Behavioral and Brain Sciences

cambridge.org/bbs

Brain disorders? Not really: Why network structures block reductionism in psychopathology research

Denny Borsboom,[a] Angélique O. J. Cramer,[b] and Annemarie Kalis[c]

[a]Department of Psychology, University of Amsterdam, 1018 WT Amsterdam, The Netherlands; [b]Department of Methodology and Statistics, Tilburg University, 5000 LE Tilburg, The Netherlands; and [c]Department of Philosophy and Religious Studies, Utrecht University, 3512 BL Utrecht, The Netherlands

d.borsboom@uva.nl https://www.dennyborsboom.com
aoj.cramer@gmail.com https://www.aojcramer.com
A.Kalis@uu.nl https://www.uu.nl/staff/AKalis

Target Article

Cite this article: Borsboom D, Cramer AOJ, Kalis A. (2019) Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences* **42**, e2: 1–63. doi:10.1017/S0140525X17002266

Target Article Accepted: 24 December 2017
Target Article Manuscript Online: 24 January 2018
Commentaries Accepted: 27 June 2018

Keywords:
networks; philosophy; psychometrics; psychopathology; reductionism

Abstract

In the past decades, reductionism has dominated both research directions and funding policies in clinical psychology and psychiatry. The intense search for the biological basis of mental disorders, however, 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 target article, we show that this conceptualization can help explain 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; 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 depend partially 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.

1. 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 article.

Many believe that symptoms, signs, and other problems associated with mental disorders – for example, 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 be explained ultimately in terms of specific dysfunctional neurobiological conditions – is partly because 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, another 2013 *Nature* editorial concluded that "despite 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. Currently, there is 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 – that is, 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 article aims to show, following through the logic of recently proposed network models of psychopathology (Borsboom 2017; Borsboom & Cramer 2013; Cramer et al. 2010) 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. This is 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 on 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. It *does* mean, however, 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.

Denny Borsboom is Professor of Psychological Methods at the University of Amsterdam. His work has focused on conceptual analyses of psychometric models, the measurement problem in psychology, and on substantive psychological research in a number of domains, including intelligence, personality, and psychopathology. He is the founder of the Psychosystems project (www.psychosystems.org), which is specifically dedicated to the development of network approaches to psychometric problems, including the construction of statistical models and formalized psychological theories.

Angélique Cramer is Associate Professor of Methodology and Statistics at Tilburg University. She is a key developer of the network perspective on psychopathology, with more than 40 published papers in internationally renowned journals. Funded by a personal grant (VENI) and a research grant from the Dutch Cancer Society, she currently works on network theory and methodology, as well as the implementation of network modeling in (clinical) practice. Dr. Cramer is the director of the Tilburg Experience Sampling Center, was a fellow at the Dutch Institute for Advanced Study (NIAS), and was recently appointed a member of the Young Dutch Academy.

Annemarie Kalis is Assistant Professor in Practical Philosophy at Utrecht University. She has a double background in psychology and philosophy and received her Ph.D. in 2009 with a thesis on weakness of will. Her main areas of research are action theory, and the philosophy of psychology and psychiatry. She is the author of the monograph *Failures of Agency* (Lexington Books, 2011) and has written various papers on the philosophy of psychology. In 2014, she was awarded a personal research grant of the Netherlands Organization for Scientific Research, on a project on the metaphysics and measurement of attitudes.

2. 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; Fodor 1974; Kievit et al. 2011; Nagel 1961; Oppenheim & Putnam 1958; Schaffner 1974). 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 identifying the cause of symptomatology (higher-level concept) with the bacteria *Treponema pallidum* (lower-level concept[2]). In general, explanatory reductionism, in psychiatry, 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 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 therefore insufficient to merely identify neural correlates of psychiatric conditions. This is the case for at least three reasons:

1. 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; American Psychiatric Association 2013). This means the lower-level description piggybacks on the higher-level description instead of, as explanatory reductionism requires, the other way around.

2. Given the plausible rejection of Cartesian substance dualism, *some* neural correlate is guaranteed to exist for any behavioral measure. As a result, 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.
3. It is unclear for most correlates 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, 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 boldest, most prominent expression of this thesis appears 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 evident in a 2014 *Nature* editorial (Ledford 2014) which, when comparing depression to cancer, notes 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 it is also evident 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 an established fact. For instance, Hoogman et al. (2017) state 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 cases 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 the lab tests that should be able to identify them (e.g., see Redei et al. 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 treatment plans that restore or ameliorate these abnormalities and, as a result, remove 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.

3. 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; 2012a). Instead, many have come to accept the ideas that (1) mental disorders are massively multifactorial in their causal background; (2) many mechanisms that sustain disorders are transdiagnostic; and (3) mental disorders require pluralist explanatory accounts (Borsboom et al. 2011; Kendler 2008; Nolen-Hoeksema & Watkins 2011). Kendler et al. (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* (Borsboom 2017; Borsboom & Cramer 2013; Cramer et al. 2010; Fried & Cramer 2017; Fried

et al. 2017). 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 (a symptom of psychosis) and sad mood (a 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. 2017), 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 (Boschloo et al. 2015; Cramer et al. 2010) and between such symptoms and external stressors (Borsboom 2017; Isvoranu et al. 2017; McNally et al. 2017a). In addition, it offers plausible accounts of several other psychiatric phenomena such as comorbidity, spontaneous recovery, and heterogeneity (Borsboom & Cramer 2013; Cramer et al. 2016; Fried et al. 2017, McNally et al. 2017b), and is compatible with theories and models that focus on the dynamic interplay among 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 to which it is directly related (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 more problems; teaching a person how to control rumination, so that late-night worries no longer lead to insomnia; Borsboom 2017). Importantly, just as symptoms can arise from both internal and external causes, interventions on symptoms can involve 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, or 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 2017; Borsboom & Cramer 2013; Fried & Cramer 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 article.[4]

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 (Borsboom & Cramer 2013; Cramer et al. 2010; Hyland 2011). 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. For suppose that the correlations between symptoms are in fact produced by their common dependence on a latent variable. In that case, equating that latent variable with a brain disorder is tantamount to achieving the hallmark event of explanatory reductionism, namely cross-level identification. Thus, identifying the common cause of the symptoms (higher-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.

If a network model is correct, however, no such common cause exists, so the above theoretical move is blocked. In the network model, mental disorders do not resemble medical disease 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 & Cramer 2013). Instead, in the network definition, disorders behave more like *fields* – for example, 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. 2018)[6] and to stable states in complex ecosystems (Scheffer et al. 2009; Van de Leemput et al. 2013; van der Maas et al. 2006). The emergent global states into which networks can get locked 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

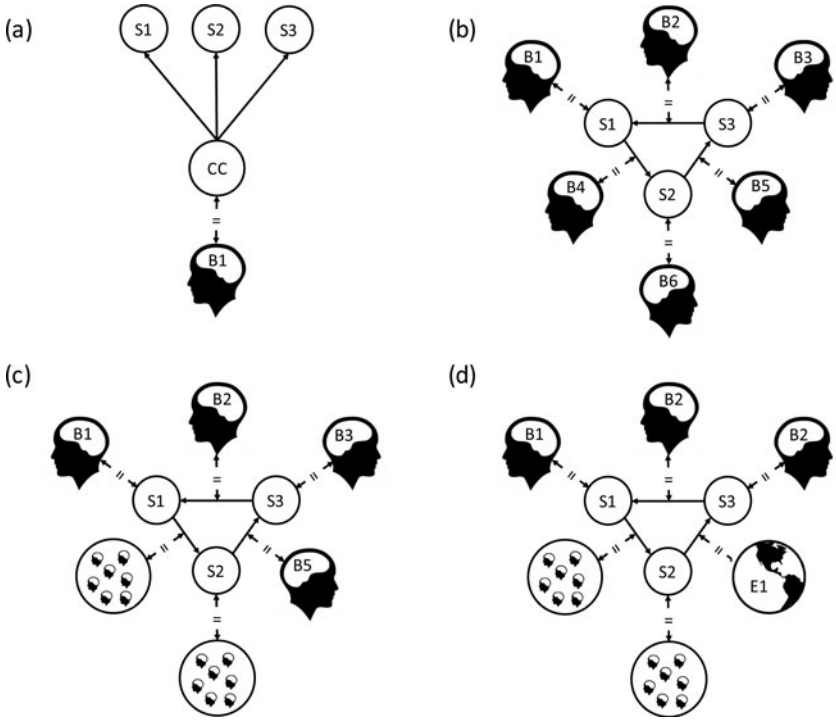

Figure 1. Four prospects for (partial) explanatory reduction in psychopathology. Panels (a) through (d) depict relations among symptoms S, brain states/mechanisms B, a common-cause CC, and external factors E. Identifications are marked "=." In panel (a), a *brain disorder* scenario, correlations among 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 among 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 and 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.

of the individual symptoms in a psychopathology network nevertheless corresponds to a 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 justifiably still be 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. Although in many cases connections between symptoms are indeed strongly grounded in biological mechanisms (e.g., insomnia → fatigue; appetite loss → weight loss, drug use → tolerance) that may support "local" reductions of certain parts of the symptom network ("patchy reductionism"; Kendler 2005; Schaffner 2006), there are other cases in which such grounding should not be suspected to support successful explanatory reduction. In the following, we show this by providing a detailed analysis of symptoms that involve the *content* of mental states.

5. 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, such as pains or undirected anxieties, may not (McGinn 1982; Searle 1983).

The symptomatology in systems such as *DSM*-5 (American Psychiatric Association 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" (American Psychiatric Association 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" (American Psychiatric Association 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 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 that 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 becomes visible only 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 (Abramowitz et al. 2006; Tolin et al. 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 the 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 explanatory only insofar 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 et al. 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 becomes 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 et al. 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 reasonable (at least to some extent) 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, based on the work of Donald Davidson (1984) and Daniel Dennett (1987), that is 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 ending his/her existence might be best. 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 be seen as an irrational deviation only insofar 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 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 previously, 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, which 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]

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 as 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 coded differently 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 houses), 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 a 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 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 him 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 & 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.

6. 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. This is 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 et al. 2016; Haroz et al. 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 World War II who has failed to defend his post to that of 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 et al. 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; likewise, 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. Because 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 instance, 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 previous 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 differ partially across place and time, which will give rise 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).

There is, however, 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 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 (American Psychiatric Association 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* – that is, it is not located in the person's head (see Clark & 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 hence 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 2005). 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, and 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 & Tabb 2014).

7. 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 and 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" (Dennett 1987, p. 27).

Dennett (1987; 1991) introduced the idea of *real patterns* in order to show how intentional explanations (which Dennett calls "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 of 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 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* 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. Instead, a mental state is characterized by the fact that it operates in the appropriate causal way in the symptom network (e.g., leads to compulsive handwashing). And it likely 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 compulsive only if it can be seen as an *excessive* response, and this cannot be determined by looking inside the person's head: What is excessive in a Western country is not excessive in a country ravaged by Ebola. This means that whether a behavior is to be considered symptomatic or normal depends not only on a person's psychological or neurobiological state, but also on the environment.

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 sides 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 mechanism), 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. 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 Fig. 1, panels a & d), and it would certainly be a great scientific discovery if such a reductive explanation were to be construed. However, given the 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

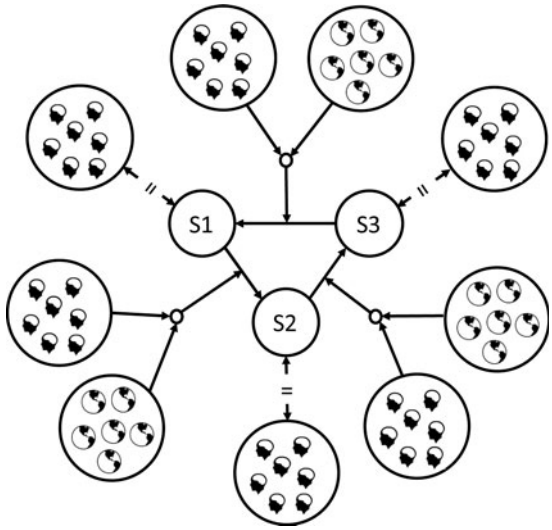

Figure 2. Massively multifactorial symptom networks. The figure shows a situation in which symptoms S1–S3 involve multiply realized mental states. Network connections depend on the combination of heterogeneous sets of biological and environmental conditions, as indicated by the concatenation symbol ¢,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

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. This is because the latter course of action may exclude lines of research that could yield crucial information about psychopathology networks. Examples include research regarding cultural variations in the contents of relevant mental states and into the role of the external environment in sustaining disorders, 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 as 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 in which we are ultimately interested. 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 taking 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, for example, 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.

8. 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 2017; Borsboom & Cramer 2013; Fried & Cramer 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 (Barabási 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 present article 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, for example, 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 2017). 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. Instead, one would search for genetic variants associated with *specific aspects* of a depression network structure (Cramer et al. 2011) – for instance, 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 on 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: for example, 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 et al. 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 et al. 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 if 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 Eronen [2013] and Wimsatt [2007]).

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. By 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 of 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. In that sense, they 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.

Notes

1. This characterization of reduction is somewhat simplified. One can question whether even these hallmark reductive explanations are in fact successful and, if so, in what sense (e.g., see Bickle 1998; Eronen 2013; Fodor 1974; Schaffner 1974; Wimsatt 2007). In the current context, however, 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.
3. https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2014/a-new-approach-to-clinical-trials.shtml.
4. Note that the current article does not argue that the network model is correct; plausibility arguments have been given elsewhere (e.g., see Borsboom 2017; Borsboom & Cramer 2013; Cramer et al. 2010; 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 et al. 2017), just that the common-cause explanation plays a very important role in medicine (Hyland 2011).
6. This comparison is not just metaphorical; the Ising model used to identify network structures (Van Borkulo et al. 2014) is, in fact, mathematically identical to models for magnetism and, as understood in terms of network dynamics, features largely the same phenomena (Cramer et al. 2016).
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 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 of the realizing conditions involved), its existence makes explanatory reductionism pragmatically unattractive.
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 be described (only) by means of the design 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). 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).
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 et al. 2012). This is because they may be invariant over individuals and, hence, may not produce individual differences at all.

