Behavioral and Brain Sciences (forthcoming)

Brain disorders? Not really…

Why network structures block reductionism in psychopathology research

Denny Borsboom

Angélique Cramer

University of Amsterdam

Annemarie Kalis

Utrecht University

Running head: Brain disorders? Not really…

Send correspondence to:

Denny Borsboom

Department of Psychology

University of Amsterdam

Nieuwe Achtergracht 129-B

1018 WT Amsterdam

d.borsboom@uva.nl

Word count:

Short Abstract: 176

Long Abstract: 264

Main text: 10320

Short Abstract

In the past decades, reductionism has dominated research directions and funding policies in clinical psychology and psychiatry. However, the intense search for the biological basis of mental disorders has not resulted in conclusive reductionist explanations of psychopathology. Recently, network models, in which mental disorders arise from the causal interplay between symptoms, have been proposed as an alternative framework. In this paper, we show that this framework can help understand why reductionist approaches in psychiatry and clinical psychology are on the wrong track. First, symptom networks preclude the identification of a common cause of symptomatology with a neurobiological condition, because in symptom networks there is no such common cause. Second, symptom network relations depend on the content of mental states and as such feature intentionality. Third, the strength of network relations is highly likely to partially depend on cultural and historical contexts as well as external mechanisms in the environment. This means that neither psychological nor biological levels can claim causal or explanatory priority, and that a holistic research strategy is necessary for progress in psychopathology research.

Long Abstract

In the past decades, reductionism has dominated both research directions and funding policies in clinical psychology and psychiatry. However, the intense search for the biological basis of mental disorders has not resulted in conclusive reductionist explanations of psychopathology. Recently, network models have been proposed as an alternative framework for the analysis of mental disorders, in which mental disorders arise from the causal interplay between symptoms. In this paper, we show that this conceptualization can help understand why reductionist approaches in psychiatry and clinical psychology are on the wrong track. First, symptom networks preclude the identification of a common cause of symptomatology with a neurobiological condition, because in symptom networks there is no such common cause. Second, symptom network relations depend on the content of mental states and as such feature intentionality. Third, the strength of network relations is highly likely to partially depend on cultural and historical contexts as well as external mechanisms in the environment. Taken together, these properties suggest that, if mental disorders are indeed networks of causally related symptoms, reductionist accounts cannot achieve the level of success associated with reductionist disease models in modern medicine. As an alternative strategy, we propose to interpret network structures in terms of D. C. Dennett's (1987) notion of *real patterns*, and suggest that, instead of being reducible to a biological basis, mental disorders feature biological and psychological factors that are deeply intertwined in feedback loops. This suggests that neither psychological nor biological levels can claim causal or explanatory priority, and that a holistic research strategy is necessary for progress in the study of mental disorders.

Introduction

Can mental disorders be conclusively explained in terms of neurobiology and genetic constitution? Yes, according to many researchers and laypeople alike. Probably not, as we will argue in the present paper.

Many believe that symptoms, signs, and other problems associated with mental disorders – e.g., depressed mood, psychomotor agitation – are caused by 'genes for mental disorders', neurobiological mechanisms, deficient brain circuits, and other biological factors. This firm belief in *explanatory reductionism* – that is, the belief that mental disorders can ultimately be explained in terms of specific dysfunctional neurobiological conditions – is partly due to the fact that the study of mental disorders traditionally belonged to the medical discipline (Andreasen, 1984; Greenberg, 2013; Guze, 1989; Kraepelin & Lange, 1927). Additionally, many laypeople and some researchers alike are convinced that a biological explanation of a mental disorder supports the notion that the disorder is 'real'; that a patient is not just sitting home feeling blue and tired, which feeds the stigma surrounding mental disorders (*Nature* editorial, Anonymous, 2013), but actually has a real disease for which one needs medication.

But what is the evidence for uniquely biological explanations of mental disorders? A sober evaluation of the research literature does not inspire much enthusiasm for explanatory reductionism (Bentall, 2003; Lacasse & Leo, 2005; van Os, 2009). For example, a 2013 *Nature* editorial concluded that "[d]espite decades of work, the genetic, metabolic and cellular signatures of almost all mental syndromes remain

largely a mystery" (Adam, 2013, p. 417). We think that today, given the current scientific record, this conclusion is still broadly correct. Interestingly, the reason that the mystery persists is not that no biological correlates for mental disorders have been found, or that no genes have been implicated. On the contrary, past research efforts have shown that neurotransmitters such as dopamine are clearly implicated in psychopathology, and there have been major advances in uncovering the structure of the polygenic background of mental disorders. However, these findings have not been translated into convincing reductive explanations of mental disorders through central pathogenic pathways rooted in neurobiology, as many had expected.

One way to respond to this situation, which is not uncommon in the psychiatric literature, is through continued optimism that someday, with even better equipment and methods and even more participants, we will hit on a reductive explanation of mental disorders (e.g., see Insel & Cuthbert, 2015). However, another way to respond to the lack of success in formulating reductive explanations is to accept the scientific record at face value. There is currently no compelling evidence for the viability of reducing mental disorders to unique biological abnormalities, both in terms of enhanced etiological understanding and of improving the effectiveness of interventions. Given this absence of compelling evidence, it seems sensible to entertain the possibility that explanatory reductionism is wrong; i.e., that mental disorders are not brain disorders, that they do not have a privileged description at the level of (neuro)biology, and that we will never find out 'what mental disorders really are' through neuroscientific and/or genetic research.

In fact, the present paper aims to show, following through the logic of recently proposed network models of psychopathology (Cramer, Waldorp, van der Maas & Borsboom, 2010; Borsboom & Cramer, 2013; Borsboom, 2017) that if it makes sense to understand mental disorders as arising from the causal interplay of symptoms and other factors in a network structure, there may be no reductive biological explanation that awaits discovery because, contrary to quite widely shared current opinion, mental disorders are not brain disorders at all. We will elaborate on three primary reasons for this. First, symptom networks preclude the identification of a common cause of symptomatology as a neurobiological/genetic condition, because in symptom networks there is no such common cause. Second, many causal connections in mental disorders cannot be understood without referring to the content of mental states and thus presuppose some form of intentionality. Third, psychopathology networks are likely to depend to some extent on cultural and historic variation, which means that they are in part context dependent. All this does not mean that mental disorders are not accessible via scientific means, or are just a social construction; but it does mean that explanatory reductionism is not a viable strategy: mental disorders cannot be explained in terms of neural mechanisms. Before we delve into these reasons for why mental disorders are not brain disorders from a network perspective, we start out by providing a working definition of the kind of reductionism that features in psychiatry and clinical psychology.

Explanatory reductionism in mental health research

Explanatory reductionism, in the context of mental health research, is the thesis that mental disorders can be explained in terms of biology. The hallmark theoretical strategy of reductionism is the identification of a phenomenon designated by a higher-

level theoretical term (i.e., a mental disorder) with a property that can be defined at a lower-level (i.e., a biological phenomenon; Oppenheim & Putnam, 1958; Nagel, 1961; Fodor, 1974; Schaffner, 1974; Kievit, Romeijn, Waldorp, Wicherts, Scholte, & Borsboom, 2013). For instance, the most famous successful reductive explanation in the history of science – the reconstruction of the ideal gas laws in terms of statistical mechanics – rests on the identification of temperature (higher-level concept) with average kinetic energy of particles in a gas (lower-level concept; Nagel, 1961) [1]. Similarly, the most famous reductive explanation in psychiatry – the explanation of General Paralysis of the Insane in terms of bacterial infection (Hurn, 1998) – rests on the identification of the cause of symptomatology (higher-level concept) with the bacteria Treponema Pallidum (lower-level concept[2]). In general, explanatory reductionism, in psychiatry and elsewhere, depends on the hypothesis that psychiatric conditions – either as currently defined, or as defined in future theoretical systems – can be identified with (a set of) neurobiological mechanisms and properties, possibly by altering or correcting the description of higher-level phenomena along the road (e.g., 'bumpy reduction'; see Bickle, 1998).

It is important to note that, for such a reductive explanation to work, the reducing science should ultimately be able to identify the lower-level properties that enter into the reductive explanation independently of the higher order science. Thus, just as one

[1] This characterization of reduction is somewhat simplified, as one can question whether even these hallmark reductive explanations are in fact successful and, if so, in what sense (e.g., see Fodor, 1974; Schaffner, 1974; Bickle, 1998; Wimsatt, 2006; Eronen, 2013). However, in the current context we take the charitable, and in our view sensible, position that the cited examples do provide successful reductive explanations.

[2] Note that, in the present context, such identification does not necessarily mean that the entire causal process by which the biological root cause operates is known, just that it has been established as a root cause.

can identify kinetic energy of particles without using the higher-level concept of temperature, and just as one can identify the bacteria Treponema Pallidum without using the behavioral symptomatology of General Paralysis of the Insane, explanatory reductionism in psychiatry requires a theoretical system that allows one to identify the hypothesized brain disorders *as* brain disorders.

For setting up a reductive explanation of mental disorders, it is thus insufficient to merely identify neural correlates of psychiatric conditions. This is the case for at least three reasons. First, because the determination of neural correlates depends methodologically on the antecedent assessment of psychiatric disorders using the concepts of the higher-level science (e.g., symptomatology as defined in the DSM-5; APA, 2013). This means the lower-level description piggybacks on the higher-level description instead of, as explanatory reductionism requires, the other way around. Second, because *some* neural correlate is guaranteed to exist for any behavioral measure, given the plausible rejection of Cartesian substance dualism, finding neural correlates by itself does not provide evidence for reductionism vis-à-vis any other thesis about the relation between biology and psychopathology – most importantly, one of the many varieties of nonreductive materialism. Third, because for most correlates it is unclear whether they are realizations, causes, or effects of psychiatric symptomatology; for instance, deviant neurotransmitter levels may be a cause of depressive symptoms, but they may also arise from the presence of these symptoms, which often include prolonged changes in sleep patterns, appetite and weight, and physical activity levels. Correlations by themselves cannot disentangle these possibilities. Importantly, because explanatory reduction implies the explanation of higher-level phenomena from lower-level phenomena, rather than the mere

identification of correlations between these levels, it also implies the possibility of constructing a biological definition of and diagnostic protocol for the identification of mental disorders.

