Forgeard, M. J. C., Fried, E. I., Hsu, K. J., Treadway, M. T., et al. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, *46*(16), 3359–3369. https://doi.org/10.1017/S0033291716002300

Beck, A. T. (1993). Cognitive therapy: Past, present, and future. *Journal of Consulting and Clinical Psychology*, *61*(2), 194–198. https://doi.org/10.1037/0022-006X.61.2.194

Bertalanffy, L. von. (2015). *General System Theory: Foundations, Development, Applications*. George Braziller, Incorporated.

Bhugra, D., Easter, A., Mallaris, Y., & Gupta, S. (2011). Clinical decision making in psychiatry by psychiatrists. *Acta Psychiatrica Scandinavica*, *124*(5), 403–411. https://doi.org/10.1111/j.1600-0447.2011.01737.x

Boorse, C. (1977). Health as a Theoretical Concept. *Philosophy of Science*, *44*(4), 542–573. https://doi.org/10.1086/288768

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*(9), 1089–1108. https://doi.org/10.1002/jclp.20503

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., Cramer, A. O. J., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, *42*. https://doi.org/10.1017/S0140525X17002266

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., et al. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L. J., & Cramer, A. O. J. (2017). False alarm? A comprehensive reanalysis of “Evidence that psychopathology symptom networks have limited replicability” by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, *126*(7), 989–999. https://doi.org/10.1037/abn0000306

Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *61*(1/2), 127–148.

Boyd, R. (1999). Homeostasis, Species, and Higher Taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141–85). MIT Press.

Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present, and future. *Current Opinion in Psychology*, *41*, 59–64. https://doi.org/10.1016/j.copsyc.2021.03.004

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., et al. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, *128*(8), 892–903. https://doi.org/10.1037/abn0000446

Brülde, B., & Radovic, F. (2006). What is Mental about Mental Disorder? *Philosophy, Psychiatry, & Psychology*, *13*(2), 99–116. https://doi.org/10.1353/ppp.2007.0001

Casali, M. (2021). Exploring the boundaries and ontology of Psychiatric Disorders (PDs) using the Homeostatic Property Cluster (HPC) model: *Lato Sensu: Revue de la Société de philosophie des sciences*, *8*(2), 15–31. https://doi.org/10.20416/LSRSPS.V8I2.3

Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D., et al. (2016). Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. *The American Journal of Psychiatry*, *173*(4), 373–384. https://doi.org/10.1176/appi.ajp.2015.14091200

Clementz, B. A., Trotti, R. L., Pearlson, G. D., Keshavan, M. S., Gershon, E. S., Keedy, S. K., et al. (2020). Testing Psychosis Phenotypes From Bipolar-Schizophrenia Network for Intermediate Phenotypes for Clinical Application: Biotype Characteristics and Targets. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *5*(8), 808–818. https://doi.org/10.1016/j.bpsc.2020.03.011

Cohen, A. S., Schwartz, E., Le, T., Cowan, T., Cox, C., Tucker, R., et al. (2020). Validating digital phenotyping technologies for clinical use: the critical importance of “resolution.” *World Psychiatry*, *19*(1), 114–115. https://doi.org/10.1002/wps.20703

Cooper. (2002). Disease. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *33*(2), 263–282. https://doi.org/10.1016/S0039-3681(02)00018-3

Cooper. (2005). *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Springer Netherlands. https://doi.org/10.1007/1-4020-3345-1

Cooper. (2012). *Is psychiatric classification a good thing?* *Philosophical Issues in Psychiatry II*. Oxford University Press. https://oxfordmedicine.com/view/10.1093/med/9780199642205.001.0001/med-9780199642205-chapter-004001. Accessed 2 September 2021

Cooper. (2013a). Avoiding false positives: zones of rarity, the threshold problem, and the DSM clinical significance criterion. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, *58*(11), 606–611. https://doi.org/10.1177/070674371305801105

Cooper. (2013b). Natural kinds. In K. Fulford, M. Davies, G. Graham, J. Sadler, G. Stanghellini, & T. Thornton (Eds.), . Oxford: Oxford University Press. https://eprints.lancs.ac.uk/id/eprint/50918/. Accessed 11 November 2021

Cooper. (2014). *Psychiatry and Philosophy of Science*. Routledge.