Open Peer Commentary

Reductionist thinking and animal models in neuropsychiatric research

Nicole M. Baran

School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332.
nicole.baran@biology.gatech.edu https://www.nicolembaran.com

doi:10.1017/S0140525X18001231, e3

Abstract

Reductionist thinking in neuroscience is manifest in the widespread use of animal models of neuropsychiatric disorders. Broader investigations of diverse behaviors in non-model organisms and longer-term study of the mechanisms of plasticity will yield fundamental insights into the neurobiological, developmental, genetic, and environmental factors contributing to the "massively multifactorial system networks" which go awry in mental disorders.

The problems identified by Borsboom et al. with regard to reductionism in neuropsychiatry are evident in the widespread use of animal models of mental disorders in modern neuroscience. Neuroscience research focused on psychiatric disorders is dominated by studies using a small number of species – predominantly, artificially housed inbred strains of laboratory mice. This work uses selective breeding, genetic engineering (to produce transgenic lines or mutant knockouts), targeted lesioning of the brain, or manipulations of the environment to recapitulate the plausible causative factor(s) thought to underlie a given diagnosis, or at least the neural or behavioral pathologies which characterize the human disorder.

Unfortunately, these animal models often are found to have weak correspondence to the phenomenology of the neuropsychiatric disorder in question (weak validation), and drugs developed using these models often have limited efficacy (poor predictive validity) (Markou et al. 2009; Nestler & Hyman 2010). For example, despite hundreds of supposed mouse models of autism spectrum disorder (ASD), no pharmacological interventions have yet been found that markedly improve ASD's characteristic deficits in either social interaction or repetitive behavior (Kazdoba et al. 2016; Varghese et al. 2017). Furthermore, several widely used animal models of depression (e.g., tests of behavioral despair or learned helplessness, such as forced-swim or tail-suspension tests) map poorly to the pathologies of chronic depression, which also include symptoms such as anhedonia, disruptions of sleep, and changes to psychomotor behavior. And, of course, animal models will never be able to recapitulate symptoms that are central to many neuropsychiatric disorders: the content of mental states. It is hard to imagine a mouse that experiences rumination, guilt, shame, or existential ennui.

Unfortunately, the reductive emphasis on rodent models of human mental disorders in neuroscience may be hindering the development of more safe and effective interventions for psychiatric patients. While much can be learned from neuroscience research which attempts to understand the genetic, anatomical,

and molecular mechanisms underlying mental disorders, more phenomenologically and taxonomically broad efforts targeted at understanding what Borsboom and colleagues call "network structures" will provide key additional insights. Comparative research that focuses on diversity and variation between species, instead of merely attempting to phenocopy a human disorder in a single – albeit convenient – species, will be necessary to build a comprehensive understanding of the general principles of brain organization and development (Striedter et al. 2014).

The problem is that reductive research efforts are prioritized over more broad research efforts by funding agencies, journals editors, hiring committees, and the popular press. Incentives in neuroscience often reward technologically complex experiments focused on dissecting neural circuits, but not careful behavioral observation, long-term studies that investigate the developmental trajectories of behavior, or work that seeks to understand the environmental and evolutionary context in which any given behavior functions (see Krakauer et al. [2017] for a review).

A broader, holistic, more evolutionarily grounded approach may ultimately provide critical insights into the brain circuits, developmental processes, genetic mechanisms, and environmental factors which contribute to the "massively multifactorial system networks" that go awry in mental disorders. Indeed, some of the most exciting discoveries about the functioning of the brain have come from long-term neuroethological studies of "non-model" organisms, such as pair bonding in prairie voles, vocal learning in songbirds, spatial attention in owls, and social plasticity in cichlid fish (Knudsen 2011; Maruska & Fernald 2018; McGraw & Young 2010; Pfenning et al. 2014). This work is especially important because many key behaviors at the core of some neuropsychiatric disorders have no parallel in the behavioral repertoire of a rodent (e.g., language-learning deficits in ASD).

Additionally, greater emphasis should be placed on understanding the specific mechanisms by which the environment and experience (stress, trauma, interactions with caregivers, etc.) are translated into changes in the brain. Learning and responding to the environment is what brains do, and brains are profoundly shaped by experience at every stage of development. In addition to genetic correlates, the majority of neuropsychiatric disorders have profoundly important, yet woefully understudied, environmental etiologies. As such, researchers should not shy away from performing long-term developmental experiments to understand these mechanisms. Fortunately, there is a renewed interest in the neurobiological mechanisms of plasticity, including the genes, molecules, and epigenetic influences which regulate the sensitivity of organisms to environmental conditions (Baran 2017; Caspi et al. 2010; Meaney 2017). However, much of this work is still in its infancy.

Reductive approaches in neuroscience research have led to an extreme focus on trying to find biologically based interventions (i.e., drug development), despite the fact that we already know that behavioral and environmental interventions are critical components in the effective treatment of many human mental disorders. Taking a network structure approach suggests that we should both include a greater diversity of organisms and behaviors in neuroscience research, as well as study complex interactions between multiple factors at multiple levels of analysis. Borsboom et al. have identified a conceptual flaw at the heart of neuropsychiatric research; preclinical neuroscience researchers would do well to heed the warning.

Beyond trait reductionism: Implications of network structures for dimensional models of psychopathology

Robert F. Bornstein

Department of Psychology, Adelphi University, Garden City, NY 11530.
bornstein@adelphi.edu
http://www.adelphi.edu/faculty/profiles/profile.php?PID=0366

doi:10.1017/S0140525X18001243, e4

Abstract

Borsboom et al. discuss the implications of network structures for neurobiology-based reductionism, but inherent in the network approach is that dimensional models of psychopathology are untenable as well. Insofar as mental disorders are complex dynamic constellations of symptoms, the "trait reductionism" of dimensional psychopathology frameworks suffers from the same limitations as neurobiological reductionism.

Just as the existence of biological correlates of mental states does not imply that those biological processes impel the mental states with which they covary, the existence of trait correlates of mental states does not mean that these mental states are accounted for by trait descriptors. Inherent in the network approach of Borsboom et al. is that the "trait reductionism" of dimensional models of psychopathology cannot provide a complete explanation of psychological dysfunction and distress.

1. Neurobiological reductionism and trait reductionism

The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) and *International Classification of Diseases* (*ICD*) have traditionally conceptualized psychological disorders categorically, but in recent years limitations of polythetic diagnoses based on combinations of symptoms have become increasingly apparent (e.g., heterogeneity within categories, excessive comorbidity across syndromes; see Herpertz et al. 2017). As a result, dimensional models of psychopathology have been presented as a psychometrically superior approach. Typically, these models conceptualize psychological syndromes using a series of trait dimensions, with patients assigned severity ratings that are combined into a trait profile to capture the central elements of a particular disorder. The *Hierarchical Taxonomy of Psychopathology* (HiTOP) is the most influential dimensional model currently in ascendance (Kotov et al. 2017), but others have been developed as well (e.g., Krueger & Markon 2014). It is clear that the next editions of the diagnostic manuals will implement dimensional frameworks for conceptualizing and diagnosing psychological disorders.

Like neurobiological models, trait models make intuitive sense to researchers and laypersons. Like neurobiological models, trait models capture one element of psychological dysfunction, but they cannot explain psychological disorders in all of their complexity. Borsboom et al.'s network model not only provides important context for understanding the strengths and limitations of neurobiological models, but of trait models as well.

2. Three neglected issues

Kotov et al.'s assertion that "not a single mental disorder has been established as a discrete categorical entity" (Kotov et al. 2017, p. 457) requires an exceedingly narrow definition of "categorical" (i.e., a wholesale shift in impairment once a threshold is met, along with virtually no symptomatic overlap or comorbidity with other syndromes). A number of categorical diagnoses in medicine (e.g., hypertension, autoimmune disorders, various pulmonary conditions) and psychology (e.g., anorexia, depression, narcissism) have imperfect thresholds and fuzzy boundaries, but are nonetheless clinically useful (see Hutchinson & Romero 2016; Lekkas & Mikhailov 2010). Kotov et al.'s (2017) statement illustrates the importance of taking a more nuanced and critical stance in evaluating the strengths and limitations of different frameworks for conceptualizing and quantifying psychological dysfunction. Three issues are crucial in this effort.

2.1. A focus on process

Like neurobiological models, trait models do not capture cultural influences on the experience and expression of psychological distress, nor do they make *a priori* predictions regarding elements of the environment that exacerbate psychopathology and moderate its severity. Studies using ambulatory assessment techniques have documented trait-driven contextual variations in responding (Wright et al. 2015); however, a complete understanding of psychological dysfunction requires use of experimental methods in which aspects of the environment are manipulated and the impact of these manipulations on cognition, behavior, and affect is assessed. Such studies exist (e.g., Bornstein 2011; Horvath & Morf 2009), and while their results are not easily accommodated by trait models, they fit well with the network approach which specifies relations among elements in a network and generates predictions regarding how altering one feature of that network will affect other elements of the system.

2.2. Reification and misplaced concreteness

Just as the existence of biological correlates of mental states does not imply that those correlates account for the mental states with which they covary, the existence of trait correlates of mental states does not mean that mental states are accounted for by traits. Resistance to acknowledging the metaphoric underpinnings of neurophysiological constructs (e.g., "brain activity"; "cortical arousal") and psychological concepts (e.g., "disinhibition"; "neuroticism") can lead to "the error of misplaced concreteness" (Gargiulo 1998, p. 416) wherein descriptors are treated as immutable entities rather than descriptive labels. The search for useful metaphors is a key feature of scientific inquiry that cuts across disciplines (e.g., *natural selection*, *molecular bonding*, *working memory*), but these labels can have negative consequences as well (Bornstein & Becker-Matero 2011). Just as certain neurobiological metaphors (e.g., "neural circuits") are so familiar that they have taken on the appearance of concrete reality, the same is true of widely used trait metaphors (e.g., "facets of antagonism").

2.3. Statistical integration and cross-method generalizability

Studies examining the underlying structure of psychological dysfunction using taxometric procedures suggest that continuous latent variables outperform categorical latent variables with respect to model fit and replicability (Haslam et al. 2012). When other statistical methods are used (e.g., latent class analysis, finite mixture modeling), some symptom clusters and dysfunctional behavior patterns coalesce into replicable individual differences that have qualities of network structures (Hallquist & Pilkonis 2012; Yun et al. 2013). Psychopathology researchers must use a broader range of methodologies to capture patient functioning on multiple levels, including dynamics not amenable to self-report. Use of contrasting assessment methods that engage different psychological processes should be the norm within, as well as between, studies (Hopwood & Bornstein 2014).

3. The integrative potential of network structures: Beyond psychological dysfunction

As Borsboom et al. have noted, insofar as mental disorders arise from the causal interplay among symptoms, it is unlikely that psychological syndromes can be explained conclusively in neurobiological terms. The same is true of trait frameworks, which do not provide a complete explanation of complex psychological phenomena. Looking ahead, network structure models may be a useful integrative framework linking psychological dysfunction with adaptation and impairment in other domains as well. Findings have documented that psychological and physiological processes interact to influence illness risk and moderate disease outcome (e.g., Cromer & Villodas 2017); as network structure models expand across multiple contexts, they have the potential to enhance our understanding of the intra- and interpersonal dynamics that moderate health and illness – biological as well as psychological.

The adaptive self: Culture and social flexibility in feedback networks

Daina Crafa[a] and Saskia K. Nagel[b,c]

[a]Integrated Program in Neuroscience, Douglas Mental Health Institute, McGill University, Montreal, Quebec, H4H 1R3, Canada; [b]Human Technology Center, RWTH Aachen University, 52062 Aachen, Germany; and [c]Department of Philosophy, University of Twente, 7500 AE Enschede, The Netherlands.
daina.crafa@mail.mcgill.ca
saskia.nagel@humtec.rwth-aachen.de
http://www.dainacrafa.com

doi:10.1017/S0140525X18001255, e5

Abstract

Culture contextualizes the contents and intentionality of many mental statuses. Cognitive mediation of cultural information shapes these contents and intentionalities, as well as many of the false beliefs of pathology. Flexibility of cognitive mediation processes and resulting beliefs and pathologies may vary by individual, be a key mechanism of the feedback loop, and help characterize network connections.

Many psychiatric disorders may be better explained by nonreductive, network models, as exemplified by Borsboom et al. in the target article, than by reductionist models that condense or dismiss distinct but interrelated cognitive processes. The influences of

culture and individual social flexibility on mental disorders and outcomes are among the processes that network models can robustly account for. We agree with the authors' descriptions of the interplay between mental state contents and culture in symptom networks. To further enrich the discussion, we examine (1) the relevant details of the relationships between mental state contents and culture; (2) how these relationships might arise and feed back into themselves; and (3) the implications these relationships have for further defining appropriate network models.

Culture influences the causal relationship between *symptoms* in a network model. Culture also largely shapes the *contents* of mental states and *intentionality* of many mental states (e.g., Ambady & Bharucha 2009; Ramos-Sánchez & Atkinson 2009; Varela & Shear 1999). The network relationships between symptoms appear causally related once their contents and intentionality are understood (Borsboom et al. in the target article, and see also Borsboom 2008). In the example by Borsboom et al. in the section on "the content of mental states" (sect. 5), someone who believes the Central Intelligence Agency (CIA) is watching will close the curtains. According to their line of argument, the CIA watching would be part of the cultural influence of being a patient in a country where the CIA is a relevant entity, such as the United States; however, the intentional act of closing the curtains and withdrawing from social life is also part of the cultural effects. These actions may be considered comprehensible in the United States, for example, because of the cultural interpretation of how to obtain safety from governmental agencies. Alternative actions, such as becoming hypersocial to find safety by remaining in a large group of people, may be equally understandable in a cultural context where groups are perceived as safer than isolation. This interpretation harmonizes recent findings in cultural psychiatry, such as variations in hallucination experiences of psychotic patients (Luhrmann et al. 2014), with findings in neuropsychiatry and on real-world patient behaviors (Bowie et al. 2008; Menon & Uddin 2010).