The idea that mental disorders are in fact brain disorders, which will in the future be diagnosable using lab tests, is not an extreme thesis. In fact, it is overtly espoused by some of the most authoritative sources in psychiatry. Perhaps the most prominent and bold expression of this thesis was given in a paper entitled "Brain disorders? Precisely", authored by the former leadership of the National Institutes of Mental Health (NIMH) and published in the leading journal *Science* (Insel & Cuthbert, 2015). As the title suggests, the core idea of the paper is that mental disorders literally *are* brain disorders: "As new diagnostics will likely be redefining 'mental disorders' as 'brain circuit disorders', new therapeutics will likely focus on tuning these circuits" (Insel & Cuthbert, 2015, p. 500). This kind of explicit explanatory reductionism regarding mental disorders is relatively mainstream. For example, it is also evident in a 2014 *Nature* editorial (Ledford, 2014) which, when comparing depression to cancer, noted that "[...] the reality of cancer is easy to grasp: tumors can be seen, monitored and removed. No such certainty exists in depression, where the affected tissue is locked inside the brain [...]" and in a citation that Solomon (2014, p. 370) attributes to the former head of the National Alliance for the Mentally Ill: "It's a chemical imbalance just like the kidney or liver [...]. We've developed a five-year campaign to end discrimination by making these illnesses understood to be brain disorders and nothing more". Some researchers suggest that the identification of mental disorders with brain conditions is not even a hypothesis anymore, but is an established fact; for instance, Hoogman et al. (2017) stated in *Lancet Psychiatry* that their data "confirm

that patients with ADHD do have altered brains and therefore that ADHD is a disorder of the brain". The prominence of these sources, and the prestigious outlets in which their theses are published, show clearly that explanatory reductionism is not a straw man. Rather, the almost casual way in which authors make their case suggests that it is a rather middle-of-the-road philosophy in the research community.

The idea that mental disorders ultimately are brain disorders has important ramifications. First, it implies that it is possible in principle to identify a common pathogenic pathway at the level of the brain that causally explains symptom patterns. This inspires the search for 'biomarkers' of mental disorders as well as lab tests that should be able to identify them (e.g., see Redei, Andrus, Kwansky, Seok, Ho, & Mohr, 2014). Second, the reductionist mindset implies that, if the common pathogenic pathway can be intervened on, such interventions should have broad effects across the symptomatology, just as killing the bacteria Treponema Pallidum in time prevents or cures the symptomatology associated with Syphilis. Although research into medical interventions can in principle be justified independently, and need not rely on reductionism, the idea that mental disorders are ultimately brain disorders may therefore also partly determine the setup of psychiatric research that revolves around Randomized Controlled Trials (RCTs) to test the effectiveness of medication. Third, the idea that mental disorders should be explained in biological terms has important funding consequences. The NIMH, for instance, has endorsed the position that fundable research proposals should show "not only that an intervention ameliorated a

symptom, but that it had a demonstrable effect on a target, such as a neural pathway implicated in the disorder"[3].

In sum, explanatory reductionism is a widely espoused thesis that holds that mental disorders can in principle be conclusively explained on the basis of biology (e.g., through neurological, biochemical, molecular, and genetic explanations). The thesis rests on the idea that mental disorders are literally brain disorders (although the exact sense in which this identification should be understood may differ across researchers). Its research strategy is aimed at the discovery of pathogenic processes that underlie mental disorders, typically taken to exist at the level of the brain. The hope is that, after identifying mental disorders as biological abnormalities, we can come up with a treatment plan that restores or ameliorates these abnormalities and, as a result, removes the symptoms that people suffer from. In a non-trivial sense, explanatory reductionism thus aims to find out *what mental disorders really are*, and it is based on the premise that the answer to this question lies at the neurobiological level of description.

The network approach to mental disorders

Despite the powerful reductionist mindset present in psychiatry, one of the main recent theoretical developments in psychiatry and clinical psychology has been to move away from monocausal explanations of mental disorders (Kendler, 2005, 2012). Instead, many have come to accept the idea that mental disorders are massively multifactorial in their causal background, that many mechanisms that sustain

[3] https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2014/a-new-approach-to-clinical-trials.shtml

disorders are transdiagnostic, and that mental disorders require pluralist explanatory accounts (Kendler, 2008; Nolen-Hoeksema & Watkins, 2011; Borsboom, Epskamp, Kievit, Cramer, & Schmittmann, 2011). Kendler, Zachar, and Craver (2011) have extended these findings to the ontology of mental disorders itself, and have suggested that this ontology should not be based on essentialism of any kind (including biological essentialism). Instead, they hold that psychopathology should be conceptualized in terms of what they call *mechanistic property clusters*: constellations of properties (defined at different theoretical levels) that hang together because they are connected by a diverse set of mechanisms, analogous to modern accounts of how properties cluster in species as developed in theoretical biology (Boyd, 1991, 1999).

A research program that has put that idea to work is the *network approach to mental disorders* (Cramer, Waldorp, van der Maas, & Borsboom, 2010; Borsboom & Cramer, 2013; Borsboom, 2017; Fried, van Borkulo, Cramer, Boschloo, Schoevers, & Borsboom, 2017; Fried & Cramer, *in press*). Instead of conceptualizing psychiatric problems as symptoms of 'underlying disorders' that are produced by some currently unknown biological pathogenic pathway, the network model explains the robust patterns of covariation in symptom data by assuming simply that symptoms directly influence one another (Borsboom, 2008). For instance, insomnia and fatigue (two symptoms of major depression) do not covary because they are caused by the same pathological (neuro)biological/genetic condition, but because they are directly related: insomnia -> fatigue (Beard et al., 2016). On the other hand, experiencing hallucinations (symptom of psychosis) and sad mood (symptom of major depression) do not covary as much because hallucinations are not very likely to directly cause sad mood or vice versa; and even if these symptoms may be causally associated, the

relevant causal relation is likely to be much more indirect (e.g., hallucinations -> anxiety -> sad mood; Isvoranu et al., 2016), which explains why the correlation between these symptoms is somewhat weaker. As such, the network approach offers a plausible explanation for robust association patterns among symptoms in empirical data (Cramer et al., 2010; Boschloo et al., 2015) and between such symptoms and external stressors (Borsboom, 2017; Isvoranu et al., 2016; McNally, Heeren, & Robinaugh, 2017). In addition, it offers plausible accounts of several other psychiatric phenomena such as comorbidity, spontaneous recovery, and heterogeneity (Borsboom & Cramer, 2013; Cramer, van Borkulo, Giltay, van der Maas, Kendler, Scheffer, & Borsboom, 2016; Fried et al., 2017, Mcnally, Mair, Mugno, & Riemann, 2017), and is compatible with theories and models that focus on the dynamic interplay between symptoms over time (Bringmann, et al., 2013; Wichers, 2014).

From a network perspective, mental disorders arise from direct interactions between symptoms in a network architecture. This happens as follows. Symptoms can be activated by factors external to the person, such as life events (e.g., loss of a loved one -> sad mood; Fried et al., 2015), or they may arise through processes inside the person, including neurobiological dysfunction (e.g., mislabeling auditory sensations that arise from the brain -> hallucinations, catastrophic misinterpretation of arousal -> panic attack, insufficient top down control of behavior -> inability to stop worrying about germs). When a symptom gets activated, it can in turn activate symptoms it is directly related to (e.g., loss of a loved one -> sad mood -> insomnia -> self-reproach). In this way, a disorder grows out of a network of symptom-symptom relations. In particular, this happens when these relations are sufficiently forceful to lead the network to sustain its own activation, producing a hysteresis effect that keeps

the system activated even if precipitating causes have waned (Cramer et al., 2016). Thus, in this conceptualization 'to suffer from a disorder' means 'to be trapped in the stable state of a self-sustaining symptom network' (Borsboom, 2017).

Interventions in the network structure can involve targeting a symptom (e.g., using antipsychotics to counter hallucinations, or using a sleep intervention to counter insomnia) or symptom-symptom connections (e.g., training a person to recognize psychotic symptoms so that timely measures can be taken to preclude them from causing more problems, teaching a person how to control rumination, so that late-night worries no longer lead to insomnia; Borsboom, 2017). Importantly, just like symptoms can arise from both internal and external causes, interventions on symptoms can involve both biologically based interventions (e.g., medication, electro-convulsive therapy, deep brain stimulation), behavioral interventions (e.g., behavioral activation), psychological interventions (e.g. cognitive restructuring), and changes in the environment (e.g., relocating a person with substance abuse disorder to a place where no drugs are available, creating jobs for individuals who suffer from certain forms of psychopathology).

Although network approaches do not rule out the importance of biology in realizing symptoms and symptom-symptom connections (Borsboom & Cramer, 2013; Fried & Cramer, *in press*; Borsboom, 2017), the general network definition of mental disorders in terms of the alternative stable state of a symptom network does not align with the idea that mental disorders are brain disorders. In fact, if the network model is

broadly correct, it pulls the rug from under the explanatory reductionist's feet in many if not all cases of psychopathology, as we will argue in the remainder of this paper[4].