Cooper. (2020). The Concept of Disorder Revisited: Robustly Value-Laden Despite Change. *Aristotelian Society Supplementary Volume*, *94*(1), 141–161. https://doi.org/10.1093/arisup/akaa010

Cooper, R. (2021). On Harm. In *Defining Mental Disorder Jerome Wakefield and His Critics*. https://direct.mit.edu/books/book/5015/chapter/2812075/On-Harm. Accessed 21 July 2022

Cramer, A. O. J., Borkulo, C. D. van, Giltay, E. J., Maas, H. L. J. van der, Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major Depression as a Complex Dynamic System. *PLOS ONE*, *11*(12), e0167490. https://doi.org/10.1371/journal.pone.0167490

Craver, C. F. (2001). Role Functions, Mechanisms, and Hierarchy. *Philosophy of Science*, *68*(1), 53–74. https://doi.org/10.1086/392866

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Cucuringu, M., Rombach, P., Lee, S. H., & Porter, M. A. (2016). Detection of core–periphery structure in networks using spectral methods and geodesic paths. *European Journal of Applied Mathematics*, *27*(6), 846–887. https://doi.org/10.1017/S095679251600022X

Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, *72*(November), 741–64. https://doi.org/10.2307/2024640

Cuthbert. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World psychiatry: official journal of the World Psychiatric Association (WPA)*, *13*(1), 28–35. https://doi.org/10.1002/wps.20087

Darden, L. (2008). Thinking Again about Biological Mechanisms. *Philosophy of Science*, *75*(5), 958–969. https://doi.org/10.1086/594538

De Block, A., & Sholl, J. (2021). Harmless Dysfunctions and the Problem of Normal Variation. In *Defining Mental Disorder: Jerome Wakefield and His Critics*. https://direct.mit.edu/books/book/5015/chapter/2812073/Harmless-Dysfunctions-and-the-Problem-of-Normal. Accessed 23 July 2022

De Vreese. (2021). Against the Disorder/Nondisorder Dichotomy. In *Defining Mental Disorder Jerome Wakefield and His Critics*. https://direct.mit.edu/books/book/5015/chapter/2812050/Against-the-Disorder-Nondisorder-Dichotomy. Accessed 21 July 2022

De Vreese, L. (2017). How to Proceed in the Disease Concept Debate? A Pragmatic Approach. *The Journal of Medicine and Philosophy*, *42*(4), 424–446. https://doi.org/10.1093/jmp/jhx011

Demazeux, S. (2019). *L’Éclipse du symptôme. L’observation clinique en psychiatrie, 1800-1950.* (Ithaque.).

Drummond, D. C., Litten, R. Z., Lowman, C., & Hunt, W. A. (2000). Craving research: Future directions. *Addiction*, *95*(Suppl2), S247–S255. https://doi.org/10.1080/09652140050111816

Dupré, J. (2002). *Humans and Other Animals*. https://doi.org/10.3366/edinburgh/9780748639403.003.0006

Eaton, W. W., Regier, D. A., Locke, B. Z., & Taube, C. A. (1981). The Epidemiologic Catchment Area Program of the National Institute of Mental Health. *Public Health Reports*, *96*(4), 319–325.

Engelhardt, H. T. (1975). The Concepts of Health and Disease. In H. T. Engelhardt & S. F. Spicker (Eds.), *Evaluation and Explanation in the Biomedical Sciences: Proceedings of the First Trans-Disciplinary Symposium on Philosophy and Medicine Held at Galveston, May 9–11, 1974* (pp. 125–141). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-1769-5\_9

Engelhardt, H. T. (1981). Clinical Judgment. *Metamedicine*, *2*(3), 301–317. https://doi.org/10.1007/BF00882077

Engelhardt, H. T. (1984). Clinical Problems and the Concept of Disease. In L. Nordenfelt & B. I. B. Lindahl (Eds.), *Health, Disease, and Causal Explanations in Medicine* (pp. 27–41). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-6283-5\_5

Ereshefsky, M. (2009). Defining ‘health’ and ‘disease.’ *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *40*(3), 221–227. https://doi.org/10.1016/j.shpsc.2009.06.005

Faucher, L., & Forest, D. (Eds.). (2021). *Defining Mental Disorder: Jerome Wakefield and His Critics* (1st edition.). Cambridge, Massachusetts: The MIT Press.