Culture, of course, is characterized by numerous continua of common beliefs and behaviors (Kemmelmeier & Kühnen 2012), and one must also ask how an individual decides which culturally informed behaviors to adhere to. Cognitive mediation of cultural information is arguably a key process that gives rise to the contents and intentionality of mental contents as well as many of the false beliefs that contribute to pathology (Crafa & Nagel 2013; in press; Kitayama & Uskul 2011). Cognitive mediation in this context refers to the process of identity and belief construction based on consciously or unconsciously subscribing to or rejecting information in the sociocultural environment. Because social information is continuously encountered throughout the lifespan, this process is a feedback loop that constantly feeds into itself, but that also shapes the biological, neurological, and psychological constitution of the individual (Crafa & Nagel 2013; in press). New social experiences may reinforce or alter existing biological, neural, and psychological processes by providing both information and impetus. For example, the reification or undermining on the levels of existing beliefs, neural networks, or behavioral outputs changes what information in the social environment is experienced (e.g., other people respond differently to you depending on how you think and behave) and perceived (e.g., if you believe the CIA is after you, you will attend to different environmental information and interpret that information differently than you might otherwise) (Archpru Akaka & Chandler 2011). The information in the social environment that is experienced and perceived then feeds back, either reinforcing or undermining existing processes, and the cycle continues *ad mortem*. This feedback loop shapes neural networks, behavioral outputs, and other biopsychological processes.

When examining the relationship between mental disorders and cognitively mediated feedback loops, individuals vary substantially in their abilities to adapt to novel or dynamically changing social situations (Folke et al. 2010). Social rigidity and hyperflexibility are symptoms of many disorders, such as obsessive-compulsive disorder, autism, and schizophrenia (Bliksted et al. 2014; Chamberlain et al. 2009; Geurts et al. 2009). From the perspective of a network model, it is useful to consider rigidity and hyperflexibility as parts of a single continuous trait of social flexibility. Where an individual falls on this continuum of social flexibility is informative for understanding whether exposure to new social information will reinforce or undermine existing processes to larger or smaller extents. Thus, in turn, an individual's degree of social flexibility may indicate how mental contents or intentionality might continue to develop across the lifespan and the magnitude of those changes. In other words, not all feedback loops are created equal. Considering the role of social flexibility as a key feature of any feedback loop can help further characterize the development, strength, and possible trajectories of network connections, and further specify how we can understand the complex reasoning of individual patients as well as the relationship between their reasoning and their underlying neurobiology.

The impact of culture on the contents of mental states and how a person cognitively mediates those experiences may vary depending on how flexible a person is. Understanding variations in human flexibility can be informative for characterizing and potentially predicting the impact network relations may have on the trajectory of individual mental states. Network models of psychiatric disorders will benefit from the inclusion of these interrelated processes in order to ultimately better understand the patient.

Beyond reduction with the representation: The need for causality with full complexity to unravel mental health

Martin Desseilles[a,b,c,d] and Christophe Phillips[a,e]

[a]GIGA Cyclotron Research Centre (CRC) in vivo imaging, University of Liege, Liège B-4000, Belgium; [b]Department of Psychology, University of Namur, Namur B-5000, Belgium; [c]Clinique Psychiatrique des Frères Alexiens, Henri-Chapelle B-4841, Belgium; [d]Transition Institute, University of Namur, Namur B-5000, Belgium; and [e]GIGA in silico medicine, University of Liege, Liege B-4000, Belgium.

martin.desseilles@unamur.be c.phillips@uliege.be
http://mentalhealthsciences.com/index_en.html
http://www.giga.uliege.be/cms/c_17732/en/home

doi:10.1017/S0140525X18001267, e6

Abstract

In this commentary on Borsboom et al.'s target article, we argue that researchers should be aware of the historical development of models in neuroscience. Considering the importance of causality in anatomo-clinical approach and stressing the complexity of mental phenomenon, we provide new insight on reductionism and representation limitation.

In the course of neuroscience history, and despite the multiplicity of studies carried out, the physiology of the nervous system was often conceived and developed along two distinct ways. These two paths coexisted for many years and were already in the work of Herbert Spencer (1864–1867). Between, on the one hand, the reaction or the reflex and, on the other, the spontaneous activity and the activity of the psyche, the evolution of the physiology of the nervous system reflects two conceptions of neuroscience. First:

> With Ivan Sechenov, Claude Bernard, Charles Richet, and Ivan Pavlov, the study of psychic reflexes leads to the definition of the concept of conditioning as an adaptive learning mechanism, by strengthening a permanent association between a conditioned stimulus and a physiological response, whose function is anticipation …. It is with the rise of cybernetics, after the Second World War, and the central role of France in the East-West rapprochement in neurophysiology, that this line of research leads to the definition of adaptive neural mechanisms of learning as strengthening synapses. (Barbara 2008, pp. 2–3)

Second:

> In an opposite way, biologists, ethologists, psychologists and neurologists characterize the adaptation of organisms by structured and innate psychic processes, which are part of the history of animal species, and not only of interactions with the animal's environment. In a Spencerian spirit, the British neurologist of the second half of the nineteenth century, John Hughlings Jackson (1835–1911), proposes hierarchical and organized psychic functioning, that is to say, elaborated over generations, and which can undergo degradations during pathological phenomena. This model leaves room for reflexes and automatics, but mainly describes sensory-motor integration and coordination at a higher level by prefrontal areas…. Jackson comes to admit that the study of the intellect is distinct and parallel to that of reflexes. (Barbara 2008, p. 3)

Beside these two paths, Francisco Varela (1946–2001) proposed his theory of enaction. This other adaptation model proposes that cognition is the permanent production of the world that emerges in the subject through the establishment of neuronal connections during a history that is not interrupted (Varela et al. 1991; cf. Barbara 2008). Thus, enaction theory conceives of the mind by emphasizing how the body and mind organize themselves and interact with the environment (Varela et al. 1991). In the same line, Stanislas Dehaene proposed that neural structures might serve predefined different fundamental functions such as reading languages and adapting to particular forms of writing (Dehaene 2007). Similarly, Collignon and colleagues showed that a new shape of the functional architecture and the connectivity of the visual cortex could take place during developmental periods of visual deprivation (Collignon et al. 2013). So, re-use of neurons would be a larger general phenomenon consisting in the diversion of cognitive functions formerly used for other purposes towards a new use in the context where the environmental conditions are new (Barbara 2008, de Ricqlès 2015). This "neural recycling" of Dehaene or the reshaping of Collignon et al. is similar to Steven Jay Gould's concept of exaptation. In their proposal, Borsboom et al. are not crystal clear whether they place their model in any of the historical pathways of neuroscience. Their proposal seems disembodied and a metaphor for psychic singularity. As such, the cognitive way would seem the option they implicitly choose. However, in the absence of causality, their model would be rather separate and apart from any adaptation or evolution model.

The representation proposed by Borsboom et al. in a graph with subjective and objective factors placed on the same foot is only a reductive visual illustration of a complexity encountered daily in clinical practice. The printed representation of a subset of factors involved in psychic illnesses seems limited compared to mental associations of the clinician. Hence, metaphor and graphical representation might reassure those unacquainted with the clinical work. Indeed, beyond a seemingly superficial reflection that remains always subjective, a meticulous work of analysis and professional reflection takes place. In the same way that writing and speaking reduce and constrain concepts and perceptions, graph representation remains a drastic reduction both of the patient's suffering and of the clinical relationship. According to Edgar Morin, "We are still blind to the problem of complexity" (Morin 2005, p. 24). Complexity requires that we try to understand the relationships between the whole thing and the parts of the whole; however, knowledge of the parts is not enough to know the whole. Thus, for the principle of reduction, we substitute a principle which conceives the relation of mutual implication between the whole and parts. Generalized complexity would thus be a paradigm that would require the combination of a principle of distinction and a principle of conjunction. If, according to Morin, we have learned from our education to separate more than to connect, to know is both to separate and to connect. We must therefore make an effort to connect the parties to each other in all areas. Thus, in order to think complexity, we need a complex thought that connects more than it cuts out knowledge in fields exclusively centered on an object. According to Morin again, we must reject the paradigm of classical thought which was well formulated by Descartes and which is based on the disjunction between, for example, spirit and matter. A paradigm of complexity associating distinction and connection in mutual involvement should replace this separation (Juignet 2015). Thus, one of the epistemological consequences of complexity is that science is invited to become multi- and even transdisciplinary. In the context of Borsboom et al.'s proposal, the non-consideration of causality and the equalization of the different factors involved in mental pathology suggests a transdisciplinary view of complexity, focusing on the link between parties; however, it reduces the distinction between different areas and therefore erases their specificities. In other words, Borsboom et al. substitute complex interactions with simple linear correlations.

Medicine has relied on the search for causality using the anatomo-clinical approach, whether at a macro or micro level. This approach led to the identification of treatment to care for patients when it is not possible to cure them. However, in mental health, the approach was based on parallel movements of (1) hypotheses generation on mental functioning through metaphors (i.e., cybernetic, psychoanalytic, biological, and so on) that gave rise to various research protocols (either connectionist or cognitivist models, as described previously); and (2) redefinition of mental illness through pharmacological compounds efficacy, such as depression (with antidepressants) or anxiety (with anxiolytics). In the biological model of mental health, the use of dynamic causal models (Desseilles & Phillips 2016) makes it possible to represent brain functioning with directional graphs maintaining a causal dimension that seem crucial to the medical approach. As opposed to Borsboom et al.'s proposal, recent original models intuit that complexity of mental phenomenon might be emerging from biological models involving causal interactions (Friston 2010).

In view of all of the points discussed previously, we argue that causality with full complexity should be the approach of choice for unravelling mental health complexity.

Symptoms are not the solution but the problem: Why psychiatric research should focus on processes rather than symptoms

Immanuel G. Elbau, Elisabeth B. Binder, and Victor I. Spoormaker

Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, 80804 Munich, Germany.
Immanuel_elbau@psych.mpg.de binder@psych.mpg.de
spoormaker@psych.mpg.de http://www.psych.mpg.de/1448291/binder

doi:10.1017/S0140525X18001000, e7

Abstract

Progress in psychiatric research has been hindered by the use of artificial disease categories to map distinct biological substrates. Efforts to overcome this obstacle have led to the misconception that relevant psychiatric dimensions are not biologically reducible. Consequently, the return to phenomenology is once again advocated. We propose a *process*-centered paradigm of biological reduction compatible with non-reductive materialism.

Historically, biomedical research has oscillated between two principle scientific paradigms: mapping of phenomena to remote causes versus rejecting the intelligibility of such remote causes and pursuing a taxonomy on the level of the phenomenon (Zachar & Kendler 2017). Most medical fields have settled successfully on the former paradigm, but in psychiatry, the pendulum continues to swing. The rejection of a biological reduction of psychiatric phenomena and the retreat to discretional mapping of symptoms, as follows from Borsboom et al.'s argumentation in the target article, is in line with the currently observed swing-back of the pendulum to the latter paradigm. This motion is motivated by the valid negative critique of the limited therapeutic success that biological research in psychiatry has had and is conceptually grounded in non-reductive materialism.

Specifically, the authors argue: "Currently, there is no compelling evidence for the viability of reducing mental disorders to unique biological abnormalities" (target article, sect. 1, para. 4). They conclude that this reasons against distinct biological causes for mental disorders. Although we agree with the premise, both from an empirical and a theoretical perspective, we do not agree that the conclusion is justified. Both historical evidence from biological reduction in other fields of medicine (Kotchen 2011) and the principle of minimal assumptions (Ockham's razor) favor the alternate conclusion that it is not the assumption of biological causes, but the artificial and arbitrary labeling of mental disorders that is problematic. Artificial, ever-changing definitions are unlikely to lead to a biologically coherent cause – or biological correlate (Katahira & Yamashita 2017). We agree with Borsboom et al. that a research paradigm that tries to unravel the causes of such artificial entities is flawed by essentialism; however, shifting the focus from disorders (i.e., symptom-aggregates) to symptoms inherits the problem of essentialism: For what is a symptom? A symptom is merely a verbal subjective or behavioral consequence of a (dysfunction in a) given process/system. In analogy to other organ systems, the relationship between disturbance in such a process and reported symptoms may be nonlinear and heterogeneous/pleiotropic. In accordance with the authors, this relationship might set off a cascade of interactions that, together, affect the course of illness. In this sense only, the symptom-networks approach, as advocated by the authors, precludes biological reduction (as does the partial-reduction approach that the authors propose). Instead of abandoning biological reduction altogether, or continuing the hunt for biological substrates of artificial entities, why not follow the formula that has proven most successful in all other medical disciplines?

For example, the cardinal symptom in cardiology, chest pain, does not reflect any relevant biological substrate, but emerges from a variety of causes (from heart ischemia to intercostal neuralgia) (Lenfant 2010). What has paved the road to reduction in this discipline was a paradigm shift toward the concept of *processes* (physiology and pathophysiology) (Granger et al. 1998). Symptoms, formerly the sole means of establishing a diagnosis, now merely serve as a guiding torch, but the characterization of patients takes place directly at the level of these processes (e.g., ECG). An approach based on the understanding of how a quantitatively measurable physiological process (myocardial oxygen delivery and extraction) can be disturbed (sclerotic plaque) and how this unreliably maps to specific symptoms (from none to severe chest pain or abdominal symptoms), has only recently started to be developed in psychiatric research (Friston et al. 2017; Peters et al. 2017; Stephan et al. 2016).

Translated into the context of psychiatry, these processes will need to reflect relevant classes of computations that the brain performs to maintain the organism's homeostasis (Friston 2010). The definitions of such processes (e.g., reward prediction processing, salience monitoring, or executive functioning), as continuous functional entities cannot be derived directly from symptom properties. Instead, they need to be informed by comparative physiological and functional anatomical studies across species and through investigations in healthy individuals.

One model of such a process that has attracted recent attention is the process of *reward prediction*, representing a proposed fundamental neural computation that takes place at multiple hierarchical levels (from sensory input to higher-order associations) and serves to update the brain's model of the world in the face of new evidence (sensory input) (Keiflin & Janak 2015). Although these processes span a space in which subjective, reportable experience might correlate with individual variance along these axes (e.g., prediction error and feeling of surprise), the success of reduction will have to be evaluated by the utility of direct assessment and manipulation with regard to relevant outcomes.

In Figure 1, we illustrate exemplary dimensions of a putative process and how we envision its non-unique relationship to commonly used symptoms.