Symptom networks versus the common cause model

In the standard disease model, which has played an extremely important role in the medical sciences, symptomatology arises from a common cause in the body (Hyland, 2011; Cramer et al., 2010; Borsboom & Cramer, 2013). For instance, the symptoms of fatigue, headaches, and foggy eyesight may be caused by a brain tumor that plays the role of common cause; accordingly, medical treatment is often targeted at such common causal factors (e.g., surgical removal of the tumor)[5]. In psychometric approaches to psychiatric symptomatology, correlations between symptoms are similarly analyzed in terms of a latent variable model, in which the disorder is conceptualized as an unobserved common cause of the symptoms (Reise & Waller, 2009).

This offers a potential inroad for explanatory reductionism to operate. Namely, if correlations between symptoms are in fact produced by their common dependence on a latent variable, then equating that latent variable with a brain disorder is tantamount to achieving the hallmark event of explanatory reductionism, namely cross-level identification. Thus, the identification of the common cause of the symptoms (higher-

[4] Note that the current paper does not argue *that* the network model is correct – plausibility arguments have been given elsewhere (e.g., see Cramer, Waldorp, van der Maas & Borsboom, 2010; Borsboom & Cramer, 2013; Borsboom, 2017; McNally, 2016). Instead, the argument is conditional: *if* this model (or something at least as complicated) is correct, *then* reductionism will fail.

[5] Note that this does not imply that all medical accounts are adequately captured in the common cause model (Guloksuz, Pries, & van Os, 2017), just that the common cause explanation plays a very important role in medicine (Hyland, 2011).

level concept) with a biological property of the brain (lower-level concept) would explain the correlations between symptoms reductively and offer a crucial advance in the reductive program. This scenario is visually represented in Figure 1a.

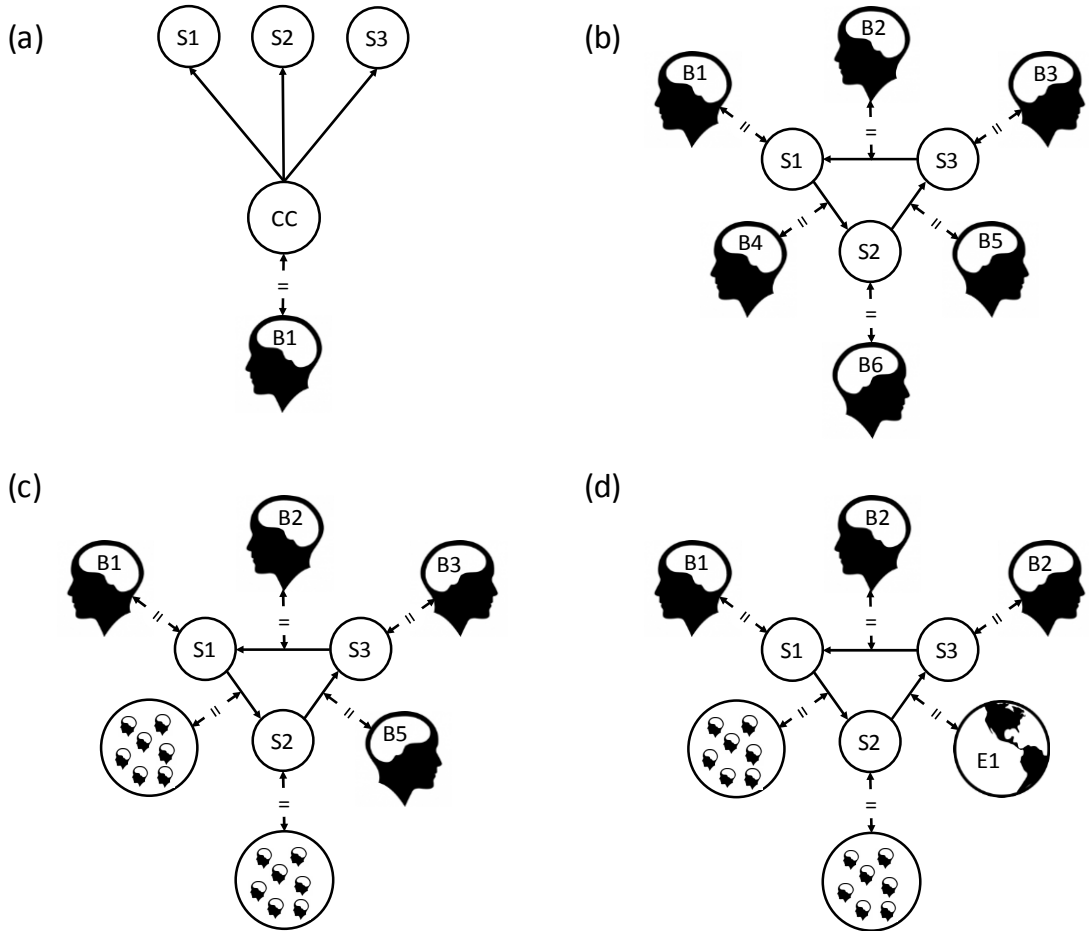

Figure 1. Four prospects for (partial) explanatory reduction in psychopathology. Panels a-d depict relations between symptoms S, brain states/mechanisms B, a common cause CC, and external factors E. Identifications are marked '='. In panel (a), a *brain disorder* scenario, correlations between symptoms are produced by a common cause, and that cause is identifiable with a particular dysfunctional brain state or mechanism. In panel (b), a *network reduction* scenario, correlations arise from a network structure, but all symptoms and relations between them are identifiable with different brain states and neural mechanisms. In panel (c), a *patchy reductionism* scenario, some symptoms are identifiable with brain states, but S2 as well as the relation S1 -> S2 arise from rational relations between multiply realizable mental states, and hence resist reduction to underlying biology. In panel (d), a *patchy reductionism with partial externalism* scenario, in addition relation S2 -> S3 is realized in a mechanism that is external to the individual.

If a network model is correct, however, there exists no such common cause, so the above theoretical move is blocked. In the network model, mental disorders do not resemble medical diseases entities, which are theoretically and empirically identifiable independently of the symptomatology (e.g., a structural MRI that shows the presence of a malignant tumor, or a blood test that proves HIV infection; see Borsboom and Cramer, 2013). Instead, in the network definition, disorders behave more like *fields* – e.g., their emergence is analogous to the appearance of magnetization through the aggregate behavior of pieces of ferromagnetic material that lock into a particular position (Epskamp et al., *in press*)[6] and to stable states in complex ecosystems (van der Maas, Dolan, Grasman, Wicherts, Huizenga, & Raijmakers 2006; Scheffer et al., 2009; Van de Leemput et al., 2013). The emergent global states that networks can get locked into are explained from the symptoms and symptom-symptom connections that make up the network structure, and not from shared dependence on a distinct common cause. Thus, the royal road into explanatory reductionism – an explanation of the "spirochete-like variety" (Kendler, 2005) as displayed in Figure 1a – is out of the question in network models.

From a reductionist position, one could argue that we have simply complicated the research question by adopting a more complex model. For example, one could hold that even while there may be no biological common cause in the system, each of the individual symptoms in a psychopathology network nevertheless corresponds to a

[6] This comparison is not just metaphorical; the Ising model used to identify network structures (Van Borkulo, Borsboom, Epskamp, Blanken, Boschloo, Schoevers, & Waldorp, 2014) is in fact mathematically identical to models for magnetism and, understood in terms of network dynamics, features largely the same phenomena (Cramer et al., 2016).

neural property, and that each of the connections in the network corresponds to a neural mechanism that connects the relevant neural properties. In this way, successful explanatory reductionism would not involve biological identification of '*the* common cause' of symptomatology, analogous to the identification of a virus or tumor, but rather of '*many* causes' of the symptoms and connections in the network. In this *network reduction* scenario, which is depicted in Figure 1b, mental disorders can be still justifiably seen as brain disorders – at least, as long as the relevant neurobiological states and mechanisms are identifiable at the biological level (i.e., without knowledge of the phenomenological and behavioral level regarding symptoms or connections between them). If such identification is possible, then a complex explanatory reduction is feasible, and mental disorders would still be brain disorders, be it complex ones.

However, while this form of reduction represents an interesting logical possibility that as far as we know has not been thoroughly analyzed in the context of psychopathology, we think it is unlikely that this scenario will play out. While in many cases, connections between symptoms are indeed strongly grounded in biological mechanisms (e.g., insomnia -> fatigue; appetite loss -> weight loss, drug use -> tolerance, etc.) that may support 'local' reductions of certain parts of the symptom network ('patchy reductionism'; Kendler, 2005, Schaffner, 2006), there are other cases where such grounding should not be suspected to support successful explanatory reduction. In the following, we will show this by providing a detailed analysis of symptoms that involve the *content* of mental states.

The content of mental states

The relations between symptoms in network structures often involve intentional information. By 'intentional information' we mean: descriptions of mental states such as beliefs, desires, emotions and intentions that indicate 'what they are about'. We have beliefs *about* being persecuted, or *about* the floor being dirty. We have fears *about* spiders or germs. Hence, a term like 'contamination fear' contains intentional information: it tells us that the fear is *about* contamination. Brentano (1874) famously argued that intentionality is 'the hallmark of the mental': that all mental states are about something (for some recent versions of this view see Dretske, 1997; McDowell, 1996). Here, we adopt the more modest assumption, accepted by almost all participants in the discussion, that *many* mental states have intentional content, while others, like pains or undirected anxieties, may not (McGinn, 1982; Searle, 1983).

The symptomatology in systems such as DSM-5 (APA, 2013) frequently relies on such intentional information. For example, the symptom of 'craving' in alcohol use disorder is defined as "a strong urge or desire to use alcohol" (APA, 2013). Here, a reference is made to the *content* of the desire (it being *about* alcohol). Or, to give another example, one of the symptoms associated with depression is described as "feelings of worthlessness or excessive or inappropriate guilt" (APA, 2013). These feelings are about things as well: worthlessness involves feelings about oneself, and guilt involves feelings about things one has done or should have done (however vague or implicit this content might be). Finally, delusions are almost always identified through their content, because they involve a mental state (i.e., a belief) which does not match reality in an appropriate way: after all, people who think their thoughts are broadcasted are considered delusional largely because of the content of their beliefs.