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of ‘precision psychiatry.’ *BMC Medicine*, *15*(1), 80. https://doi.org/10.1186/s12916-017-0849-x

Ferrante, M., Redish, A. D., Oquendo, M. A., Averbeck, B. B., Kinnane, M. E., & Gordon, J. A. (2019). Computational psychiatry: a report from the 2017 NIMH workshop on opportunities and challenges. *Molecular Psychiatry*, *24*(4), 479–483. https://doi.org/10.1038/s41380-018-0063-z

Frances, A. (2013). *Essentials of Psychiatric Diagnosis: Responding to the Challenge of DSM-5*. Guilford Publications.

Fried. (2015). Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00309

Fried, E. I., von Stockert, S., Haslbeck, J. M. B., Lamers, F., Schoevers, R. A., & Penninx, B. W. J. H. (2020). Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. *Psychological Medicine*, *50*(16), 2682–2690. https://doi.org/10.1017/S0033291719002770

Fried, Eiko I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are “good” depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Fried, Eiko I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017a). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Fried, Eiko I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017b). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Fusar-Poli, P., Hijazi, Z., Stahl, D., & Steyerberg, E. W. (2018). The Science of Prognosis in Psychiatry: A Review. *JAMA psychiatry*, *75*(12), 1289–1297. https://doi.org/10.1001/jamapsychiatry.2018.2530

Gagné-Julien, A.-M. (2020). Towards a socially constructed and objective concept of mental disorder. *Synthese*. https://doi.org/10.1007/s11229-020-02647-7

Gerin, M. I., Puetz, V. B., Blair, R. J. R., White, S., Sethi, A., Hoffmann, F., et al. (2017). A neurocomputational investigation of reinforcement-based decision making as a candidate latent vulnerability mechanism in maltreated children. *Development and Psychopathology*, *29*(5), 1689–1705. https://doi.org/10.1017/S095457941700133X

Glennan, S. (2017). *The New Mechanical Philosophy*. Oxford University Press.

Gómez-Carrillo, A., Langlois-Thérien, T., & Kirmayer, L. J. (2018). Precision Psychiatry—Yes, but Precisely What? *JAMA Psychiatry*, *75*(12), 1302–1303. https://doi.org/10.1001/jamapsychiatry.2018.2651

Gross, F. (2011). What Systems Biology Can Tell Us about Disease. *History and philosophy of the life sciences*, *33*, 477–96.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62. https://doi.org/10.1136/jnnp.23.1.56

Haslam, N. (2002). Kinds of Kinds: A Conceptual Taxonomy of Psychiatric Categories. *Philosophy, Psychiatry, & Psychology*, *9*(3), 203–217. https://doi.org/10.1353/ppp.2003.0043

Hauswald, R. (2016). The Ontology of Interactive Kinds. *Journal of Social Ontology*, *2*(2), 203–221. https://doi.org/10.1515/jso-2015-0049

Herrman, H., Stewart, D. E., Diaz-Granados, N., Berger, E. L., Jackson, B., & Yuen, T. (2011). What is Resilience? *The Canadian Journal of Psychiatry*, *56*(5), 258–265. https://doi.org/10.1177/070674371105600504

Hesslow, G. (1993). Do we need a concept of disease? *Theoretical Medicine*, *14*(1), 1–14. https://doi.org/10.1007/BF00993984

Hitchcock, P., Fried, E. I., & Frank, M. (2021, April 1). Computational Psychiatry Needs Time and Context. PsyArXiv. https://doi.org/10.31234/osf.io/xf2u3

Illari, P., & Williamson, J. (2011). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*, 119–135. https://doi.org/10.1007/s13194-011-0038-2