Whether a process-centered approach will yield a complete description of mental states (in terms of content) in material terms within an individual's brain is not relevant for psychiatric research or praxis. Relevant are the dimensions of mental states, such as anxiousness, vigilance, or relation to reality. Importantly, acute manipulation along these dimensions – for instance, pharmacologically – is possible (e.g., benzodiazepines, propofol, or LSD, respectively).

Conceptualizing psychiatric phenomena as the result of disturbances in functional brain processes allows us to overcome another limitation that Borsboom et al. suggest as intangible with biological reduction: accounting for social and normative

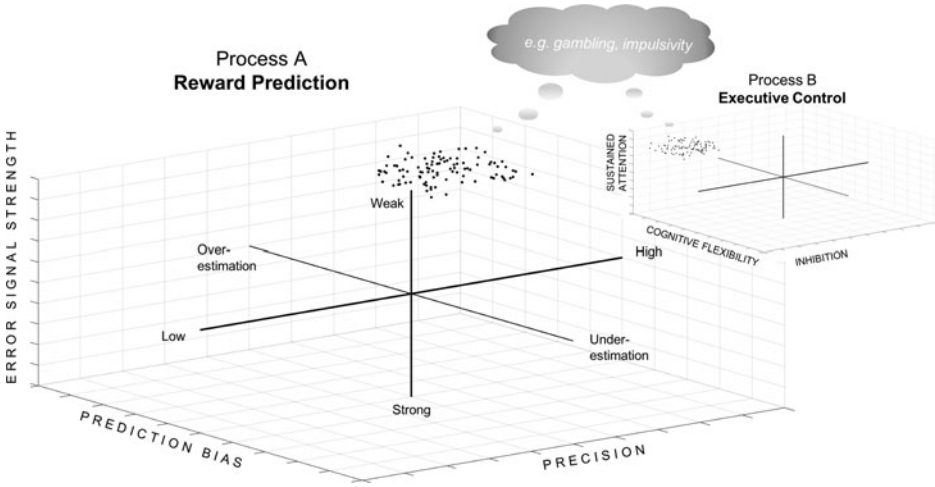

Figure 1 (Elbau et al.). The putative process of *reward prediction*, with three exemplary dimensions/axes (out of many). The *x* axis depicts the accuracy of a prediction (i.e., the tendency to underestimate or overestimate an outcome). The *y* axis reflects the precision of a prediction (i.e., ranging from a point prediction to broad interval predictions), and the *z* axis represents the strength of the error signal that drives learning (i.e., from weak to strong) An overestimation of positive outcomes combined with a high precision and a weak error signal (see black dots for exemplary individual scores) may lead to dysfunctional behavior such as gambling. Importantly, the resulting behavior and subjective interpretation (e.g., symptoms) can range across a broad spectrum and might not be unique to this process (see insert at right upper corner).

factors. The authors exemplify this with gambling, which is strongly contingent on environmental (presence of gambling booths) and normative factors. When we regard gambling not as the substrate of reduction but as a phenomenon that can emerge from dispositions in different processes (e.g., reward processing, executive functioning), such an account becomes tangible. One can now investigate how environmental and societal normative factors shape these processes through an interaction between genomic setpoints and outside influences (e.g., stressors) on cellular, circuit, and behavioral levels. Such an integrated research paradigm was put into practice recently in the National Institute of Mental Health Research Domain Criteria (Insel et al. 2010) and the ROAMER initiative (Schumann et al. 2014). In conclusion, the pendulum continues to swing, and there is not yet reason to abandon the effort of biological reduction that has been most fruitful in all other fields of medicine.

Networks, intentionality and multiple realizability: Not enough to block reductionism

Markus I. Eronen[a] and Laura F. Bringmann[b]

[a]Department of Theory and History of Psychology, University of Groningen, 9712 TS Groningen, The Netherlands; and [b]Department of Psychometrics and Statistics, University of Groningen, 9712 TS Groningen, The Netherlands.
m.i.eronen@rug.nl, l.f.bringmann@rug.nl http://www.markuseronen.com
https://www.rug.nl/staff/m.i.eronen/ https://www.rug.nl/staff/l.f.bringmann/

doi:10.1017/S0140525X18001012, e8

Abstract

Borsboom et al. propose that the network approach blocks reductionism in psychopathology. We argue that the two main arguments, intentionality and multiple realizability of mental disorders, are not sufficient to establish that mental disorders are not brain disorders, and that the specific role of networks in these arguments is unclear.

We are sympathetic to the idea that mental disorders are not just brain disorders, and the target article by Borsboom et al. does an excellent job in conveying this antireductionist message to the psychological community. However, in this commentary, we will show that the two main arguments provided (more specifically, intentionality and multiple realizability of symptoms) are not yet sufficient to block reductionism, and moreover, that defending antireductionism does not require taking a network perspective.

The core idea of the network approach is that mental disorders should be seen as networks of causally interacting symptoms. A new insight that the authors put forward in the target article is that symptoms also often have intentional content (i.e., they are about something) and are meaningfully connected to one another, and that these contents and connections are not visible at the biological level. This seems to make explanatory reductionism impossible or, at least, very unlikely.

However, this intentionality argument is unlikely to sway a sophisticated reductionist. She could accept the importance of intentional contents and their meaningful relationships, but nevertheless argue that the real causal work is done by brain states. For example, it is pragmatically useful to describe and predict human behavior in terms of beliefs and desires, but this is consistent with the idea that the real causes of behavior are biological or neural (this was, roughly speaking, Dennett's [1987] original view on intentional explanation). That is, even though the intentional contents of symptoms may have an important pragmatic or predictive role in studying mental disorders, they need not figure in the causal mechanisms of mental disorders. This can be illustrated with an analogy: The Ptolemaic system of astronomy, in which Earth is at the center of the universe and planets follow circular trajectories with epicycles, was a very useful predictive and descriptive tool for centuries, but as a representation of celestial mechanics it is radically false.

Thus, in order to block reductionism, it would have to be shown that intentional states are not only pragmatically important, but are also part of the causal mechanisms of mental disorders. One step toward this would be to show that intentional states as such, and not just the underlying brain states, can be treated as (interventionist) causes, and can have genuine causal relevance (along the lines of Eronen 2017; see also Note 9 in the target article). A second step would then be to show that

the particular intentional states that appear in psychopathological networks *actually* satisfy the conditions for causal relevance: For example, by showing that intervening on the intentional content of a symptom, while holding other factors fixed, would result in a change in another symptom. For this purpose, the models presently used in the network approach (e.g., vector autoregressive models or Markov random fields) are not yet sufficient, as they are not causal models, and the extent to which they give causal information is unclear (e.g., Bulteel et al. 2016). Moreover, even if a reductionist is forced to accept that symptoms with intentional contents are real causes, she could still maintain that these higher-level causes will in the end be reduced to neural or biological causes. Thus, more is needed to stop reductionism.

Indeed, Borsboom et al. anticipate this kind of response, and in order to counter it, argue that mental states (in this case, symptoms with intentional contents) are *multiply realizable*: A given mental state can be realized in different ways in different individuals, which seems to make it impossible to identify it with a single biological state. However, also this between-individuals multiple realizability is not yet enough to block explanatory reductionism. Even if "fear of heights" is realized by brain state *X* in John, but by brain state *Y* in Mary, it can still be *locally* reducible: "Fear of heights" is brain state *X* in John's case, and brain state *Y* in Mary's case (Kim 1992). That is, even though "fear of heights" may not be identical to a single brain state, in each specific context it could be locally reduced to a specific brain state (e.g., "fear of heights" in John is identical to a brain state *X*). This is sufficient for explanatory reductionism: For example, temperature is widely regarded to be a reducible property, although it is realized in a different way in a solid, gas, or plasma (Bickle 2016). More generally, Polger and Shapiro (2016) have recently put forward a book-length skeptical account on the relevance of multiple realizability, arguing that most putative cases of multiple realizability can be explained away by a closer look at the scientific details. Thus, multiple realizability is unlikely to provide a strong foundation for the irreducibility of symptoms or networks.

Finally, we would like to point out that taking the network perspective is not necessary for defending antireductionism in psychopathology. Debates on intentionality and multiple realizability of mental states have a long history in philosophy of mind, and the same applies to the other arguments put forward in the target article (e.g., the context-dependence and individual variation of mental states). Thus, these arguments predate the recent network approach, and are not tied to it. What seems to be different compared to the earlier philosophical literature is that now the focus is on symptoms and their meaningful connections, and not just on intentional states in general. However, it is not clear why *networks* would play an indispensable role in this kind of reasoning: Studying mental disorders by focusing on psychological symptoms and their meaningful connections does not require using network models or conceptualizing mental disorders as network structures (e.g., Bringmann & Eronen 2018; Miller 2010; Persons 1986). Thus, the specific and distinctive role of networks in the antireductionist arguments still needs to be clarified. Until this is done, there seems to be no need to adopt the recent network approach in order to argue that mental disorders are not brain disorders.

Acknowledgments. In the preparation of this commentary, we have benefited greatly from discussions with Hanna van Loo (University Medical Center Groningen) and Jan-Willem Romeijn (University of Groningen).

Indeed, not really a brain disorder: Implications for reductionist accounts of addiction

Matt Field,[a] Nick Heather,[b] and Reinout W. Wiers[c]

[a]Department of Psychology, University of Sheffield, Sheffield, S1 2LT, United Kingdom; [b]Department of Psychology, Northumbria University, Gosforth, Newcastle-upon-Tyne, NE3 1LU, United Kingdom; and [c]Department of Psychology, University of Amsterdam, PB 15916, 1001 NK Amsterdam, The Netherlands.

matt.field@sheffield.ac.uk nick.heather@unn.ac.uk r.wiers@uva.nl
https://www.sheffield.ac.uk/psychology/staff/academic/professor_matt_field
https://www.northumbria.ac.uk/about-us/our-staff/h/nick-heather/
http://www.uva.nl/profiel/w/i/r.w.h.j.wiers/r.w.h.j.wiers.html

doi:10.1017/S0140525X18001024, e9

Abstract

Borsboom et al.'s formulation provides an opportunity for a fundamental rethink about the "brain disease model" of addiction that dominates research, treatment, policy, and lay understanding of addiction. We also demonstrate how the American opioid crisis provides a contemporary example of how "brain disease" is not moderated by the environmental context but is instead crucially dependent upon it.

The dominant explanation of addiction (substance-use disorder) is that it is an acquired brain disease (Leshner 1997; Volkow et. al 2016). In recent years, many academic researchers, clinicians, and philosophers have objected to this characterization (e.g., Davies 2018; Heather 2018; Heather et al. 2018; Levy 2013; Lewis 2017; Satel & Lilienfeld 2014). We commend Borsboom and colleagues for outlining a convincing alternative to biological reductionism as an explanation for mental and behavioral disorders. In this commentary, we outline how their approach provides the foundation for a fundamental rethink about the role of the brain in addiction, one that is able to retain many of the important contributions of neurobiological research to our understanding of the disorder without the requirement to accept the "greedy reductionism" (Dennett 1995) inherent in the "brain disease model of addiction" (Volkow et al. 2016).

First, consideration of Borsboom et al.'s notions of *rational relations* and *intentionality* highlights the lack of explanatory power of the brain disease model of addiction. Current brain disease model of addiction accounts are able to characterise the molecular, structural, and functional adaptations in distinct brain regions that are correlated with distinct symptoms or "stages" of addiction, that is, multiple overlapping "brain diseases." For example, Volkow et al. (2016) distinguish three recurring stages of addiction, each of which has a distinct neural substrate: (1) binge and intoxication, characterised by rapid learning about the incentive-motivational properties of the drug and associated cues; (2) withdrawal and negative affect, characterised by hyposensitivity of the brain reward system and an exaggerated stress response; and (3) preoccupation and anticipation, characterised by impaired decision-making and inability to resist strong urges.

Borsboom et al.'s notions of *rational relations* and *intentionality* can be applied to make sense of the addict's behaviour and how it is related to, but not fundamentally determined by, the underlying neurobiological changes. For example, regarding *intentionality*, within the "withdrawal and negative affect" stage, one must invoke intentionality in order to understand why the addict uses the drug to manage negative mood (because the drug has provided short-term relief in the past), and why medications that can alleviate withdrawal symptoms, such as nicotine replacement therapy or methadone, can reduce tobacco smoking and opiate use, respectively (Mattick et al. 2009; Stead et al. 2012). Regarding *rational relations*, to give one example, the observed "impaired control" over substance use seen in the "preoccupation and anticipation" stage (stage 3) can be understood as a direct consequence of increased valuation of the drug coupled with reduced valuation of alternatives (to drug use) that characterise stages 1 and 2, respectively (Berkman et al. 2017; Heyman 1996). Thus, there is no requirement to interpret the observed structural and functional changes in prefrontal brain regions as indicative of "impaired ability to resist strong urges" (Volkow et al. 2016). Our point is that attempts to use neurobiological changes to explain behaviour can lead to very misleading explanations that are contradicted by behavioural data.

Second, consideration of intentionality can account for an important observation about the long-term course of addiction: Most addicts eventually recover from addiction, and most of those that recover do so without any treatment (Heyman 2013). If addiction is an acquired chronic brain disease, how can this be so? Demonstrations that addicts are less likely to recover if they believe that they suffer from a chronic disease (rather than, for example, an unhealthy habit that could be overcome; see Eiser & Van der Pligt 1986; Eiser et. al 1985; Miller et al. 1996) make sense when viewed through Borsboom et al.'s framework: Addicts can change their behaviour and give up drugs, but only if their attributions for their addiction permit them to do so.

Finally, we suggest that the current "opioid crisis" in the United States provides a pertinent demonstration that addiction can be *primarily* determined by the broader social, environmental, cultural, and historical context (cf. Hart 2013). The origins of this crisis coincided with the de-industrialization, economic decline, and urban decay in the "Rust Belt" and Appalachian regions (Quinones 2016). Together with alcohol poisonings, suicide, and chronic liver disease, increasing death rates from opioid overdose occurred among middle-aged, white, non-Hispanic men and women of low educational levels – the "deaths of despair" (Case & Deaton 2015). There is also a strong inverse correlation between levels of "social capital" in United States counties and age-adjusted drug overdose mortality (Zoorob & Salemi 2017). More generally, there is evidence that deaths and emergency department visits related to opioid use vary with macroeconomic conditions (Hollingsworth et al. 2017).