Thus, psychiatric symptomatology often refers to the content of mental states and thus involves intentional information.

Importantly, because symptoms are often described in intentional terms, the covariations observed between symptoms in a network can be seen to *make sense*: only at this level of description can we *understand why* the presence of one symptom (e.g., a person believing that the CIA spies on him or her) leads to another (e.g., the person closes the curtains and withdraws from social life). That is, the intentional description allows us to put ourselves in the patient's shoes (Jaspers, 1923/1963). And although the main criterion for the success of network models is empirical validation, the fact that the observed covariations make sense from that perspective greatly adds to their explanatory value. We suggest that these connections make sense because there is a *rational relation* (Nordenfeld, 2007) between, say, contamination fear and washing, which does not exist between, say, contamination fear and binge eating. This rational relation involves a connection through the *content* of the mental states in question which only becomes visible because symptoms are described in intentional terms. To illustrate this point, we discuss three examples in detail.

First, covariation between the occurrence of obsessional beliefs and the occurrence of compulsive behaviors, such as washing or hoarding behavior (Abramovitz, Khandker, Nelson, Deacon & Rygwall, 2006; Tolin, Brady & Hannan, 2008) can be explained by understanding compulsive rituals as a response to obsessional beliefs. That is, rituals are performed in order to relieve the fear raised by those beliefs (Rachman, 1998; Salkovskis et al., 2000). To understand compulsive rituals as a response, one needs to refer to the content of the beliefs and fears involved: because the patient has

a belief *about the spreading of germs*, he or she becomes afraid *of contamination*, and because this is the content of his fearful feeling, he or she responds to it by excessive washing rituals. Now of course, it is possible that although on the surface one's fear is about germs, on a deeper psychological level the fear may be about something else (e.g., losing control, death). However, this just shows that intentional explanations might be available on different levels. The point is that the connection between handwashing and fear is only explanatory in so far as one takes into account what the fear is about (i.e., its intentional content). This also shows that the compulsive rituals engaged in by the patient are not 'completely crazy': given that a patient has those beliefs and fears, it is actually to some extent understandable that he or she responds in this way. After all, washing is generally a reasonable strategy to counter the spread of germs.

A second example is found in the relation between certain symptoms occurring in panic disorder. Two symptoms that show a connection are so-called anxiety sensitivity ('fear of fear' or a fear response to signals of anxiety; McNally, 1990, 2002), and avoidance behavior (Borsboom & Cramer, 2013; Reiss, 1991; White, Brown, Somers, & Barlow, 2006). This connection can also be shown to 'make sense' by taking the content of the involved beliefs and emotions into account (Reiss, 1991). If one is afraid of having a panic attack (and thus become fearful in response to signals of such an attack) *and* believes that certain situations increase the chance of an attack (such as being in a large group of people, or in an enclosed space such as an airplane), it certainly makes some sense to avoid such situations.

Third, in major depression thoughts of self-reproach and low self-worth are correlated with thoughts of suicide and actual suicide attempts (Cramer et al. 2010; Dori & Overholser, 1999; Wild, Flisher & Lombard, 2004). Here a similar form of 'making sense' can be observed: thoughts of self-reproach and low self-worth can involve the belief that one's life is not worth living, or that one is a burden to one's environment. Such a belief may in turn lead to the conclusion that suicide is the best solution. Again, this is not completely unintelligible or crazy: if one is really convinced of the lack of worth of one's own life, it seems at least to some extent reasonable to consider the possibility of ending it.

But what exactly does it mean to say that a connection 'makes sense'? So far, we have said things such as: if one has symptom A, it is 'understandable', 'not completely crazy', or maybe even 'reasonable' that one also has symptom B. But what kind of claim is being made here? The view we adopt here is rooted in a tradition of thinking about understanding, that is based on the work of Donald Davidson (1984) and Daniel Dennett (1987), generally known as interpretivism (for recent contributions see Francken & Slors, 2014; Mölder, 2010; Thornton, 2010). Interpretivism emphasizes the pragmatic nature of belief-desire talk: we ascribe mental states with specific content to others and ourselves, in order to better explain and predict behavior. Within the interpretivist tradition, saying that it *makes sense* that the presence of self-reproach correlates with suicidal tendencies, means that we can explain the fact that people who feature self-reproach also feature suicidal tendencies by referring to their *basic rationality* (Davidson, 1984). Here one should think of basic skills such as the capacity for deduction, or means-end reasoning. By applying such skills to for example the belief *that one is worthless*, a patient can reach the conclusion that it

might be best to end his/her existence. Due to the rational relation between these two phenomena, the ascription of such thoughts - combined with basic rationality - helps in predicting suicidal tendencies.

Such ascription of basic rationality does not mean that there is nothing unreasonable going on. A patient's fear may be out of proportion, or may involve false beliefs. In fact, it almost certainly will involve some irrationality, otherwise there would be no reason to categorize the person's behavior or thinking or feeling as psychopathological in the first place (Kalis, 2011). But in order to understand a fear as disproportional, or a belief as false, we already need to ascribe a large background of proportional fears and true beliefs to a person. Something can only be seen as an irrational deviation in so far as there is a background of rationality in place (Davidson, 1984). What we do when we 'make sense' of symptom covariation, is to make explicit this background of rationality in the person's behavior, fears, beliefs and so on (for recent critical discussion of this view see Bortolotti, 2010; Campbell, 2009).

Thus, the content of mental states plausibly plays a crucial role in causally connecting symptoms to each other (Baker, 1995), and in addition allows us to recognize them as patterns that make sense. This poses a problem for explanatory reductionism, even if it is reformulated as to apply to the symptoms and symptom-symptom connections as in the network reduction scenario in Figure 1b. There are two reasons for this.

First, the rational relations between symptoms, as discussed above, depend essentially on the intentional description of these symptoms. Even if one could describe the symptoms themselves at the level of neurobiology, it is unlikely that relations

between symptoms, that are immediately recognized as rational at their intentional level of description, will (or can) be recognized as such at their neurobiological level of description. Thus, the explanatory force of symptom networks partly depends on the intentional level of description of the symptomatology, and this may very well turn out to be a matter of principle.

Second, mental states as they arise in symptomatology are almost certainly multiply realizable. Multiple realizability exists when there are multiple physical ways to 'realize' a given object or property (Fodor, 1974; Horgan, 1993; Putnam, 1967; Pylyshyn, 1984). For example, one can realize the abstract concept of *a dollar* physically in many ways; as a set of coins, a bill, or a set of bits in a computer. None of these realizations is privileged; that is, it is senseless to ask whether money 'really is' paper or coins or bits, and one cannot investigate the nature of money by, say, chemically analyzing dollar bills. There is a one-to-many mapping between the higher-level concept (money) and its lower-level realizers (bills, coins, bits), and unless this mapping has at least an element of necessity, no systematic connection between them can be forged as a matter of scientific law; in this case, the higher-level concept is said to be 'wildly disjunctive', as is the case for money, which can be realized in indefinitely many ways. Fodor (1974) argued that, for this reason, multiple realizability blocks the classical scheme of explanatory reductionism (Nagel, 1961) as it does not allow the identification of higher-level concepts with lower-level concepts[7].

[7] The exact force of the multiple realizability argument vis-à-vis intertheoretical reduction is contested (e.g., see Bechtel & Mundale, 1999; Bickle, 1998; Gillet, 2003). However, we do not need to take sides on this issue. In the current context, it is sufficient to establish that multiple realizability almost certainly obtains to *some*

For many psychopathology symptoms, multiple realizability is highly likely because they explicitly depend on the content of mental states: symptoms with intentional content (such as delusional beliefs, or fears about heights or spiders) can be physically realized in different ways in different people, just like a photograph can be stored on a microfilm, as some digits in your computer, or on a piece of printed paper. In the context of psychopathology, two people may both believe they are being spied on by the CIA, but this belief may be differently coded in their brains (Aizawa & Gillett, 2009; Endicott, 1993). In both cases, however, the belief in question may instigate deviant behavior (e.g., obsessional searching for hidden cameras in their house), and thus instantiate a relevant connection in the symptom network structure (i.e., delusion -> behavior change). Given what we know about the distributed character of representations and the plasticity of the brain (Endicott, 1993), such a scenario appears rather likely: even if one accepts the possibility that a reductive account could be given of, say, beliefs in general, there is no reason why one should expect the neural realization of the *content* of these beliefs to be invariant across individuals.

In fact, there is a growing chorus of voices in contemporary philosophy of mind arguing that we should not primarily think about beliefs, fears, and other mental states *as* being brain states. According to these approaches, a sentence such as 'John believes that his neighbor is a secret agent for the CIA' does not find its truth

---

extent in psychopathology networks. This is because, even if multiple realizability should not block reduction in principle (i.e., if symptom networks are not wildly disjunctive so that with infinite time and money one *could* find all the realizing conditions involved), its existence makes explanatory reductionism pragmatically unattractive.

conditions in the fact that John has a certain brain state (as in identity theory; Lewis, 1966) but rather in the fact that a coherent set of dispositional ascriptions or counterfactual conditionals is true of him (Baker, 1995). For example, his belief is characterized through a set of conditionals, such as: *if the neighbor were to start a friendly talk, John would respond nervously*; *if John were to pass the CIA headquarters, he would expect his neighbor to be there*; and so on. From a network perspective, *what* a person believes may thus also determine *which* relations in the network are activated: A person who believes the CIA is spying on him may start distrusting his neighbors and start avoiding contact, while a person who believes Napoleon returned from Elba to prosecute them for tax evasion may not.