Ivleva, E. I., Clementz, B. A., Dutcher, A. M., Arnold, S. J. M., Jeon-Slaughter, H., Aslan, S., et al. (2017). Brain Structure Biomarkers in the Psychosis Biotypes: Findings From the Bipolar-Schizophrenia Network for Intermediate Phenotypes. *Biological Psychiatry*, *82*(1), 26–39. https://doi.org/10.1016/j.biopsych.2016.08.030

Jones, P. J., Ma, R., & McNally, R. J. (2019). Bridge Centrality: A Network Approach to Understanding Comorbidity. *Multivariate Behavioral Research*, *0*(0), 1–15. https://doi.org/10.1080/00273171.2019.1614898

Jones, P. J., & Robinaugh, D. (2021, June 18). So What? Implications of Network Theory for Research and Practice. PsyArXiv. https://doi.org/10.31234/osf.io/6bqdx

Kendler, K. S. (2008). Explanatory models for psychiatric illness. *The American Journal of Psychiatry*, *165*(6), 695–702. https://doi.org/10.1176/appi.ajp.2008.07071061

Kendler, Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine*, *41*(6), 1143–1150. https://doi.org/10.1017/S0033291710001844

Keuck, L. (2011). How to Make Sense of Broadly Applied Medical Classification Systems: Introducing Epistemic Hubs. *History and philosophy of the life sciences*, *33*, 583–601.

Kim, N., & Ahn, W. (2002). Clinical psychologists’ theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology General*, *131*, 451–476. https://doi.org/10.1037//0096-3445.131.4.451

Kincaid, H. (2008). Do We Need Theory to Study Disease? Lessons from cancer research and their implications for mental illness. *Perspectives in biology and medicine*, *51*, 367–78. https://doi.org/10.1353/pbm.0.0019

Kirmayer, L. J., Lemelson, R., & Cummings, C. A. (2015). *Re-Visioning Psychiatry: Cultural Phenomenology, Critical Neuroscience, and Global Mental Health*. Cambridge University Press.

Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, *5*(11), 826–837. https://doi.org/10.1038/nrg1471

Kleinman, A. (1988). *The illness narratives: Suffering, healing, and the human condition* (pp. xviii, 284). New York, NY, US: Basic Books.

Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., et al. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of abnormal psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

Krueger, J. (2015). Theoretical Health and Medical Practice. *Philosophy of Science*, *82*(3), 491–508. https://doi.org/10.1086/681628

Lahey, B. B., Flagg, E. W., Bird, H. R., Schwab-stone, M. E., Canino, G., Dulcan, M. K., et al. (1996). The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: Background and Methodology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(7), 855–864. https://doi.org/10.1097/00004583-199607000-00011

Lesne, A. (2008). Robustness: confronting lessons from physics and biology. *Biological Reviews*, *83*(4), 509–532. https://doi.org/10.1111/j.1469-185X.2008.00052.x

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield’s “harmful dysfunction” analysis. *Journal of Abnormal Psychology*, *104*(3), 411–420. https://doi.org/10.1037/0021-843X.104.3.411

Lux, V., & Kendler, K. S. (2010). Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychological Medicine*, *40*(10), 1679–1690. https://doi.org/10.1017/S0033291709992157

Ma, A., & Mondragón, R. J. (2015). Rich-Cores in Networks. *PLOS ONE*, *10*(3), e0119678. https://doi.org/10.1371/journal.pone.0119678

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, *67*(1), 1–25. https://doi.org/10.1086/392759

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–264.

Marková, I. S., & Berrios, G. E. (2012). Epistemology of Psychiatry. *Psychopathology*, *45*(4), 220–227. https://doi.org/10.1159/000331599

McGorry, Keshavan, M., Goldstone, S., Amminger, P., Allott, K., Berk, M., et al. (2014). Biomarkers and clinical staging in psychiatry. *World Psychiatry*, *13*(3), 211–223. https://doi.org/10.1002/wps.20144

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, *86*, 95–104. https://doi.org/10.1016/j.brat.2016.06.006

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In *Similarity and analogical reasoning* (pp. 179–195). New York, NY, US: Cambridge University Press. https://doi.org/10.1017/CBO9780511529863.009

Meehl, P. (1967). What Can the Clinician Do Well? *Problems in human assessment*, pp. 594–599.