It could perhaps be argued that these variables exert their effects on rates of addiction merely by increasing the prevalence of drug use, so that more people are susceptible to the brain changes that then lead to the development of addiction. But in our view it is far more likely that the variables in question are significant elements in the kind of broad causal network that Borsboom et al. describe. For example, "people discover that opioids are an excellent short-term balm for existential maladies like self-loathing, emptiness, erosion of purpose, and isolation. Years of heavy use condition people to desire drugs at the first stab of distress" (Satel & Lilienfeld 2017). So, too, the easy availability of opioids, whether by prescription from local medical practitioners or through the skillful marketing of illicit suppliers (Quinones 2016), make attempts at behavioral change less likely, and relapse (if change is attempted) more likely to occur. The overarching point is that these broad contextual determinants should be regarded as part of the casual nexus of the disorder of addiction, not merely as "social factors" that might moderate the expression of an underlying brain disease.

Acknowledgments. We thank Keith Humphreys and Sally Satel for useful comments on an earlier draft of this commentary, although responsibility for the final version is entirely our own.

Conceptualizing neurodevelopmental disorders as networks: Promises and challenges

Kristien Hens,[a,b] Kris Evers,[c] and Johan Wagemans[d]

[a]Research Group Psychiatry, University of Leuven (KU Leuven), 3000 Leuven, Belgium; [b]Department of Philosophy, University of Antwerp, Stadscampus, S.D.409, 2000 Antwerp, Belgium; [c]Parenting and Special Education, University of Leuven (KU Leuven), 3000 Leuven, Belgium; and [d]Laboratory for Experimental Psychology, University of Leuven (KU Leuven), 3000 Leuven, Belgium.
kristien.hens@kuleuven.be kristien.hens@uantwerpen.be
kris.evers@kuleuven.be johan.wagemans@kuleuven.be
https://www.uantwerpen.be/en/staff/kristien-hens/
https://www.kuleuven.be/laures/ http://www.gestaltrevision.be/en/

doi:10.1017/S0140525X18001218, e10

Abstract

The target article by Borsboom et al. proposes network models as an alternative to reductionist approaches in the analysis of mental disorders, using mood disorders such as depression and anxiety as examples. We ask how this framework can be applied to neurodevelopmental disorders such as autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD). Specifically, we raise a number of promises and challenges when conceptualizing neurodevelopmental disorders as networks.

Neurodevelopmental disorders such as autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD) are considered mental disorders: Diagnostic manuals such as the *DSM-5* define them by behavioral characteristics and by the impact of those characteristics on daily-life functioning. Because of their high heritability, ASD and ADHD are also considered genetic and neurobiological conditions. As a result, the vast majority of research is trying to explain their genetic and neurocognitive etiology. However, using simple genetic models or a unified neurocognitive explanatory model, no one has been able to carve the neurodevelopmental disorders at their natural joints (Waterhouse et al. 2016). As such, they are exemplars of the thesis of Borsboom et al. that "findings [regarding the biology of mental disorders] have not been translated into convincing reductive explanations of mental disorders through central pathogenic pathways rooted in neurobiology, as many had expected" (target article, sect. 1, para. 3).

By looking at co-occurring symptoms and their relationships, instead of trying to find a common underlying cause, the network perspective also acknowledges that diagnostic boundaries are not always clear and that behavioral symptoms commonly co-occur. Hereby the approach gives place to a trans-diagnostic and more dimensional perspective, instead of the purely categorical approach. Moreover, applying a network approach to neurodevelopmental disorders not only challenges current reductionist approaches, but at the same time also offers a research strategy that does not deny the realness of the lived experience. For example, the network approach incorporates cultural and environmental contexts. Indeed, when we look at the conceptualization of autism through history, a shift can be noted of the behaviors that are considered "autistic," and this probably also differs in different cultural contexts (Kim 2012). Also, new approaches such as enactivism and discoveries regarding epigenetics suggest that a conception of neurodevelopmental disorders as challenges that can be attributed solely to a malfunctioning in the individual is wrong because they also depend on their cultural and environmental context (De Jaegher 2013). The network approach also stresses the importance of intentional information – information about mental states – as conveyed by those with the disorder. This leads to a better *Verstehen* of what it means to have the disorder. Especially in specific types of research into the origins of neurodevelopmental disorders, the latter has been completely ignored. One only needs to look at the many mouse models for ASD or ADHD for examples of how the actual meaning of experience has often been neglected in favor of reductionist and mechanistic explanations.

We also pinpoint a challenge that may hamper a successful implementation of a network approach to the neurodevelopmental disorders: The symptoms that are historically associated with these disorders, and that are used as diagnostic criteria, may not fully grasp what is to be considered ASD or ADHD. In a network approach, symptoms do not co-occur because they are symptoms of a common underlying disorder, but because they directly influence each other. As an illustration, for ASD, one could say that the preference for repetitive behavior influences social interaction, and vice versa. Or for ADHD, being inattentive and being hyperactive may influence each another without being caused by one underlying condition that *is* ADHD. Although similar problems exist for mood disorders such as depression, especially with regard to neurodevelopmental disorders, the network approach will have to deal with a certain circularity: "I am hyperactive; therefore, I get a diagnosis of ADHD"; "I have ADHD; therefore, I am hyperactive." This circularity suggests that, for many people diagnosed with a neurodevelopmental disorder, the symptoms used for diagnosis do not completely grasp or coincide with what they experience as being at the core of their disorder, and this mismatch or gap may be larger in neurodevelopmental disorders than in other mental disorders. Indeed, in recent years, we have seen an explosion of autobiographical accounts by persons with neurodevelopmental disorders. In these, it is often claimed that the way these disorders are defined, and the behavioral characteristics that define them, are not accurate descriptions of what it actually means to have such a disorder. Some individuals with a diagnosis of ASD may claim that they are not socially challenged but instead have a misaligned reaction time. In order to come up with a socially adequate response, they need more than the standard window of opportunity for an appropriate social response. In the context of adult women with ASD, it is often claimed that they do not exhibit the standard behavioral characteristics, and that their autism is more difficult to grasp in DSM terms (see Mandy & Lai [2017] for a special issue on female ASD). Hence, we believe that in the case of neurodevelopmental disorders, it will be of utmost importance to correctly define the different nodes of the network used to conceptualize them. This task will have to be done on the basis of phenomenological studies of what it *means* to have a neurodevelopmental disorder. Moreover, especially in the case of ASD, it will be difficult to relate different symptoms with one another without committing to an explanatory theory. Such theory may in itself be too reductionist or not reflect the experience of those diagnosed. For example, can we consider social challenges and challenges with regard to information processing as separate nodes in the network that influence each other? Or, do we regard social challenges as a result of challenges in information processing, or vice versa? Depending on the explanatory model that is taken as a basis, the network will look entirely different. In trying to approach neurodevelopmental disorders as networks of symptoms, perhaps such disorders will turn out to be too heterogeneous even to be captured in this way.

The value of clinical and translational neuroscience approaches to psychiatric illness

Juyoen Hur,[a] Rachael M. Tillman,[a] Andrew S. Fox,[d,e] and Alexander J. Shackman[a,b,c]

[a]Department of Psychology, University of Maryland, College Park, MD 20742; [b]Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742; [c]Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742; [d]Department of Psychology, University of California, Davis, CA 95616; and [e]California National Primate Research Center, University of California, Davis, CA 95616.

jhur1@umd.edu rmtillma@umd.edu dfox@ucdavis.edu shackman@umd.edu http://shackmanlab.org[1] http://foxlab.ucdavis.edu

doi:10.1017/S0140525X18001036, e11

Abstract

Borsboom et al. confuse biological approaches with extreme biological reductionism and common-cause models of psychopathology. In muddling these concepts, they mistakenly throw the baby out with the bathwater. Here, we highlight recent work underscoring the unique value of clinical and translational neuroscience approaches for understanding the nature and origins of psychopathology and for developing improved intervention strategies.

Borsboom et al. conflate biological approaches to psychopathology with extreme biological reductionism and common-cause models of psychiatric illness. In fusing these three distinct ideas, Borsboom et al. use evidence against extreme reductionism and common causes to dismiss clinical and translational neuroscience – effectively throwing the baby out with the bathwater. But like the paper-and-pencil approaches favored by Borsboom et al., biological approaches do not necessitate either extreme reductionism or singular causes. And

while mental illness is undeniably based in brains and genes (Geschwind & Flint 2015; Turkheimer 1998), we agree with Borsboom et al. that biological interventions are not the only or even the best way of tackling every mental illness (Kendler 2012b; Lilienfeld 2014; Miller 2010). We also agree that psychopathology reflects the interaction of multiple contexts and causes – from molecular pathways to culture – with their relative importance varying across individuals, development, sexes, and disorders (Birnbaum & Weinberger 2017; Kendler 2012b; Shackman & Fox 2018).

The network framework championed by Borsboom et al. describes patterns among symptoms, but it fails to provide a deeper explanation – biological, cognitive, or computational – for where those patterns come from. With respect to risk and etiology, it focuses on symptoms, environmental factors (e.g., stress), and the connection strengths (covariance) among them. Although this framework can provide important new insights, it cannot explain why some individuals and their biological relatives are predisposed to experience specific symptoms in maladaptive ways or how environmental factors interact with particular symptoms to produce psychopathology. In contrast, biological approaches are beginning to do just that. For example:

1. Anxiety patients and individuals at risk for developing anxiety disorders show increased reactivity (Fox & Shackman, in press; Fox et al. 2015; Shackman et al. 2016b) and aberrant functional connectivity in the extended amygdala (Birn et al. 2014).
2. Like the anxiety disorders, extended amygdala function is heritable (Fox et al. 2015; 2018), associated with specific molecular pathways (Fox et al. 2012; Roseboom et al. 2014), and amplified by stress (Shackman et al. 2016b).
3. Heightened amygdala reactivity confers risk for the development of future internalizing symptoms, particularly among those exposed to stress (Shackman et al. 2016b).
4. Amygdala reactivity is amplified by exposure to the same kinds of stressors and psychological pathogens that can precipitate acute psychopathology (Shackman et al. 2016a; Shackman et al. 2016b).
5. Anxiolytics transiently dampen amygdala reactivity (e.g., Del-Ben et al. 2012) and amygdala damage markedly reduces signs and symptoms of fear and anxiety in humans, monkeys, and rodents (Feinstein et al. 2011; Oler et al. 2016).
6. Stimulation of the extended amygdala elicits subjective feelings of fear and anxiety in humans (Inman et al., in press) and heightened defensive responses to threat in monkeys (Kalin et al. 2016).

These observations motivate the hypothesis that circuits centered on the extended amygdala causally contribute to the development of maladaptive anxiety (Shackman et al. 2016a). Such observations are hardly limited to the amygdala and anxiety. Other work highlights the importance of ventral striatal circuits to anhedonia (Bewernick et al. 2012; Greer et al. 2014; Nugent et al. 2014; Pizzagalli 2014; Schlaepfer et al. 2008; Stringaris et al. 2015).

In rejecting common-cause models, Borsboom et al. neglect evidence that uncorrelated and dissimilar disease phenotypes can reflect common substrates (Kotov et al. 2017; Zhu et al. 2014), a pattern not readily explained by symptom-network models. Individual differences in amygdala metabolism, for example, are associated with both neuroendocrine and behavioral signs of anxiety – two phenotypes that are only weakly correlated with one another (Shackman et al. 2013). Likewise, lesions and other perturbations of the amygdala produce coherent changes in a range of disease-relevant phenotypes – neuroendocrine activity, passive avoidance, vigilance, and anxious feelings – suggesting that the amygdala-centered circuits represent a common cause (but likely not the only one) for some (but certainly not all) key features of pathological anxiety (Feinstein et al. 2011; Fox & Shackman, in press; Oler et al. 2016).

Mental illness imposes a staggering burden on global public health, and there is an urgent need to develop better treatments (Global Burden of Disease Collaborators 2016; U.S. Burden of Disease Collaborators 2018). Symptom-network treatment approaches represent, at best, incremental improvements over current clinical practice. Many, perhaps even most clinicians already focus more on symptoms and their interconnections than on *DSM* diagnoses and their myriad specifiers (e.g., Waszczuk et al. 2017). In contrast to symptom-network approaches, recent biological research highlights the possibility of developing completely novel interventions, reducing heterogeneity in clinical trials, more efficiently matching patients to treatments ("stratified medicine"), and more accurately predicting clinical course (Drysdale et al. 2017; Koutsouleris et al. 2018; Woo et al. 2017). Ongoing genomics research represents one of the few feasible paths to identifying and prioritizing new molecular targets, a prerequisite for developing improved drugs (Evangelou et al. 2018; Gandal et al. 2016; Pankevich et al. 2014). In short, biological approaches afford opportunities for improving the lives of patients that go beyond those afforded by symptom-centric frameworks.

So where do we go from here? Borsboom et al. remind us that clinical and translational neuroscience has historically been oversold and under-delivered. (For a related perspective, see Gordon & Redish 2016.) Billions of dollars have failed to uncover new assays or cures (Shackman & Fox 2018). Although Borsboom et al. tell us that this reflects the futility of biological reductionism, a growing number of neuroscientists – including the architects of the National Institute of Mental Health Research Domain Criteria (RDoC) – have concluded that past underperformance reflects limitations of *DSM* diagnoses, rather than any intrinsic limitation of biological approaches (Gordon & Redish 2016; Kozak & Cuthbert 2016). Categorical diagnoses pose several critical barriers to discovering the nature and origins of psychopathology, including rampant co-morbidity, low symptom specificity, marked disorder heterogeneity, and poor reliability (Conway et al. 2018; Fried & Nesse 2015; Galatzer-Levy & Bryant 2013; Hasin et al. 2015; Kessler et al. 2005; Olbert et al. 2014; Regier et al. 2013; Watson & Stasik 2014). Addressing these problems requires that we focus on understanding the computational, cognitive, and biological bases of circumscribed symptoms or symptom clusters (e.g., anxiety, anhedonia). This "symptoms-not-syndromes" approach (Fried 2015) would also align more naturally with mechanistic work in animals (Fox & Shackman, in press).