Importantly, the truth conditions for mental states to have a certain specific content, in this view, are thought to involve the occurrence of a set of meaningfully related phenomena, embedded in a certain context. This also relates to the by now almost mainstream assumption that cognition is embedded (the extended mind hypothesis; Clark and Chalmers, 1998; Menary, 2010). The ontology of a belief such as "my neighbor is a secret agent for the CIA" does not involve only processes within the individual, but also certain elements of the environment, such as responses of others, or the way one organizes one's surroundings.

Because so many symptoms involve intentional states, we submit that the corresponding parts of symptom networks are multiply realizable. This means that at least some individual symptoms and symptom-symptom connections are expected to map to a variety of realizing brain states and mechanisms, as visualized in Figure 1c. If these realizing brain states cannot be characterized as neurobiological phenomena

(i.e., one cannot say that John believes the CIA spies on him solely on the basis of knowledge about his biological constitution), they cannot be identified independently of the phenomenology. In this case, explanatory reductionism fails to provide theoretical concepts at the biological level that can be systematically identified with theoretical concepts at a higher-level, and the idea that 'mental disorders are brain disorders' becomes void. Given such a situation, the strongest viable position that is still available would be nonreductive materialism along the lines of the supervenience thesis in the philosophy of mind (Kim, 1982, 1984). This thesis roughly holds that there cannot be distinct mental states without there being distinct physical states (i.e., changes in physical states are necessary but not sufficient for changes in mental states). In the current scheme of thinking, this would merely imply that there cannot be differences in symptom network states without there being *some* differences in physical states. Apart from the benefit of allowing one to reject Cartesian dualism, this position has no reductionist teeth.

The context dependence of network structures in psychopathology

Some connections in symptom networks are likely to be highly stable across cultures and historical circumstances. For instance, it is likely that, whatever culture you live in, insomnia will lead to fatigue and concentration problems; panic attacks will lead to worry about the implications of these attacks; a sudden decrease in appetite will lead to loss of energy; and phobic fears of spiders are not likely to promote feelings of happiness anywhere on the planet. These relations are likely to be stable across time and place because they depend on uniform biological and psychological homeostatic mechanisms, possibly grounded in the evolutionary background of our species. Elucidating these mechanisms is highly important and, in our view, it is not at all

impossible that some of the connections in symptom networks do allow for a reductionist analysis or something close to it, because in these cases both the symptom states and the connection between these states may be amenable to a description at the biological level.

However, at least some of the connections in network structures are likely to show variations across time and place that cannot be captured by such descriptions (Haroz, Bolton, Gross, Michalopoulos, & Bass, 2016; Haroz, Ritchey, Augustinavicius, Michalopoulos, Burkey, & Bolton, 2017). For instance, consider the relation between feelings of guilt and suicidal ideation. It stands to reason that this connection has a different strength in different individuals, and that such differences can be culturally loaded – one may compare the connection between these symptoms in the case of a Japanese soldier in WW-II who has failed to defend his post to a Catholic Priest who has committed a cardinal sin; for the Japanese soldier, his cultural background facilitates the connection between feelings of guilt or shame and suicide attempts, while for the Catholic priest, his background inhibits it. Similarly, while somatic complaints are stably associated with depression across cultures, *which* somatic complaints are associated with depression may vary across cultures (Marmanidis, Holme, and Hafner, 1994; Ma-Kellams, 2014). Finally, the probability that alcohol withdrawal symptoms in substance abuse will lead to legal problems (both diagnostic criteria in DSM-5) is clearly different for an American living now, as compared to one living at the time of the Prohibition, and the consequences of public drunkenness are completely different in, say, Riyadh versus New Orleans.

In the realm of explanatory reductionism, such contextual variations are noise, or merely concern the way in which disorders are 'expressed' in different times or cultures: behind context-varying symptomatology lurks a homogenous constellation of brain dysfunctions. Perhaps depressed Americans have stomach aches, depressed Israeli get heavy legs, and depressed Japanese suffer from headaches, but in all cases we would find a stable (constellation of) biological factor(s) responsible for these different expressions. That is, from a reductionist viewpoint, heterogeneity in manifestation - due to contextual factors - does not preclude homogeneity of biological essence.

This mode of thinking is not available in the network account. Since disorders *are* states of a network that are determined by patterns of causal interactions between symptoms, there is no independent mode of observation that could serve to ascertain that American, Israeli, and Japanese depressions are 'really the same' even though they feature different symptomatology. All we could perhaps say is that the *role* that stomach aches play in the depression network of Americans is the same as the role that heavy legs play in the Israeli depression network (e.g., a role as a central symptom in the network that is, for example, connected to fatigue and depressed mood). We could also say that the resulting stable network states are *similar*, to the extent that they largely involve the same set of symptoms and characteristic reactions to external events (e.g., stressful life events) and treatment (e.g., medication or psychotherapy). But because in a network model there is no way to identify depression independently of the symptoms and relations between them, there is no way to truly equate different disorders in different people independently – let alone in different historical periods or cultures. Note that this is not a practical or

methodological problem that we could expect to be solved with the advent of better measurement techniques – no such development can be expected, because the relevant identification criteria simply *do not exist* in networks.

In the above examples, cultural and historical variations impinge upon the network structure through the content of the mental states involved in the symptom network: precisely because rational relations between mental states and behavior can produce causal relations at the symptom level, different contents of these mental states can produce different symptom-symptom couplings. Thus, cultural and historical variations can lead symptom networks to partially differ across place and time, which will lead to differences in the kinds of stable problem states that networks create and thus can lead to different disorders. In this way, the network model naturally accommodates an integrationist picture in which biological and cultural factors together shape mental disorders (Murphy, 2005; Hacking, 1999).

However, there is another reason why contextual differences may change network structures, and that is that some causal relations between symptoms are literally realized outside the person. That is, they rest on or invoke mechanisms in the environment. A clear example is the relation between the excessive gambling and the desperate financial situations it leads to, both of which play a role in the DSM-5 diagnostic criteria for gambling disorder (APA, 2013). This connection is forged entirely outside of the person, namely by the operational specifications of gambling setups (e.g., fruit machines, Roulette tables, etc.). Importantly, even if the desire to gamble is taken to be a mental state that is realized in the person's brain, the operating characteristics of the fruit machine are not; and these operating characteristics realize

the causal connection between gambling and the debts it leads to. Thus, insofar as a symptom network rests on interactions with and specifications of the environment, its ontology is *extended*, i.e., it is not located in the person's head (see Clark and Chalmers, 1998, for a similar thesis in the philosophy of mind). This possibility is represented visually in Figure 1d, in which a causal connection between two of the symptoms in the network is sustained by a mechanism external to the individual.

Thus, in network models, not only cultural and historical features, but also the environment itself may become part of the network structure, and thus part of the disorder. More or less by definition, this means that parts of the network structure will defy a purely biological explanation, and that cultural and historical factors as well as external mechanisms, to some extent, shape mental disorders. Importantly, however, this does not mean that psychopathology is out of scientific reach, or that mental disorders are 'just a social construction' (see also Murphy, 2006). Differences in network structure across cultural and historical backgrounds are amenable to theoretical and empirical research: one can use network approaches to model the process by which cultural factors shape mental disorders (e.g., by including cultural factors as additional nodes in a disorder network), one can simulate such processes, one can test the resulting models against relevant data. Thus, that mental disorders are partly a function of historical and cultural variations does not make them less real or render them inaccessible to scientific study (see also Schaffner and Tabb, 2014).

Prospects for reductionism in psychopathology networks

The fact that organization, content, and context matter so much in symptom-symptom interactions makes a general reductive move, which would explain such interactions

entirely based on biology, highly unlikely. Building on early work by Daniel Dennett, in accordance with Kalis (2014) we suggest that mental disorders should instead be understood as *real patterns*[8]*:* "there are patterns in human affairs that impose themselves, not quite inexorably but with great vigor, absorbing physical perturbations and variations that might as well be considered random; these are the patterns that we characterize in terms of the beliefs, desires, and intentions of rational agents" (1987, p. 27).

Dennett (1987; 1991) introduced the idea of *real patterns* in order to show how intentional explanations (Dennett calls it 'taking up the intentional stance') have unique explanatory value. The information provided by taking up the intentional stance is *unique* in that these patterns cannot be made visible by analyzing symptoms in lower-level functional or physical terms; the patterns are nevertheless *real* in the sense that all the elements of the pattern are physically instantiated, and that there are real causal processes involved.

Even though Dennett himself has sometimes been called a reductionist, we think Dennett's notion of a *real pattern* can convincingly show why the reductive strategy for understanding mental disorders will not work if the network approach is roughly correct[9]. Even if one were to gain perfect knowledge about the physical processes that

[8] During the review procedure of this article, we learned that Tabb and Schaffner (2017) independently also developed an explanatory account referring to Dennett's notion of real patterns. Our analyses are complementary, although Tabb and Schaffner put slightly more emphasis on the pragmatic aspect of Dennett's approach.

[9] Dennett could be called a reductionist in that he holds that there is no such thing as 'intrinsic intentionality' as opposed to 'derived intentionality' (Dennett, 1997). Also, he claims that the intentionality that can be ascribed to human beings ultimately originated from characteristics that can (only) be described by means of the design

instantiate a symptom network, and would list them all, one would end up with an unorganized set of relations between distinct sets of biological states and processes that make no sense at the level of biology (Fodor, 1974). This is because the *pattern* that is present in this set is not visible at the lower-level as it involves the organization of sets of symptoms in terms of causal and intentional relations, which can only be made sense of by taking both the content of mental states and the world outside the patient's head into consideration. In addition, if multiple realizability obtains then what makes a given constellation of biological conditions a realization of a mental state (e.g. contamination fear) is not a fact about the biology of the system, but the fact that it functions in the symptom network in the appropriate causal way (e.g., leads to compulsive handwashing); and it is highly likely that it has this causal function because of the rational relation between mental states and behavior.