Messick, S. (1989). Validity. In *Educational measurement, 3rd ed* (pp. 13–103). American Council on Education.

Moghaddam-Taaheri, S. (2011). Understanding pathology in the context of physiological mechanisms: the practicality of a broken-normal view. *Biology & Philosophy*, *26*(4), 603. https://doi.org/10.1007/s10539-011-9258-2

Murphy, D. (2006). *Psychiatry in the Scientific Image*. Cambridge, MA, USA: A Bradford Book.

Nelson, B., McGorry, P., & Fernandez, F. (2020). Integrating Clinical Staging and Phenomenological Psychopathology to Add Depth, Nuance, and Utility to Clinical Phenotyping: A Heuristic Challenge. Accessed 25 August 2021

Nervi, M. (2010). Mechanisms, malfunctions and explanation in medicine. *Biology & Philosophy*, *25*(2), 215–228. https://doi.org/10.1007/s10539-009-9190-x

Nordenfelt, L. (2006). The concepts of health and illness revisited. *Medicine, Health Care and Philosophy*, *10*(1), 5. https://doi.org/10.1007/s11019-006-9017-3

Oliver, M. (1996). *Understanding disability: From theory to practice* (pp. x, 192). New York, NY, US: St Martin’s Press. https://doi.org/10.1007/978-1-349-24269-6

Powell, R., & Scarffe, E. (2019a). “Rethinking ‘Disease’: a fresh diagnosis and a new philosophical treatment.” *Journal of Medical Ethics*, *45*(9), 579–588. https://doi.org/10.1136/medethics-2019-105465

Powell, R., & Scarffe, E. (2019b). Rehabilitating “Disease”: Function, Value, and Objectivity in Medicine. *Philosophy of Science*, *86*(5), 1168–1178. https://doi.org/10.1086/705520

Putnam, H. (1975). Philosophy and Our Mental Life. In *Mind, Language, and Reality*. Cambridge University Press.

Reznek, L. (1987). *The Nature of Disease*. Routledge & Kegan Paul.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research. Brain Research Reviews*, *18*(3), 247–291. https://doi.org/10.1016/0165-0173(93)90013-p

Rombach, M. P., Porter, M. A., Fowler, J. H., & Mucha, P. J. (2014). Core-Periphery Structure in Networks. *SIAM Journal on Applied Mathematics*, *74*(1), 167–190. https://doi.org/10.1137/120881683

Sadegh-Zadeh, K. (2001). The Fuzzy Revolution: Goodbye to the Aristotelian Weltanschauung. *Artificial Intelligence in Medicine*, *21*(1–3), 1–25. https://doi.org/10.1016/S0933-3657(00)00071-3

Sadegh-Zadeh, K. (2008). The Prototype Resemblance Theory of Disease. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *33*(2), 106–139. https://doi.org/10.1093/jmp/jhn004

Sadowsky, J. (2021). Before and After Prozac: Psychiatry as Medicine, and the Historiography of Depression. *Culture, Medicine, and Psychiatry*, *45*(3), 479–502. https://doi.org/10.1007/s11013-021-09729-2

Sartorius, N., & World Health Organization (Eds.). (1983). *Depressive disorders in different cultures: report on the WHO collaborative study on standardized assessment of depressive disorders*. Geneva : [Albany, N.Y: World Health Organization ; available from WHO Publications Centre].

Sass, L., Pienkos, E., Skodlar, B., Stanghellini, G., Fuchs, T., Parnas, J., & Jones, N. (2017). EAWE: Examination of Anomalous World Experience. *Psychopathology*, *50*, 10–54. https://doi.org/10.1159/000454928

Sayette, M. A. (2016). The Role of Craving in Substance Use Disorders: Theoretical and Methodological Issues. *Annual Review of Clinical Psychology*, *12*, 407–433. https://doi.org/10.1146/annurev-clinpsy-021815-093351