In conclusion, there is a real intellectual danger to adopting Borsboom et al.'s framework wholesale. Although symptom-network approaches are valuable, they steer us away from deeper explanations for why some individuals and their biological relatives are prone to particular symptoms. A more holistic approach – one that embraces both biological and non-biological approaches (e.g., assessing relations between symptom networks and neural circuits) – is likely to yield greater dividends for understanding the nature and bases of psychopathology and accelerate the development of improved interventions for patient suffering.

Acknowledgments. The commentary authors acknowledge assistance from K. DeYoung and L. Friedman. This work was supported by the University of California, Davis; University of Maryland, College Park; and National Institutes of Health (Grant numbers DA040717 and MH107444).

Note

1. This web address, http://shackmanlab.org, applies to 3 authors: Juyoen Hur, Rachael Tillman, and Alexander Shackman. Andrew Fox's different web address is given separately.

Functional disorders can also be explained through a non-reductionist application of network theory

Michael E. Hyland

School of Psychology, University of Plymouth, Plymouth PL4 8AA, United Kingdom.
mhyland@plymouth.ac.uk
https://www.plymouth.ac.uk/staff/michael-hyland

doi:10.1017/S0140525X18001048, e12

Abstract

A network structure explains why reductionism is not possible for mental illness, but the same argument applies for the somatic symptoms of functional disorders. Because the covariation of symptoms of functional disorders cannot be explained in terms of symptom-to-symptom causality, explanation requires a network of biological mechanisms having emergent properties that cannot be reduced to biology.

Authors have argued for several years that psychology cannot be reduced to physiology because the emergent properties of the body cannot be deduced from biological theories (Bem & Jong 2013; Kirsch & Hyland 1987). The target article by Borsboom et al. provides a good rationale for why those emergent properties arise and therefore provides a convincing argument for why reductionism is not possible. However, networks apply not only to mental illness, and the same argument can apply to functional disorders.

The term functional disorder refers to patterns of somatic symptoms with no unique pathophysiology, also known as medically unexplained symptoms (MUS). Common examples of MUS include irritable bowel syndrome, fibromyalgia syndrome, and chronic fatigue syndrome. Less common examples include non-epileptic seizures and functional blindness. Functional disorders respond poorly to pharmacological interventions and, although evidence-based psychological interventions are recommended, outcome is often poor, with some patients rejecting a psychological interpretation.

There are two types of network. First, there are networks where symptoms cause each other. This is the type of network proposed by Borsboom et al., which I will call a symptom network. Second, there are networks where the biological mechanisms that cause the symptoms are themselves part of a network, and which I will call a mechanism network. Symptom networks can occur without mechanism networks, but mechanism networks lead to the expectation that causality between symptoms will also form part of the network structure. Thus, mechanism networks are consistent with symptom networks but not necessarily vice versa.

Mechanism networks can be treated in two ways. First, because the mechanisms are biological, it may be assumed that they can be treated as any other form of biological theory. The mechanism network is simply a more complex form of biological theory than that normally envisaged in medicine, where cascades of chemical events are described in terms of a causal sequence of events (Sturmberg et al. 2017). From this perspective, the application of the network concept is consistent with a reductionist perspective and requires a biological approach to diagnosis and treatment. However, this reductionist approach fails to accommodate the emergent properties of networks. Networks have emergent properties because the behaviour of the network as a whole depends on the strength and rate of change between the different connections among the nodes of the network – not on the properties of the individual nodes or mechanisms. One of the emergent properties of networks is rule-following behaviour, which has been demonstrated elsewhere (Wolfram 2002). Repetition of complex rules leads to complexity that cannot be predicted from an initial state. Therefore, the initial state of a biological network system cannot predict the outcome of that repetition when the network is of a kind that produces rule-following behaviour.

It is for the above reason that I and some others (Hyland 2011; Hyland 2017; Melidis et al. 2018; Martinez-Lavin et al. 2008) believe that functional disorders can be only partially understood in terms of biological – or psychological – theories and that a more helpful theoretical approach comes from the insights gained from artificial intelligence, machine learning, and complexity theory. That is, one should take an instrumentalist rather than a realist approach to the science of functional disorders, and develop theories that explain how particular patterns of events, both biological and psychological, produce the particular pattern of symptoms that are observed. The reason for assuming a mechanism network in contrast to only a symptom network is that covariation of some somatic symptoms of functional disorders cannot be explained in terms of symptom-to-symptom causality – for example, the covariation of diarrhoea and constipation of irritable bowel syndrome. However, there are many similarities between functional disorders and mental illness, so mechanism networks may also apply to both types of illness. Functional disorders and mental disorders are often co-morbid, and distinct diagnostic categories have been challenged for both types of illness (Wessely et al. 1999).

Mental illness and functional disorders share one important characteristic that has bearing on the issue of reductionism. For both types of illness, a unique pathophysiology has not, as yet, been identified, so diagnosis is based on symptoms. In the case of functional disorders, diagnosis is based on symptoms after exclusion of other biological causes of those symptoms. There are, of course, many biological differences between healthy individuals, on the one hand, and those with mental illness or functional disorders, on the other. For example, both exhibit a tendency (though not found in all cases) for raised pro-inflammatory cytokines, but there is no one-to-one relationship between a particular type of cytokine and a particular symptom. This failure to discover a specific pathophysiology is not for lack of trying, and continues to the present day. Prizes have been offered for a biological diagnostic test for chronic fatigue syndrome.

When something cannot be found but is known to exist, it may be that one is looking in the wrong place. Network theory suggests that the specific pathology of functional disorders – and there must be a different biological basis for each pattern of symptomatology – is not in any local mechanism, but in the strength of connections between the different mechanisms across the network. This possibility, for which there is preliminary support (Melidis et al. 2018), suggests that the specific pathophysiology is of a kind that will not be discovered using biology, and that it would be better to understand functional disorders – and mental illness – as a form of program error in a complex system. Although patients with mental health problems accept psychological interpretations of their illnesses, patients with functional disorders seldom do so (Stone et al. 2002), but they do find acceptable a narrative based on a program error (Hyland et al. 2016). Artificial intelligence may prove a better heuristic for the treatment of these troubling conditions than either biology or psychology.

Therapy and prevention for mental health: What if mental diseases are mostly not brain disorders?

John P. A. Ioannidis

Departments of Medicine, Health Research and Policy, and Biomedical Data Science, Stanford University School of Medicine; and Department of Statistics, Stanford University School of Humanities and Sciences; and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA 94305.
jioannid@stanford.edu
https://profiles.stanford.edu/john-ioannidis

doi:10.1017/S0140525X1800105X, e13

Abstract

Neurobiology-based interventions for mental diseases and searches for useful biomarkers of treatment response have largely failed. Clinical trials should assess interventions related to environmental and social stressors, with long-term follow-up; social rather than biological endpoints; personalized outcomes; and suitable cluster, adaptive, and n-of-1 designs. Labor, education, financial, and other social/political decisions should be evaluated for their impacts on mental disease.

In the target article Borsboom et al. argue convincingly that mental diseases are (mostly) not brain disorders, but represent highly complex network relations that depend on cultural, historical, and environmental mechanisms. They are highly variable across settings, across individuals, and even for the same individual in different settings, circumstances, and life periods. This narrative has major implications for the treatment and prevention of these conditions.

The new narrative explains why focusing on neurobiology (e.g., neurochemistry) to explain mechanisms and to develop effective treatments for mental conditions has achieved limited progress. The failure is so prominent that big pharma has largely abandoned new drug development in this field, despite its huge burden of disease and potential market (Chandler 2013). The largest meta-analyses to date show that for most mental health diseases, available drug treatments result in modest average treatment effects (d = 0.2–0.4) (Cipriani et al. 2018; Huhn et al. 2014; Leucht et al. 2017), with small, incremental benefits over placebo. True treatment effects may be even smaller, if we consider biases (Ioannidis 2008). Some scientists even argue that extremely widely used drugs such as antidepressants are entirely ineffective and cause more harm than good (Gotzsche 2013), although this is probably an extreme position. For example, the recent largest meta-analysis on antidepressants found that almost all antidepressants were better than placebo for moderate/severe major depression, but the summary effect size for efficacy on a continuous scale was d = 0.30. There was also novelty bias: In head-to-head comparison trials, antidepressants seemed to work better when they were first marketed but then seemingly lost in efficacy as they became older (Cipriani et al. 2018). This further erodes the credibility of the estimated treatment effects in placebo-controlled trials, since these trials are mostly performed early in the licensing process when expectations are heightened.

Responses to drug and psychological therapies also show large between-person variability. Few patients have excellent responses, a modest proportion achieves some response, and many have no response. There is enormous investment in basic neuroscience research and intensive searches for informative biomarkers of treatment response and toxicity. The yield is close to nil. Even optimists acknowledge that, currently, there is still no clinically useful way to predict which patients will respond best to widely used medications such as antidepressants (Thase 2014). If mental health problems are mostly not brain disorders, the dearth of useful neuroscience-derived biomarkers is only to be expected.

To overcome this dead end, we should shift emphasis away from the research paradigm that considers mental health problems to be mostly brain disorders and move towards exploring other, potentially more fruitful paths. First, this would mean reducing emphasis on identifying etiological brain pathways, and through them, biological markers and surrogate outcomes. If consistently strong and clinically useful biological markers/surrogates do not exist, perpetually searching for them would be in vain.

Second, the design of clinical trials in the field needs to be radically recast. Instead of running thousands of small trials of short-term duration and short-term response assessments, we could focus on larger simple trials with long-term follow-up (Ioannidis 2008). These trials should use a completely different core of non-biological, social outcomes likely to have relevance for most individuals. Such endpoints include suicides (completed/attempts), loss of job, marital and social relationships, social disability, personal finances (e.g., bankruptcy), and major quality of life and patient-related outcomes (Macefield 2014).

Third, while such major endpoints are likely to be quite important to everyone, there are many other outcomes that are highly personalized. These personalized outcomes have to be defined for and by each patient, capturing what matters most under the specific personal and social context. One may discuss and choose before any treatment (either in an experimental trial setting or in real life) what context-specific outcomes have highest value in each case. What matters most may vary a lot across patients and may even change over time for the same patient, as life priorities and values evolve. Admittedly, evidence is weak to date on whether routine use of patient-reported outcome measures for feedback during the course of treatment improves the outcomes of mental disease (Kendrick et al. 2016). Simply sharing some information between patients and

physicians may not suffice. Full personalized choice of the outcomes that matter may be needed, as has been described for other diseases, for example, substance use and chronic obstructive pulmonary disease (Alves et al. 2017; Braid et al. 2016).

Fourth, we could focus more on research for therapeutic and preventive interventions that have non-biological speculated mechanisms. In particular, we could prioritize understanding and ameliorating environmental and social stressors (Radua et al. 2018). Many of these may be context-specific, and they may vary across different cultures, times, epochs, and civilizations.

Fifth, some interventions that might be effective may need to be applied and delivered at a group or community level, or to the entire society, while others may need to be tailored to single patients. This means that we need to develop expanded research agendas both for community-level/society-level interventions and for personalized interventions. The most appropriate study designs would be different, depending on the goal of these trials. Cluster randomized trials may be most appropriate for group-level interventions. Personalized options may include adaptive trials (to account for strategies of sequential choices in long-term follow-up as response and/or treatment goals change) and n-of-1 trials.

Finally, much of the mental-health–related burden of disease may be induced or prevented by decisions in areas that have nothing to do with the brain, and go beyond the traditional remit of biomedical science. Our societies may need to consider more seriously the potential impact on mental health outcomes when making labor, education, financial and other social/political decisions at the workplace, state, country, and global levels. Mental health should be part of the conversation when different opinions exist on which decisions are preferable. Evidence on the impact of contesting actions could inform these decisions. Instead of thinking of mental disease as a narrow problem of brain tissue, brain cells, and brain molecules, we may need to think of it as an evolving, ever-changing challenge for society at large.

Network models can help focus research on the role of culture and context in psychopathology, but don't discount latent variable models

Nuwan Jayawickreme,[a] Andrew Rasmussen,[b] Alison Karasz,[c] Jay Verkuilen,[d] and Eranda Jayawickreme[e]

[a]Department of Psychology, Manhattan College, The Bronx, NY 10471; [b]Department of Psychology, Fordham University, The Bronx, NY 10458; [c]Department of Family and Social Medicine, Albert Einstein College of Medicine, The Bronx, NY 10461; [d]Department of Education Psychology, City University of New York Graduate Center, New York, NY 10016; and [e]Department of Psychology, Wake Forest University, Winston-Salem, NC 27109.
nuwan.jayawickreme@manhattan.edu arasmussen@fordham.edu
alison.karasz@einstein.yu.edu jverkuilen@gc.cuny.edu
jayawide@wfu.edu
http://manhattan.edu/campus-directory/nuwan.jayawickreme
http://www.fordham.edu/info/21660/faculty_and_staff/5435/andrew_rasmussen
http://www.einstein.yu.edu/faculty/309/alison-karasz/
http://www.gc.cuny.edu/Faculty/Core-Bios/Jay-Verkuilen
http://college.wfu.edu/sites/eranda-jayawickreme/

doi:10.1017/S0140525X18001061, e14

Abstract

Borsboom et al. correctly note that the use of latent variable models in cross-cultural research has resulted in a futile search for universal, biological causes of psychopathology; however, this is not an inevitable outcome of such models. While network analytic approaches require further development, network models have the potential to better elucidate the role of cultural and contextual variables related to psychopathology.

In the target article Borsboom et al. assert that network models have the potential to highlight the important role of cultural and contextual variables in psychopathology, and that they allow for the modeling of such variables so that their roles can be properly elucidated. As the authors correctly note, these variables have been understudied, in part because of past emphasis on universalist, biological explanations in mainstream psychopathology research. A growing body of research suggests that cultural and contextual variables within individuals' local social worlds play key roles in the development of psychopathology (see Hinton & Good 2016; Kirmayer & Ryder 2016).