In addition, symptom networks have holistic properties precisely because the causal relations between symptoms make some sense. What makes a set of phenomena into a case of obsessive-compulsive disorder is that a person suffers from compulsive handwashing *in response to* fear of contamination, a relation that exhibits basic rationality. And the 'sense' that we see in the relations between symptoms in turn depends on the context: compulsive handwashing is only compulsive if it can be seen as an *excessive* response, and this cannot be determined by looking inside the person's

---

and the physical stances (Dennett, 1997, for a critical response to this idea see Searle, 1998, p. 90-91). However, Dennett is clearly an anti-reductionist in the sense we are concerned with here: he argues that intentional explanations have independent value: they are not derived from, nor shorthand for, lower-level explanations (Elton, 2003), and recent analyses have also stressed the compatibility of Dennett's position with the minimal realism involved in interventionist conceptions of causation that nicely match the network paradigm (Eronen, 2017).

head: what is excessive in a Western country is not excessive in a country ravaged by Ebola. This means that whether behavior is to be considered symptomatic or normal not only depends on a person's psychological or neurobiological state, but also on the environment.

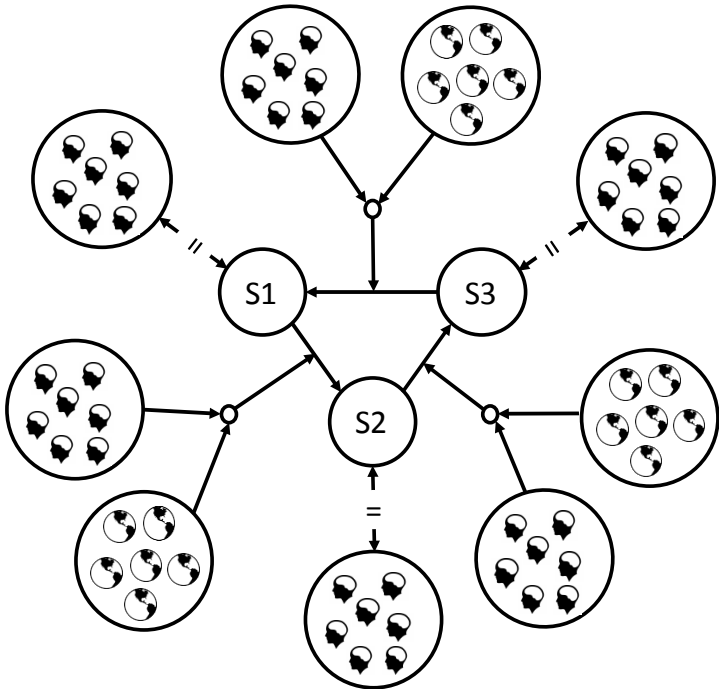

Figure 2. Massively multifactorial symptom networks. The figure shows a situation in which symptoms S1-S3 involve multiply realized mental states. Network connections depend on the combination of heterogeneous sets of biological and environmental conditions, as indicated by the concatenation symbol **O**, that may moreover differ over individuals.

Given these considerations, even the *patchy reductionism with externalism* scenario as represented in Figure 1d may be too simplistic. Many connections between symptoms are unlikely to be fully realized in the biology of the human system (as, e.g., lack of appetite -> weight loss may be) or in the external world (as e.g., excessive gambling -> debts may be). Instead, the presence and strength of these connections is likely to depend on a mixture of biological and external factors and mechanisms, many of

which may be multiply realizable at both the biological and the external side of the equation and which may combine in complicated ways.

For instance, the presence and strength of a causal relation such as feelings of worthlessness -> suicidal ideation probably depends on a mix of variables that are most efficiently described as biological states (e.g., dopamine dysregulation), psychological processes (e.g., self-blame as a dysfunctional coping mechanisms), and external conditions (e.g., the amount of social support). Moreover, it is unlikely that these factors combine in a simple manner (e.g., in a purely additive way). For this reason, even if one takes the actual symptom states of mental disorders to reside 'in the person' in the sense that they are multiply realized functions of the biological condition (internalism), the strength of the causal connections between these states themselves is highly unlikely to be realized solely in the biology of the human system[10].

In this scenario, which features *massively multifactorial symptom networks* as visualized in Figure 2, not even patchy reductionism is feasible, as basically every element of the system is dependent on a heterogeneous set of biological and external factors. Although we cannot rule out *a priori* that some mental disorders (either as currently defined, or under a future diagnostic scheme) may be fully or partly reducible to underlying biology (as in Figures 1a-1d), and it would certainly be a great scientific discovery if such a reductive explanation were to be construed, given the

[10] Note that the interactions described here do not imply gene-environment interaction effects in the sense of common models for population statistics, which operate on individual differences (Franić, Dolan, Borsboom, and Boomsma, 2012), because they may be invariant over individuals and hence may not produce individual differences at all.

current scientific record we think the massively multifactorial situation of Figure 2 is most likely to hold across the psychopathological board. Also, it is pragmatically preferable to take Figure 2 as a point of departure, instead of betting entirely on a reductive scenario in which mental disorders are to be identified as brain disorders because the latter course of action may exclude lines of research that could yield crucial information about psychopathology networks (e.g., regarding cultural variations in the contents of relevant mental states, or the role of the external environment, both of which are likely to get little attention if one commits *a priori* to the brain disorder perspective).

While wholesale explanatory reductionism is unlikely, the question to what extent causally interacting symptoms could be productively analyzed at biological levels of description is currently open. Mental disorders are not brain disorders, but that does not imply that psychopathology research should not be interested in the physical processes involved in psychopathological symptoms. However, to know which physical processes to investigate, and what to conclude from one's investigations, one needs to see these processes for what they are: as physical phenomena that might help us understand the bigger picture of symptoms and symptom networks that we are ultimately interested in. And given that many symptoms involve relations between thoughts, desires, and emotions, investigating physical processes underlying mental disorders is only a relevant scientific enterprise provided one *also* keeps looking at these processes from a perspective that takes the organization, content and context of phenomena into account.

From this point of view, even 'purely bodily' symptoms such as agitation and weight change should be analyzed in relation to other symptoms, many of which will have intentional characteristics that may, in a sense, spill over to these bodily symptoms if these symptoms themselves become part of the content of relevant mental states; for instance, the bodily symptom of weight gain may affect one's self-image and lead to feelings of worthlessness. If this happens, it is not so much the bodily symptom itself that causes feelings of worthlessness, but the mental representation of that symptom, e.g., as a negative aspect of the self (Beck, 2008), which may be overlooked if one focuses too strongly on the biological dimension of such a symptom.

Discussion

Mental disorders are not brain disorders. To the extent that mental disorders arise from the causal interplay between symptoms, as represented in network models (Borsboom & Cramer, 2013; Fried & Cramer, *in* press; Borsboom, 2017), it is highly unlikely that the symptomatology associated with psychopathology can ever be conclusively explained in terms of neurobiology. Therefore, sticking to the idea that mental disorders are brain disorders may be counterproductive and can lead to a myopic research program, because it assumes the implausible position of explanatory reductionism *a priori*. As we have purported to show in the present paper, this position does not stand up to empirical and theoretical scrutiny. Therefore, we tentatively conclude that research programs on mental disorders may be better based on the working assumption that psychopathology is massively multifactorial, not only in its causes but also in its constitution. Neuroscientific research has an extremely important role to play, but only if it leaves behind the explanatory reductionism that characterizes some of its mainstream thinking.

We noted throughout this paper that the argumentation provided, and the conclusions that it affords, are conditional on the assumption that the network model is broadly correct. It is, of course, possible that this assumption is not justified, or is not accurate for a subset of disorders. It may even be that a biological basis for, say, Major Depressive Episode will in fact be found after all; that is, a brain circuit or chemical imbalance may be identified that acts as a root cause for the disorder. However, if our analysis is anywhere near the truth, such a scenario is unlikely. In fact, we should expect to find interactions between symptoms to be grounded in an even more complex set of biological, social, and cultural factors involved in psychopathology. If so, then psychopathology must, like so many other phenomena in nature, be understood as intrinsically complex, so that the simplification of this complexity that explanatory reductionism aims at is unlikely to offer a productive strategy for research; rather, approaches that naturally accommodate complexity are called for (Barabasi, 2012). In this respect, it is important to note that the network models that have so far been suggested in the literature (Borsboom, 2017; Cramer et al., 2016; Marsman et al., 2017) are very simple pairwise interaction models that, in our view, should be seen as a lower bound on the true complexity of the system. Naturally, if our analysis is correct about these relatively simple models, it must *a fortiori* be correct about anything more complicated. This, in our view, means that the argumentation provided in the paper has considerable reach.

Importantly, the complexity of common mental disorders includes biological mechanisms and processes. Thus, while we strongly oppose the explanatory reductionism as voiced, e.g., in Insel and Cuthbert (2015), we do *not* argue that

biological approaches to understanding psychopathological systems are worthless or should be stopped. There is, however, a considerable difference between searching for the biological essence of a disorder and investigating the role of biological processes in a network structure (e.g., the biological underpinnings of the insomnia → fatigue link in a depression network). We conclude that the former approach is unlikely to pay off, but the latter approach should be pursued with vigor, as it can constrain and inform network structures and their resulting dynamics (see also Fried & Cramer, *in press*). In genetics research for example, such an approach would entail that one would no longer be searching for genetic variants that are associated with the presence of a *disorder* such as major depression – either quantified by a sum score of symptoms or by a case-control variable that is based on a cut-off of the sum score – but, rather, for genetic variants that are associated with *specific aspects* of a depression network structure (Cramer, Kendler, and Borsboom, 2011) – for example a specific connection in that network (e.g., insomnia → fatigue) or a specific symptom (e.g., concentration problems).