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Serre, F., Fatseas, M., Denis, C., Swendsen, J., & Auriacombe, M. (2018). Predictors of craving and substance use among patients with alcohol, tobacco, cannabis or opiate addictions: Commonalities and specificities across substances. *Addictive Behaviors*, *83*, 123–129. https://doi.org/10.1016/j.addbeh.2018.01.041

Shah, J. L., Scott, J., McGorry, P. D., Cross, S. P. M., Keshavan, M. S., Nelson, B., et al. (2020). Transdiagnostic clinical staging in youth mental health: a first international consensus statement. *World Psychiatry*, *19*(2), 233–242. https://doi.org/10.1002/wps.20745

Shapiro, S., Skinner, E. A., Kessler, L. G., Von Korff, M., German, P. S., Tischler, G. L., et al. (1984). Utilization of health and mental health services. Three Epidemiologic Catchment Area sites. *Archives of General Psychiatry*, *41*(10), 971–978. https://doi.org/10.1001/archpsyc.1984.01790210053007

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, *24*(5), 399–411. https://doi.org/10.1016/0010-440X(83)90032-9

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

Spitzer, & Wakefield. (1999). DSM-IV diagnostic criterion for clinical significance: does it help solve the false positives problem? *The American Journal of Psychiatry*, *156*(12). https://doi.org/10.1176/ajp.156.12.1856

Stein, D. J. (2021). *Problems of Living: Perspectives from Philosophy, Psychiatry, and Cognitive-Affective Science* (1st edition.). San Diego: Academic Press.

Stelling, J., Sauer, U., Szallasi, Z., Doyle, F., & Doyle, J. (2004). Robstness of Cellular Functions. *Cell*, *118*, 675–85. https://doi.org/10.1016/j.cell.2004.09.008

Szasz, T. S. (1960). The myth of mental illness. *American Psychologist*, *15*(2), 113–118. https://doi.org/10.1037/h0046535

Tabb, K. (2019). Philosophy of psychiatry after diagnostic kinds. *Synthese*, *196*. https://doi.org/10.1007/s11229-017-1659-6

Tekin, Ş. (2016). Are Mental Disorders Natural Kinds?: A Plea for a New Approach to Intervention in Psychiatry. *Philosophy, Psychiatry, & Psychology*, *23*(2), 147–163. https://doi.org/10.1353/ppp.2016.0013

Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again* (Illustrated edition.). New York: Basic Books.

Tsou. (2007). Hacking on the Looping Effects of Psychiatric Classifications: What Is an Interactive and Indifferent Kind? *International Studies in the Philosophy of Science*, *21*(3), 329–344. https://doi.org/10.1080/02698590701589601

Tsou. (2015). DSM-5 and Psychiatry’s Second Revolution: Descriptive Vs. Theoretical Approaches to Psychiatric Classification. In Steeves Demazeux & P. Singy (Eds.), *The DSM-5 in Perspective: Philosophical Reflections on the Psychiatric Babel* (pp. 43–62). Springer.

Tsou. (2016). Natural kinds, psychiatric classification and the history of the DSM. *History of Psychiatry*, *27*. https://doi.org/10.1177/0957154X16656580

Tsou, J. Y. (2021a). *Philosophy of Psychiatry*. S.l.: Cambridge University Press.

Tsou, J. Y. (2021b). *Philosophy of Psychiatry*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108588485

Tsou, J. Y. (2022). Biological Essentialism, Projectable Human Kinds, and Psychiatric Classification. *Philosophy of Science*, 1–21. https://doi.org/10.1017/psa.2022.64

Vafaie, N., & Kober, H. (2022). Association of Drug Cues and Craving With Drug Use and Relapse: A Systematic Review and Meta-analysis. *JAMA psychiatry*, *79*(7), 641–650. https://doi.org/10.1001/jamapsychiatry.2022.1240

van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of Symptom Network Structure With the Course of [corrected] Depression. *JAMA psychiatry*, *72*(12), 1219–1226. https://doi.org/10.1001/jamapsychiatry.2015.2079

Wachbroit, R. (1994). Normality as a Biological Concept. *Philosophy of Science*, *61*(4), 579–591.