As psychologists who conduct research on the impact of culture and context on psychopathology, we agree with Borsboom et al.'s critique that the use of latent variable models in cross-cultural research has supported a fruitless search for universal, biological origins of psychopathology (see Littlewood 2002). However, we caution that this may not be an inevitable outcome of the latent variable approach per se. There are several instructive examples of studies in cultural-clinical psychology that use latent variable models to explore how cultural constructs of distress covary (e.g., Rasmussen et al. 2011) and test their construct validity (e.g., Chhim 2012). Furthermore, as both network analytic approaches and latent variable models are based on covariance, both may lead researchers to the discovery of similar patterns. Studies using network analysis to examine patterns of daily stressors, traumas, and symptoms (e.g., De Schryver et al. 2015; Jayawickreme et al. 2017) have reported similar results to studies examining similar variables but using a latent variable approach (e.g., Jordans et al. 2012; Rasmussen et al. 2010). This similarity in findings makes even more sense if one conceptualizes latent variable models as a pragmatic way to specify distributions that summarize the associations among a set of variables (Skrondal & Rabe-Hesketh 2004), rather than purely as reflections of underlying variables.

Another area in which we urge caution concerns the sources of data that have been used in dynamic network analysis thus far. Most studies to date have used data from responses to pre-existing (and often well known) psychological measures. But the vast majority of psychological measures have been constructed based on the assumption that items should reflect an underlying latent variable. Thus, these measures, which are constructed using methods such as factor analysis or item response theory, consist of items that have been selected to correlate highly with each other. One of the touted strengths of network analysis is its ability to identify symptoms that are key, or central, to the disorder in question (Borsboom & Cramer 2013), yet it is unlikely that this will be accomplished if one is using a measure in which all of the items are developed by design to correlate with one another. Studies that use expanded symptom pools or other theoretically relevant variables will be more valuable.

Overall, we believe that network analytical approaches as currently used have yet to deliver on the promise of network models (Wright 2017). Further methodological work is needed to determine the degree to which network analysis is able to elucidate the role of cultural and contextual variables in the development of psychopathology beyond latent variable models. Furthermore, network researchers should avoid including individual items from measures constructed to identify latent variables (e.g., through the use of factor analysis) and should consider using the average scores or total score instead. It should be noted that a number of these limitations of current network analytical approaches have already been observed by at least one of the authors of the target article (i.e., Cramer in Fried & Cramer 2017).

That said, the history of psychology suggests that research methods and analytic strategies often shape research questions and even epistemologies. The authors' critique of the latent variable model, especially to the degree to which it led to a fruitless (at least so far) attempt to discover the "underlying" biological substrate of mental disorder, is thus refreshing. Network models may lead more researchers towards an appreciation of the impact that cultural and contextual variables have on the development of person-level psychopathology. Such models may have the potential to move cultural-clinical research beyond a conceptualization of culture as a static, monolithic variable and towards a dynamic model in which multiple cultural, contextual as well as biological variables interact with one another (Morris et al. 2015). Recognition of these important causal factors also has implications for the reorganization of our current psychiatric nosological system. It has long been argued by cultural-clinical researchers that nosological systems such as the DSM are cultural products (e.g., Gone & Kirmayer 2010; Ryder et al. 2011) that, if rigidly applied, conceal culturally specific expressions of psychopathology (Kleinman 1988). Network models have the potential to help us develop more cross-culturally valid diagnostic systems that are flexible enough to take into account cultural and contextual variables. Ryder et al. (2011) have proposed that, rather than be restricted by diagnostic classifications, cultural-clinical researchers should consider both "lumping" syndromes together (e.g., examine anxiety disorders broadly rather than focus specifically on panic disorder, generalized anxiety disorder) and focusing on cross-cultural variability in specific symptoms. Network models lend themselves to such a conceptual approach.

The network takeover reaches psychopathology

Richard J. McNally

Department of Psychology, Harvard University, Cambridge, MA 02138.
rjm@wjh.harvard.edu
https://www.mcnallylab.com

doi:10.1017/S0140525X18001073, e15

Abstract

Borsboom et al. have written a trenchant critique of biological reductionism in psychopathology. After commenting on recent controversies concerning the network perspective, I discuss ways of integrating biology into the network enterprise.

The controversies embroiling the latest version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; American Psychiatric Association 2013) culminated in the head of the National Institute of Mental Health (NIMH) proclaiming that the DSM would cease to be the essential framework for grant applications submitted to the Institute (Insel 2013). NIMH would now be "re-orienting its research away from DSM categories" to fund proposals targeting trans-diagnostic mechanisms as embodied in the Research Domain Criteria (RDoC) initiative (Insel et al. 2010). Remarkably, even the chief architect of the new manual had seemingly lost faith in the categorical approach. He and his co-authors announced, "We are now coming to the end of the neo-Kraepelinian era" (Regier et al. 2013, p. 68).

As the diagnostic manual was undergoing its periodic overhaul, a team of clinical psychometricians developed a hierarchical dimensional alternative to the DSM's categorical system (Kotov et al. 2017). Capping this structure was the *p factor* (Caspi et al. 2014), the general factor of psychopathology akin to *g*, the general factor of intelligence. And just as intelligence researchers have sought to identify the biological referent of *g* (e.g., Deary 2012), so have some psychopathologists suggested that the *p* factor "may have a physical reality" (Lahey et al. 2011, p. 187), perhaps genetic. However, it is questionable whether the *p* factor amounts to an empirical discovery about a major cause of psychopathology. As van Bork and her colleagues have emphasized (Van Bork et al. 2017), any dataset consisting of highly intercorrelated measures is mathematically bound to yield a general factor, even if the factor has no existential referent independent of the data themselves.

The categorical and dimensional perspectives have long been the only nosological games in town (McNally 2011, pp. 184–211). Yet, the landscape of our field dramatically changed when Borsboom, Cramer, and their associates introduced the network approach to psychopathology (Borsboom 2017; Borsboom & Cramer 2013; Cramer et al. 2010). On the heels of the first few empirical network studies (e.g., Cramer et al. 2012; McNally et al. 2015; Robinaugh et al. 2014), a tsunami of studies has appeared with many more in the pipeline (for reviews, see Fried et al. 2017; McNally 2016).

Pushback has been inevitable, especially from those favoring traditional approaches. As Bringmann and Eronen (2018) observe, many psychometric critics have argued that network and latent variable models are mathematically equivalent. Although that may often be the case, they are certainly not *ontologically* equivalent (cf. geocentric versus heliocentric models of the solar system; Galilei 1615/2012). Just as it makes a great deal of difference whether the earth revolves around the sun or vice versa, so does it matter whether symptoms arise from a latent variable having an existential referent (e.g., undetected brain tumor) or not. Treating a headache with aspirin is inadvisable in the former case, but not in the latter (Borsboom & Cramer 2013).

Analyzing two epidemiologic datasets, proponents of the latent dimensional approach launched a spirited critique of network analysis in psychopathology (Forbes et al. 2017a; 2017b). Their central claim was that "networks" do not replicate – or at least not as much as one would have thought. Unfortunately, their work was hobbled with errors, and when Borsboom et al. (2017) correctly re-did their analyses, the substance of Forbes et al.'s critique evaporated.

In their excellent target article, Borsboom et al. convincingly criticize contemporary biological reductionist accounts of mental disorders. The purpose of my commentary is to elaborate on several themes adumbrated by these authors.

Given the unsuitability of common-cause biological reductionist theories of mental disorder, what roles might biology play in the

psychopathology enterprise? One possibility would be to use biological measures as grouping variables in network analyses. For example, some studies have shown that people with at least one copy of the short allele of the serotonin transporter promoter polymorphism exhibit more symptoms of depression in stressful circumstances than do those with two copies of the long allele (e.g., Caspi et al. 2003). Van Borkulo et al.'s (2015) network comparison test applied to depression symptoms of subjects grouped by genotype (or other biological variables) could illuminate how such risk variables predict network topologies in psychopathology.

Another possibility would be to unpack the biology of signs and symptoms rather than syndromes. The classic psychiatric distinction between *form* and *content* may help us here (e.g., Taylor 1981, pp. 2–7). *That* a person hears voices in the absence of anyone speaking denotes the form of psychopathology, whereas *what* the voices are saying denotes its content. The intentional content – "aboutness" – of a symptom renders it elusive to biological reductionist explanations, as Borsboom et al. emphasize. On the other hand, the formal features of symptoms such as auditory hallucinations, experiences of influence, thought broadcasting, and so forth have better prospects for biological elucidation. And some symptoms appear to lack intentional content altogether. Consider posttraumatic stress disorder: Although flashbacks and nightmares possess intentional content, as they are about the trauma, emotional numbing and exaggerated startle do not.

Finally, the discipline of biology itself has been undergoing its own network makeover (Barabási & Oltvai 2004). Network genomics (Forst 2002), network neuroscience (Bassett & Sporns 2017), and network medicine (Barabási et al. 2011) are flourishing, nonreductionist enterprises. The goal is to build bridges between networks emerging at these levels of analysis rather than to reduce "higher" levels to "lower" ones. And now the "network takeover" (Barabási 2012, p. 14) has reached psychopathology, inspired by Borsboom and his colleagues.

Making a case for constructive reductionism

Christian P. Müller

Section of Addiction Medicine, Department of Psychiatry and Psychotherapy, Friedrich-Alexander-University of Erlangen-Nuremberg, 91054 Erlangen, Germany.
christian.mueller@uk-erlangen.de
http://www.psychiatrie.uk-erlangen.de/wir_ueber_uns/mitarbeiter/prof_dr_rer_nat_christian_p_mueller/index_ger.html

doi:10.1017/S0140525X18001085, e16

Abstract

Borsboom and colleagues argue that reductionism in psychopathology research has not provided the expected insights. Instead, they suggest a systems approach of interacting syndromes, which, however, falls short of a perspective for empirical testing. Here, a combination of both approaches is suggested: a reductionistic empirical approach allowing testability, synergistic with a constructivistic systems appraisal of syndrome networks – a *constructive reductionism*.

Borsboom et al. argue that reductionism in psychopathology research has not yielded sufficient insights to understand the complexity of the systems that cause psychopathologies, nor has it provided effective treatments. This somewhat pessimistic verdict may apply when the philosophically driven and publicly nourished expectations are a full understanding and full cure of the known psychiatric disorders (Müller 2018). The leading framework for the last decades of neurobiological research in psychopathology was reductionism. Leading hypotheses were generated that suggested a single psychiatric disorder can be reduced, that is, causally explained and treated, to a dysfunction in a single target or single functional system of the brain. It can be argued that this may have been supported by the technical limitations of new empirical techniques. For instance, the view that a dysfunction in the dopaminergic (DA) system of the brain is the sole causal mediator for drug addiction development was massively driven by the advent of *in vivo* microdialysis (Westerink 1995). This technique proved that all drugs with addiction potential acutely enhance DA activity in the brain's reward circuitry, while nonaddictive drugs do not (Di Chiara & Imperato 1988). Thereby, a single prominent and many-times-replicated finding was generalized to a functional theory (Koob 1992; McBride et al. 1999; Wise 2002) which did not pay tribute to the emerging complexity of the system (Salamone 1996). The therapeutic predictions based on this model, however, failed in practice (McCreary et al. 2015; Spanagel & Kiefer 2008), suggesting that the DA theory of addiction is at least incomplete (Nutt et al. 2015). One possible reason for that could have been the initial limitation of a key technique, in that only DA was measured and only dopaminergic innervated brain structures were considered. Empirical techniques advanced and showed that many more transmitter systems of the brain are dysregulated during and after drug consumption (Heilig & Koob 2007; Müller & Huston 2006; Schneider et al. 2017; Williams & Adinoff 2008). What is required in this field is not to deny the massive gain in knowledge based on single-target empirical approaches (Koob & Volkow 2016), but rather, a constructive synthesis of those findings (Spanagel 2009).

Borsboom and colleagues suggest a systems approach in which syndromes of distinct diagnostic categories interact. Only a systems understanding that also comprises proximal and distal environmental factors would allow for overcoming what they consider an epistemological failure. While this radical approach has the merit of considering syndrome-syndrome interactions (e.g., Schuckit et al. 1997), as well as environmental factors in psychopathology (Caspi & Moffitt 2006), it falls short of a major criterion for a scientific approach: its testability. Here it is argued that a major driving force for a bio-reductionistic approach in psychopathology is actually its testability. Hypotheses can be falsified on empirical grounds, when single variables – biological and environmental – can be addressed. This is difficult, and probably impossible, with syndromes that naturally involve a plethora of single causal elements which cannot be systematically manipulated at once. Empirical network approaches simply purport testability by bioinformatic large-data analysis. Although they are not worthless, they are just correlative by nature. By that way, they may generate hypotheses; however, these require explicit testing (Easton et al. 2013; Mielenz et al. 2018; Schumann et al. 2016). Borsboom et al. need to be asked for this crucial but elegantly avoided point in their theory: How do they want to test causal factors at the systems (network) level without reducing it to measurable units? So far, the examples they provide are just disguised reductionistic approaches.

In order to rescue the innovative momentum of the authors' proposal, a combination of both approaches may be considered: a

reductionistic empirical frame which allows empirical testability, interwoven with a constructivistic systems approach which brings single elements together to a higher-order network of syndromes: a *constructive reductionism*. Instances for that can already be found, for example, in drug addiction research. A systems approach proposed the interaction of disorders, like depression or schizophrenia causing the abuse and addiction to various psychoactive drugs (Müller & Schumann 2011). The reductionistic testing of mechanisms has recently yielded first testing of elements of a depression-alcoholism (Gulbins et al. 2013; Müller et al. 2017) and schizophrenia-nicotine addiction syndrome interaction (Koukouli et al. 2017). A constructivistic approach is now picking that up to a systems view that includes various syndrome-syndrome interactions (Müller & Kornhuber 2017; Schneider et al. 2017). Given the increasingly recognized complexity of the biological base of psychiatric disorder syndromes (Schumann et al. 2014), we should not expect a full-blown constructive synthesis happening in a single step, that is, by just interlinking single reductionistic findings.