As such, biological processes may be fruitfully integrated in symptom networks even if they are not common causes of a specific cluster of symptoms. We note that investigating biological underpinnings of symptom networks is probably best directed at connections and symptoms that do not depend on the content of mental states and/or represent highly contextualized variables: that is, connections between symptoms such as depressed mood and suicidal ideation might prove hard to associate with biological underpinnings, given their dependence the content of mental states. To the contrary, prospects are better for finding genetic variants and biological processes that are implicated in connections between symptoms that are grounded in

homeostatic mechanisms: e.g., sleep, appetite, and maybe elementary behavioral variables such as agitation or retardation.

When integrating the distinct levels of biological and psychological information in this way, we deem it unlikely that any one of these levels will gain uniform causal priority, because many relevant processes feature feedback relations that work across different levels. As an example, consider the biological clock – one of the best understood neural mechanisms (Partch, Green, and Takahashi, 2013) – which is involved in sleep regulation and implicated in disorders such as depression and generalized anxiety. The biological clock features a regulatory system that involves gene expression, and in this sense may seem to be an excellent target for an explanatory chain that runs from genes to brain to behavior. However, although the biological clock is highly important in controlling the sleep-wake cycle, it is also very sensitive to environmental cues. It is easy, for instance, to give rats a jet lag by changing the light conditions in their cages (Deboer, Détári & Meijer, 2007). In contrast to rats, humans can change these light conditions themselves, for example by pressing a light switch. A changing light condition then becomes willful behavior of the switch-pressing human and that behavior will alter patterns of gene expression involved in the biological clock. Now it is a small step to envision a feedback loop between genes and behavior: a person who has sleep problems due to dysregulation of the biological clock may keep the light on at night, thereby intervening directly in the gene expression involved in the circadian rhythm, which may lead to further sleep problems. Thus, this relatively simple example already features effects that run from behavior to genes, as well as from genes to behavior (see also Kendler, 2005). Such examples clearly establish that, even one takes the world to feature a bottom-up

mereological ordering in levels (atoms make up molecules, which make up cells, which make up brains, etc.), this does not imply a parallel causal ordering (e.g., chemistry->biology->psychology; see also Wimsatt, 2007, and Eronen, 2013).

Human beings are the most complicated systems ever studied in science. In certain cases, the optimal way of studying such systems is to take them apart to see how each individual component works and what it is made of. However, in other cases, a research strategy that is holistic is more likely to bear fruit: with holistic we mean a research strategy that is focused on the interaction between parts rather than on their individual realization. Mental disorders likely involve feedback loops that cross all the traditional divides between levels of explanation, none of which can claim the status of 'basis' for the others; many of these feedback loops may well be driven by the basic rationality that characterizes human beings. Network theory offers tantalizing possibilities to integrate the biological, psychological, behavioral, and environmental mechanisms that create causal relations between symptoms, and in that sense may offer a starting point for bridging the Cartesian schism that has divided the mental and the biological realms in psychopathology research for so long.

References

Abramowitz, J. S., Khandker, M., Nelson, C., Deacon, B.J., & Rygwall, R. (2006). The role of cognitive factors in the pathogenesis of obsessive–compulsive symptoms: a prospective study. *Behaviour Research and Therapy*, *44*, 1361–1374.

Adam, D. (2013). On the spectrum. [Editorial]. *Nature*, *496*, 416-418.

Aizawa, K. and Gillett, C. (2009). The (multiple) realization of psychological and other properties in the sciences. *Mind and Language*, *24*, 181–208.

Andreasen, N. C. (1984). *The broken brain: The biological revolution in psychiatry*. New York: Harper & Row.

Anonymous (2013). No dishonour in depression. [Editorial]. *Nature, 498*, 137.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baker, L. R. (1993). Metaphysics and mental causation. In: J. Heil & A. Mele (Eds.). *Mental causation*. Oxford: Clarendon Press, 75-96.

Baker, L. R. (1995). Explaining attitudes: a practical approach to the mind. Cambridge University Press.

Barabási, A.-L. (2012). The network takeover. *Nature Physics*. doi:10.1038/nphys2188

Beard, C., Millner, A. J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M. T., Leonard, C. V., Kertz, S.J , & Björgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, 2016, 46, 3359-3369.

Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: linking cognitive and neural states. *Philosophy of Science*, *66*, 175-207.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*, 969-977.

Bentall, R. (2003). *Madness explained: Psychosis and human nature.* London: Penguin.

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology, 9*, 1089-1108.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*, 5-13.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*.

Borsboom, D., Epskamp, S., Kievit, R. A., Cramer, A. O. J., & Schmittmann, V. D. (2011). Transdiagnostic networks: A comment on Watkins and Nolen-Hoeksema. *Perspectives on Psychological Science, 6*, 610-614.

Boschloo, L., van Borkulo, C. D., Rhemtulla, M., Keyes, K. M., Borsboom, D., & Schoevers, R. (2015). The network structure of symptoms of the Diagnostic and Statistical Manual of Mental Disorders. *PLoS ONE, 10*, e0137621.

Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford: Oxford University Press.

Boyd, R. (1991). Realism, antifoundationalism and the enthusiasm for natural kinds. *Philosophical Studies, 61*, 127–148.

Boyd, R. (1999). Homeostasis, species, and higher taxa. In: *Species: New Interdisciplinary Essays* (Ed. R. A. Wilson), pp. 141–185. Cambridge, MA: MIT Press.

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

Brentano, F. (1874). *Psychology from an empirical standpoint*. London: Routledge and Kegan Paul.

Bringmann, L.F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013) A network approach to psychopathology: new insights into clinical longitudinal data. *PLoS One 8(4)*: e60188.

Campbell, J. (2009). What does rationality have to do with psychological causation? Propositional attitudes as mechanisms and as control variables. In: Broome, M., & Bortoletti, L. (Eds.). *Psychiatry as cognitive neuroscience: Philosophical perspectives*. Oxford: Oxford University Press, 137-149.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7-19.

Cramer, A. O. J., Kendler, K. S., & Borsboom, D. (2011). Where are the genes? The implications of a network perspective on gene hunting in psychopathology. *European Journal of Personality, 25*, 270-271.

Cramer, A. O. J., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS ONE, 11*, e0167490.

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: a network perspective. *Behavioral and Brain Sciences, 33*, 137-193.

Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford: Clarendon Press.

Deboer, T., Détári, L., & Meijer, J. H. (2007). Long term effects of sleep deprivation on the mammalian circadian pacemaker. *Sleep, 30*, 257-62.

Dennett, D. C. (1987). *The intentional stance.* Cambridge: MIT Press.

Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, *88*, 27-51.

Dennett, D.C. (1997). *Kinds of minds. Towards an understanding of consciousness*. New York: Basic Books.

Dori, G. A., & Overholser, J. C. (1999). Depression, hopelessness, and self-esteem: Accounting for suicidality in adolescent psychiatric inpatients. *Suicide and Life-Threatening Behavior*, *29*, 309-318.

Dretske, F. I. (1997). *Naturalizing the mind*. Cambridge, MA: MIT Press.

Endicott, R. P. (1993). Species-specific properties and more narrow reductive strategies. *Erkenntnis*, *38*(3), 303-321.

Elton, M. (2003). *Daniel Dennett. Reconciling science and our self-conception.* Cambridge: Polity Press.

Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (*in press*). Network Psychometrics. To appear in: Irwing, P., Hughes, D., & Booth, T. (Eds.), *Handbook of Psychometrics.* New York: Wiley.

Eronen, M. I. (2013). No levels, no problems: Downward causation in neuroscience. *Philosophy of Science*, *80*, 1042-1052.

Eronen, M. I. (2017). Interventionism for the intentional stance: True believers and their brains. *Topoi*. https://doi.org/10.1007/s11245-017-9513-5

Fodor, J. A. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese*, *28*, 97-115.

Franić, S., Dolan, C. V., Borsboom, D., & Boomsma, D. I. (2012) Structural equation modeling in genetics. In: Hoyle, R. (Ed.). Handbook of Structural Equation Modeling. New York: Guilford Press.

Francken, J. C., & Slors, M. (2014). From commonsense to science, and back: The use of cognitive concepts in neuroscience. *Consciousness and Cognition, 29*, 248-258.

Fried, E. I., & Cramer, A. O. J. (*in press*). Moving forward: challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science.*

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*, 1-10.

Fried, E. I., Bockting, C., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, A., Epskamp, S., Tuerlinckx, F., Carr, D., & Stroebe, M. (2015). From loss to loneliness: The relationship between bereavement and depressive symptoms. *Journal of Abnormal Psychology*, *124*, 256-265.

Gillett, C. (2003). The metaphysics of realization, multiple realization and the special sciences. *Journal of Philosophy*, *100*, 591–603

Greenberg, G. (2013). *The book of woe: The DSM and the unmaking of psychiatry*. Penguin Group US.

Guloksuz, S., Pries L.K., & van Os, J. (2017). Application of network methods for understanding mental disorders: pitfalls and promise. *Psychological Medicine,* 1-10.

Guze, S. B. (1989). Biological psychiatry: is there any other kind? *Psychological Medicine, 19*, 315-323.

Hacking, I. (1999). *The social construction of what?* Cambridge: Harvard University Press.

Haroz, E. E, Bolton, P., Gross, A., Chan, K. S, Michalopoulos, L., & Bass, J. (2016). Depression symptoms across cultures: an IRT analysis of standard depression symptoms using data from eight countries. *Social Psychiatry and Psychiatric Epidemiology*, *51*, 981-91.