Wakefield. (1992). The concept of mental disorder. On the boundary between biological facts and social values. *The American Psychologist*, *47*(3), 373–388. https://doi.org/10.1037//0003-066x.47.3.373

Wakefield. (1999). Mental disorder as a black box essentialist concept. *Journal of Abnormal Psychology*, *108*(3), 465–472. https://doi.org/10.1037/0021-843X.108.3.465

Wakefield, J. C. (2006). Are there relational disorders? A harmful dysfunction perspective: Comment on the special section. *Journal of Family Psychology*, *20*(3), 423–427. https://doi.org/10.1037/0893-3200.20.3.423

Wakefield, J. C. (2012). Are you as smart as a 4th grader? Why the prototype-similarity approach to diagnosis is a step backward for a scientific psychiatry. *World Psychiatry*, *11*(1), 27–28.

Wakefield, J. C. (2013). The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review*, *33*(7), 825–845. https://doi.org/10.1016/j.cpr.2013.03.007

Wakefield, J. C., & Conrad, J. A. (2019). Does the harm component of the harmful dysfunction analysis need rethinking?: Reply to Powell and Scarffe. *Journal of Medical Ethics*, *45*(9), 594–596. https://doi.org/10.1136/medethics-2019-105578

Wakefield, J. C., & First, M. B. (2013). Diagnostic Validity and the Definition of Mental Disorder: A Program for Conceptually Advancing Psychiatry. *The Canadian Journal of Psychiatry*, *58*(12), 653–655. https://doi.org/10.1177/070674371305801201

Wakefield, J. C., & Spitzer, R. L. (2002). *Why requiring clinical significance does not solve epidemiology’s and DSM’s validity problem: Response to Regier and Narrow* (p. 40). Arlington, VA, US: American Psychiatric Publishing, Inc.

Westen, D. (2012). Prototype diagnosis of psychiatric syndromes. *World Psychiatry*, *11*(1), 16–21. https://doi.org/10.1016/j.wpsyc.2012.01.004

World Health Organization. (2017). International Classification of Diseases 11th Edition (ICD-11). In *Encyclopedia of Personality and Individual Differences* (pp. 1–6). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8\_904-1

Zachar, P. (2000). Psychiatric Disorders Are Not Natural Kinds. *Philosophy, Psychiatry, and Psychology*, *7*(3), 167–182.

Zachar, P. (2014). *A Metaphysics of Psychopathology*. MIT Press.

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

Zachar, P. (2021). Psychiatric Disorders and the Imperfect Community: A Nominalist HDA. In *Defining Mental Disorder Jerome Wakefield and His Critics*. https://direct.mit.edu/books/book/5015/chapter/2812055/Psychiatric-Disorders-and-the-Imperfect-Community. Accessed 21 July 2022

1. This would not preclude the existence of an underlying physiological cause, but at least it supposes the possibility of a “spectrum of essentiality” (Hitchcock et al. 2021) in which a cause could be identified for some disorders such as schizophrenia, while in others such as depression, it is the interaction of symptoms that is more relevant in naming the disorder. [↑](#footnote-ref-1)
2. However, we do not retain the definition of *dysfunctional* mechanisms of (Moghaddam-Taaheri 2011) or (Darden 2008). Of course, any dysfunction can account for a clinical manifestation. In this case, the mechanism would be based on a concept of role-function according to the account of Cummins (Cummins 1975), which conceptualizes “functions as causal contributions of a component part to a capacity of a larger system”, or according to a “contextual system” as posited by Craver (Craver 2001). However, we criticize this ‘broken-normal’ view and do not retain the notion of “failing function” (Nervi 2010) owing to the lack of necessary and sufficient criteria (Krueger 2015). We prefer to consider the potential robustness of a mechanism (see fourth part). [↑](#footnote-ref-2)
3. This example reinforces the idea of using a symptom-based approach, while debate in the philosophy of psychiatry is complicated by the analysis of the category of autism (diagnostic kind) in its entirety. [↑](#footnote-ref-3)
4. The weighting depends on the importance within the network. It refers especially to centrality and can be statistically and computationally measurable – there are more than 70 methods for compute centrality in complex networks. [↑](#footnote-ref-4)