Haroz, E. E., Ritchey, M., Bass, J. K., Augustinavicius, J., Michalopoulos, L., Burkey, M. D., & Bolton, P. (2017). How is depression experienced around the world? A systematic review of qualitative literature. *Social Science and Medicine*, *183*, 151-162.

Hoogman, M., Bralten. J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S. J., …, & Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: A cross-sectional mega-analysis. *The Lancet Psychiatry*, *4*, 310-319.

Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology. In S. Wagner and R. Warner (eds.), *Naturalism: A critical appraisal*. Notre Dame: University of Notre Dame Press, 295-320.

Hurn, J. D. (1998). *The history of general paralysis of the insane in Britain, 1830 to 1950*. Unpublished Ph.D. thesis, University College London. URL: http://discovery.ucl.ac.uk/1349281/1/339949.pdf

Hyland, M. E. (2011). *The origins of health and disease.* Cambridge University Press, Cambridge.

Insel, T. R., & Cuthbert, B.N. (2015). Brain disorders? Precisely. *Science*, *348*, 499-500.

Isvoranu, A. M.**,** van Borkulo, C. D., Boyette, L. L., Wigman, J. T. W., Vinkers, C. H., Borsboom, D. (2017). A network approach to psychosis: Pathways

between childhood trauma and psychotic symptoms. *Schizophrenia Bulletin, 43,* 187-196.

Jaspers, K. (1923/1963). *General psychopathology*. Manchester: Manchester University Press.

Kalis, A. (2014). Mentale toestanden in de psychologie. [Mental states in psychology.] *Algemeen Nederlands Tijdschrift voor Wijsbegeerte, 106,* 197-206.

Kalis, A. (2011). *Failures of agency. Irrational behavior and self-understanding*. Lanham: Lexington Books.

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

Kendler, K. S. (2008). Explanatory models for psychiatric illness. *American Journal of Psychiatry, 165,* 695-702.

Kendler, K. S. (2012). Levels of explanation in psychiatric and substance use disorders: implications for the development of an etiologically based nosology. *Molecular Psychiatry, 17,* 11-21.

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

Kraepelin, E. and Lange, J. (1927) *Psychiatrie, 9th Edition, Vol. I.* Leipzig: J. A. Barth.

Kievit, R.A., Romeijn, J.W., Waldorp, L.J., Wicherts, J. M., H. Steven Scholte & Borsboom, D. (2011) Mind the gap: A psychometric approach to the reduction problem. Psychological Inquiry, 22, 67-87.

Kim, J. (1982). Psychophysical supervenience. *Philosophical Studies, 41,* 51–70.

Kim, J. (1984). Concepts of supervenience. *Philosophy and Phenomenological Research, 45,* 153–176.

Ledford, H. (2014). Medical research: If depression were cancer. [Editorial]. *Nature, 515,* 182-184.

Lacasse, J. R., & Leo, J. (2005). Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine, 2,* e392.

Lewis, D. K. (1966). An argument for the identity theory. *The Journal of Philosophy,* 17-25.

McNally, R. J. (1990). Psychological approaches to panic disorder: A review. *Psychological Bulletin, 108,* 403-419.

McNally, R. J., (2002). Anxiety sensitivity and panic disorder. *Biological Psychiatry, 52,* 938-946.

McNally R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy,* 86, 95–104.

McNally, R. J., Heeren, A., & Robinaugh, D. J. (2017). A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *European Journal of Psychotraumatology, 8*(sup3), 1341276. http://doi.org/10.1080/20008198.2017.1341276

McNally, R. J,, Mair, P., Mugno, B., & Riemann, B. (2017). Co-morbid obsessive–compulsive disorder and depression: A Bayesian network approach. *Psychological Medicine, 47,* 1204-1214. doi:10.1017/S0033291716003287

Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., van Bork, R., Waldorp, L.J., van der Maas, H.L.J. & Maris, G.K.J (2017). An introduction to Network Psychometrics: Relating Ising network models to item response theory

models. *Multivariate Behavioral Research.* DOI:10.1080/00273171.2017.1379379

Ma-Kellams, C. (2014). Cross-cultural differences in somatic awareness and interoceptive accuracy: A review of the literature and directions for future research. *Frontiers in Psychology*, *5*, 1379.

Marmanidis, H., Holme, G., & Hafner, R. J. (1994). Depression and somatic symptoms: A cross cultural study. *Australian & New Zealand Journal of Psychiatry*, *28*, 274-278.

McDowell, J. (1996). *Mind and world*. Cambridge: Harvard University Press.

McGinn, C. (1982). *The character of mind.* Oxford: Oxford University Press.

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, *86*, 95-104.

Murphy, D. (2005). *Psychiatry in the scientific image*. Cambridge, MA: MIT press.

Menary, R. (Ed.). (2010). *The extended mind*. Cambridge: MIT Press.

Mölder, B. (2010). *Mind ascribed: An elaboration and defence of interpretivism*. Amsterdam: John Benjamins Publishing.

Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation.* London, UK: Routledge & Kegan Paul.

Nolen-Hoeksema, S., & Watkins E. R. (2011). A heuristic for developing transdiagnostic models of psychopathology: Explaining multifinality and divergent trajectories. *Perspectives on Psychological Science*, *6*, 589–609.

Nordenfeld, L. (2007). *Rationality and compulsion.* Oxford: Oxford University Press.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. In: *Minnesota Studies in the Philosophy of Science*, *2*, 3–36.

Partch, C. L., Green, C. B., & Takahashi, J. S. (2013). Molecular architecture of the mammalian circadian clock. *Trends in Cell Biology, 24*, 90-99.

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Rachman, S. (1998). A cognitive theory of obsessions: Elaborations. *Behaviour Research and Therapy, 36*, 385-401.

Redei, E. E., Andrus, B. M., Kwansy, M. J., Seok, J., Ho, J., & Mohr, D. C. Blood transcriptomic biomarkers in adult primary care patients with major depressive disorder undergoing cognitive behavioral therapy. *Translational Psychiatry, 4*, e442. doi:10.1038/tp.2014.66

Reise, S.P. and Waller, N.G. (2009). Item Response Theory and clinical measurement. *Annual Review of Clinical Psychology*, 5, 27–48.

Reiss, S. (1991). Expectancy model of fear, anxiety, and panic. *Clinical Psychology Review, 11*, 141-153.

Rhemtulla, M., Fried, E. F., Aggen, S., Tuerlinckx, F., Kendler, K. S., & Borsboom, D. (*in press*). Network analysis of substance abuse and dependence symptoms. *Drug and Alcohol Dependence*.

Robb, D. & Heil, J. (2013). Mental causation. *The Stanford Encyclopedia of Philosophy* (Spring 2013), E.N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2013/entries/mental-causation

Robinaugh, D.J., Leblanc, N.J., Vuletich, H.A., McNally, R.J., 2014. Network Analysis of persistent Complex Bereavement Disorder in conjugally bereaved adults. *Journal of Abnormal Psychology, 123*, 510–522.

Salkovskis, P. M., Wroe, A. L., Gledhill, A., Morrison, N., Forrester, E., Richards, C., Reynolds, M., & Thorpe, S. (2000). Responsibility attitudes and

interpretations are characteristic of obsessive compulsive disorder. *Behaviour Research and Therapy*, *38*, 347-372.

Searle, J.R. (1983). *Intentionality.* Cambridge: Cambridge University Press.

Searle, J. R. (1998). *Mind, language and society: Philosophy in the real world*. New York: Basic Books.

Schaffner, K. (1974). Reductionism in biology: Prospects and problems. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 613-632.

Schaffner, K. (2006). Reduction: the Chesire cat problem and a return to roots. *Synthese*, *151*, 377-402.

Schaffner, K. F., & Tabb, K. (2014). Varieties of social constructionism and the problem of progress in psychiatry. *Philosophical Issues in Psychiatry III: The nature and sources of historical change.* Oxford: Oxford University Press.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009) Early-warning signals for critical transitions. *Nature*, 461, 7260, 53-59.

Solomon, A. (2014). The noonday demon. New York: Scribner.

Tabb, K., & Schaffner, K. F. (2017). Causal pathways, random walks and tortuous paths: Moving from the descriptive to the etiological in psychiatry. In: Kendler, K. S., & Parnas, J. (Eds.): *Philosophical Issues in Psychiatry IV: Nosology*. Oxford: Oxford University Press.

Thornton, T. (2010). Psychiatric explanation and understanding. *European Journal of Analytic Philosophy*, *6*, 95-111.

Tolin, D. F., Brady, R. E., & Hannan, S. (2008). Obsessional beliefs and symptoms of obsessive–compulsive disorder in a clinical sample. *Journal of Psychopathology and Behavioral Assessment*, *30*, 31-42.

Van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports, 4*, 5918.

Van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*, 842-861.

Van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., Van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen S. H., Derom, C., Jacobs, N., Kendler, K. S., Van der Maas H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2013). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1312114110

Van Os, J. (2009). 'Salience syndrome' replaces 'schizophrenia' in DSM-V and ICD-11: Psychiatry's evidence-based entry into the 21st century? *ACTA Psychiatrica Scandinavica, 120*, 363-372.

White, K. S., Brown, T. A., Somers, T. J., & Barlow, D. H. (2006). Avoidance behavior in panic disorder: The moderating influence of perceived control. *Behaviour Research and Therapy*, *44*, 147-157.

Wichers, M. (2014). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, *44*(07), 1349-1360.

Wild, L. G., Flisher, A. J., & Lombard, C. (2004). Suicidal ideation and attempts in adolescents: associations with depression and six domains of self-esteem. *Journal of Adolescence*, *27*, 611-624.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality.* Harvard: Harvard University Press.