Reductionism is, in fact, a multi-level reductionism, where functional networks are step-wise reduced to functional elements and sub-elements. Mirroring this, a multi-level constructivism should also be acknowledged. At the lowest level, reductionistic findings with the highest possible technical resolution are brought together to a nano-systems view (e.g., Kreek et al. 2005; Ungless et al. 2010). At the next constructive level, nano-systems views are brought together to constitute micro-systems views (e.g., Hyman et al. 2006; Robinson & Kolb 2004). This can be continued up to the level of the whole brain (Müller & Homberg 2015; Spanagel 2009) and syndrome behaviours (Schumann et al. 2014). This is not the end, however. An extended endophenotype of psychiatric disorders should comprise the whole body periphery as a proximal environment, and the social as well as physical environments of the organisms as distal environments (Badiani 2013; Zinberg 1984). The constructivistic synthesis may then end at a level where also the time component (i.e., personal, social and historical developments) is incorporated in this systematic construction. However, there will always be an error term in the reductionistic as well as constructivistic part of the approach. Simple chance events in biological systems, for example, *de novo* mutations (Michaelson et al. 2012), may act as single-case determinants. This cannot be overcome and controlled by either reductionism or constructivism, nor by their combination, and we may have to live with a certain degree of uncertainty.

Acknowledgments. The work of the author was supported by the Interdisciplinary Center for Clinical Research Erlangen, Project E22, and by research grant MU2789/8-1 from the Deutsche Forschungsgemeinschaft.

Neither biological nor symptomatology reductionism: A call for integration in psychopathology research

Benjamin C. Nephew,[a,b] Marcelo Febo,[c] and Hudson P Santos, Jr.[d]

[a]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01605; [b]Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA 01655; [c]Department of Psychiatry, University of Florida College of Medicine, Gainesville, FL 32610; and [d]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.
bcnephew@aol.com febo@ufl.edu hsantos@unc.edu
https://www.researchgate.net/profile/Benjamin_Nephe
https://www.researchgate.net/profile/Marcelo_Febo
https://www.researchgate.net/profile/Hudson_Santos_Jr

doi:10.1017/S0140525X18001279, e17

Abstract

We agree with Borsboom et al. in challenging neurobiological reductionism, and underscore some specific strengths of a network approach. However, they do not acknowledge that a similar problem is present in current psychosocial frameworks. We discuss this challenge as well as describe valuable parallels between symptom and neurobiological network theories that will substantially augment psychopathological research when integrated.

Borsboom et al.'s target article challenges the reductionism of psychopathology as a brain disorder. The authors argue against reductionism in neurobiological psychopathology research and propose that the focus should be on symptom-to-symptom interaction. We agree with Borsboom et al. in questioning the value of current trends in neurobiological reductionism, however, a similar problem exists in current psychosocial frameworks. We explore this challenge and underscore the specific strengths of an integrated network approach which includes both biological and psychological variables that will enhance psychopathological research.

Borsboom's network theory of mental disorders shifts our attention to symptoms and their relationship, positing that they can influence each other, creating a cascade of causal relationships (Borsboom 2017) with no need for a common cause – whether neurobiological or psychosocial. We commend the symptom network approach because it proposes that symptoms might have unique relationships and thus differential functional effects. This framework allows for the exploration of multilevel mechanistic approaches at both the micro (e.g., emotion, molecular) and macro levels (e.g., symptoms, coping) that involve biological, psychological, and sociocultural processes and perspectives. This assessment is in agreement with emerging literature indicating that underlying biology, risk factors, functional impairment, and life events are differentially related to specific symptoms (Fried & Nesse 2015; Olbert et al. 2014; Santos et al. 2017).

The debate around reductionism in the etiology of psychopathology is not new, nor is it limited to neurobiological research. Reductionism is also a hallmark of psychological theories of psychopathology. For example, Beck's Cognitive Theory of Depression is centered on underlying dysfunctional beliefs of the individual and does not consider potential biological changes. Neurobiological and psychosocial approaches to understand the etiology of psychopathologies have long had a competing history, and this has been seen as counterproductive (Kendler 2008). Both approaches to psychopathology are limited in many ways by their monocausal attempts at defining these disorders: (i) All psychopathology is best explained in terms of neurobiology; and (ii) all psychopathology is best explained in terms of specified mental and/or social mechanisms which cannot be deduced from biology (Kendler 2008). As described in Wittenborn et al. (2016), this monocausal approach is reflected

in a large number of studies, where 93% of articles from a literature search of major depressive disorder from 1980–2014 focused on only one key explanatory variable of depression (Wittenborn et al. 2016). What is often seen is an intense focus on implicating a popular research topic as the key causal psychopathology factor. This focus changes from year to year without cumulative integration into a more complex and inclusive causal framework.

The network symptomology approach and neurobiology are not as mutually exclusive as Borsboom et al. suggest. The complexity of symptom network models as discussed in the target article is remarkably similar to neural network imaging studies in humans and animal models, which are just beginning to comprehensively identify and characterize the intricate biological pathways affected in mental disorders. Aberrant neural network connectivity and activity in brain networks is associated with various psychopathologies, such as depression, anxiety, autism, schizophrenia, and addiction (Woodward & Cascio 2015). These brain imaging studies specifically target the complexity of mental disorders, in contrast to the portrayal in Figure 1 of Borsboom et al. On the contrary, neural network research is often more similar to the holistic approach of their Figure 2, and includes several quantifiable holistic measures, such as small worldness and correlation coefficient, that reflect the complexity of the etiology and symptomology of a multitude of psychopathologies (Drysdale et al. 2017). Clustering coefficient, a global measure of coordinated activity within individual brain regions, has been documented in schizophrenic patients and their non-psychotic relatives, suggesting that this measure can be used to identify endophenotypes and as a biomarker and risk factor (Lo et al. 2015). Changes in clustering coefficient are also disrupted in pediatric posttraumatic stress disorder (PTSD) (Suo et al. 2015) and may be a marker for depression-related brain network pathology (Gong & He 2015). Small worldness, an overall measure of neural network interactions, is often disrupted in schizophrenia (Kambeitz et al. 2016). It is also increased in individuals exposed to childhood maltreatment, a major psychological risk factor (Ohashi et al. 2017). While many of these imaging studies do incorporate symptom data, they often depend on a single overall diagnostic value or a limited subset of symptoms, in contrast to the network symptomology approach discussed by Borsboom et al.

Observations of the limitations of current reductionist neurobiology research should be used as opportunities to enhance future investigations and strengthen communication and collaboration between psychologists and neuroscientists, especially with regard to disease etiology and behavioral symptomology. While psychiatry and psychology are moving away from monocausal explanations, the same process has been happening in behavioral neuroscience, including increased focus on multilayered gene networks (Wei 2017), multi-hit gene x environment models (Samek et al. 2017), a greater consideration for complex transgenerational etiology (Nephew et al. 2017), and the holistic neural network analyses (Zhang et al. 2017) discussed in this commentary. We suggest that what is needed is more clinically informed translational neuroscience which fully embraces the vast biological, environmental, and symptomology complexity of most psychopathologies through the inclusion of integrated psychological and biological focused network approaches. Reasonable sacrifices of experimental reductionism will reap substantial gains in replication, robustness, and translation, and ultimately result in clinical progress.

Getting to the bottom of things: The value of evolutionary approaches in discerning the origin of psychopathology

O Jiaqing[a] and Martin Brüne[b]

[a]Department of Psychology, Aberystwyth University, Penglais Campus, Aberystwyth, Ceredigion, SY23 3UX, United Kingdom and [b]Department of Psychiatry, Psychotherapy and Preventive Medicine, LWL University Hospital Bochum, Division of Cognitive Neuropsychiatry and Psychiatric Preventive Medicine, Ruhr-University Bochum, D-44791 Bochum, Germany.
ojthehuman@gmail.com martin.bruene@ruhr-uni-bochum.de
http://www.aber.ac.uk/en/psychology/staff-profiles/listing/profile/jio2
http://www.ruhr-uni-bochum.de/philosophy/mibra/bruene.html

doi:10.1017/S0140525X18001097, e18

Abstract

The network approach as a novel way of understanding psychopathology has helped address some of the existing issues associated with traditional biological interpretations. Nonetheless, it has similarly failed in explaining the fundamental etiology of mental conditions – a persistent conundrum that arguably could be adequately addressed only by evolutionary formulations, specifically evolutionary mismatch and life history theories.

Borsboom and colleagues provide an informative account of the challenges inherent in the widespread adoption of traditional biology-oriented explanations of mental conditions. Nonetheless, while the network approach may offer a more integrative depiction of psychopathology, we contend that it is, in part, similarly based on an impoverished view of the term "biological," as is generally the case in the fields of medicine and psychology (Brüne 2016). Specifically, the prevailing undertaking of biological interpretations of mental conditions has largely focused on so-called proximate descriptions (e.g., biological mechanisms), while disregarding their "ultimate" or evolutionary dimensions (i.e., their phylogeny and adaptive value; Nesse 2013). When viewed through the lens of Tinbergen's framework encompassing both the proximate and the ultimate levels of explanations of specific design features or traits, a couple of key issues arise that render the network approach no more useful than the standard biological model in medicine and psychology (Brüne 2014; Nesse 2013).

The first claim is that the network approach offers a unique way of looking at psychopathology, whereby the synergism involving associated symptoms in a system of connections is purported to be key in the development of such conditions. Although it does make sense that symptoms of many psychological conditions are likely to affect one another in a progressive and reciprocal manner, the network approach fails to clarify as to *why* this is the case. Is it because some symptoms are naturally influential on others? If so, why is this the case? Borsboom et al. contend that the logical interpretation of intentional narratives behind certain symptoms could explain why. However, it begs the question as to why such relationships are rational or comprehensible.

The second claim is that, in the context of a network approach, psychopathological features are believed to be engendered either as a result of individual developmental trajectories or as an outcome of environmental circumstances (i.e., akin to the proximate level). It is similarly ambiguous as to why these proposed antecedents would effectively bring about psychopathological symptoms. As an example, why exactly might exposure to thin models on media induce self-esteem issues and extreme eating habits in females? If it is because of social comparison, then why do humans compare so much? It is quite apparent that such explanations seem to pose more questions than answers.

We suggest that an evolutionary framework could plug these gaps by providing a comprehensive, ultimate account to complement such an otherwise detailed theoretical perspective relating to the proximate viewpoint. Specifically, two main evolutionary concepts, one relating to evolutionary mismatch (Durisko et al. 2016), and the other to life history theory (LHT) (Stearns 2000; cf. Del Giudice 2014), respectively, are proposed to be highly valuable in shedding light on the actual origins of psychopathology.

For instance, the notion of evolutionary mismatch has been touted as a major underlying cause of a variety of mental conditions (Durisko et al. 2016). It refers to the manifestation of problems among people because relevant genes and cognitive/biological mechanisms (broadly useful in the ancient context) could not evolve fast enough to match the dramatic metamorphoses in existential conditions that have materialized in the contemporary world (Grunspan et al. 2018; O 2018a). To illustrate, consider the proposition that depression is an outcome of residing in an evolutionarily novel setting whereby one's personal, occupational, and social experiences are drastically different from that of prehistoric individuals (Hidaka 2012). Such a contextual framework would have provided important elucidations for the existence of depressive symptoms (e.g., having a sense of hopelessness and anhedonia or entrapment in a seemingly hopeless situation), which the network approach might merely attribute to the occurrence of a precipitating event (e.g., failing an exam). Although it is conceivable that an exam failure could indeed play a role in the emergence of depression, such a proximate explanation precludes a sufficient understanding of why one is susceptible to feeling hopeless (and subsequently develop depression) following such an event. Apart from banking mainly on rationality and the general acknowledgement of extrinsic influences as is the case with the network model, Hidaka's (2012) theorization (e.g., the existence of evolutionarily novel inadequacy of social support in the current context) could uniquely explain why contemporary humans may be more vulnerable to develop the condition. Similar formulations relating to evolutionary mismatch could likewise describe the pathogenesis of a wide assortment of many other disorders, ranging from animal phobias (O 2018b) to schizophrenia (Abed & Abbas 2011).

Likewise, LHT is believed to be comparatively useful in deciphering the etiology of psychiatric conditions (Del Giudice 2014), because the concept is not only valuable in comparing species but also in explaining within-species variations (Stearns 2000). It is a concept derived from behavioral ecology, which focused on the adaptive tradeoffs (e.g., a faster LH strategy involving spreading meagre resources across many offspring who will experience earlier sexual maturity and a higher mortality rate, versus a slower LH strategy encompassing heavy investment in a few offspring with sexual maturity/death occurring at a later age) in relation to the nature of external circumstances (e.g., residing in a highly dangerous vs. predictably safer environment) as a means to realize favourable procreative outcomes (Stearns 2000).

According to this approach, psychopathological problems are manifestations of either a slow (e.g., autism) or a fast LH strategy (e.g., attention deficit hyperactivity disorder [ADHD]) (Del Giudice 2014). While a network model would argue that a child develops ADHD because of parental neglect, for instance, LHT can explain why the condition may emerge in the context of gene-environment correlation, which is fully compatible with the mismatch approach. For example, some psychological traits nowadays associated with ADHD were once adaptive in harsh and uncertain environments in relation to survival and reproduction over much of human evolutionary history. These traits are less profitable in the evolutionarily novel world and may become "symptoms," depending on the quality of parental input (Bakermans-Kranenburg & Van Ijzendoorn 2006).

Taken together, we contend that a combination of the LHT and the evolutionary mismatch approach, to name just two major evolutionary concepts, would provide a fundamental framework to complement the network model in understanding psychopathology.

Evolutionary-developmental modeling of neurodiversity and psychopathology

D. Kimbrough Oller

School of Communication Sciences and Disorders, the University of Memphis, Community Health Building, Memphis, TN 38152.
koller@memphis.edu
https://umwa.memphis.edu/fcv/viewprofile.php?uuid=koller

doi:10.1017/S0140525X18001103, e19

Abstract

Modeling the extremes of mental/emotional conditions requires explicit accounts of evolutionary-developmental sources of human neurodiversity, not merely psychopathology. The target article's approach could be improved by incorporation of a hierarchical scheme wherein mental/emotional infrastructure interacts across differentiated layers of function. The notion of "symptom networks" thus calls for differentiation into hierarchically interacting components of mental/emotional evolution and development.

Borsboom et al. argue that the contents of thoughts interact to yield psychiatric disorders, proposing that psychiatric disorders are best seen as interactions among "symptom networks." The inherently abstract nature of these mental/linguistic interactions and the "multiply realized" nature of mental states reveals that the mental states are, according to the authors, detached from biological foundations. Similarly, it is claimed that symptoms themselves cannot be reduced to being direct consequences of a brain disorder. Thus, thoughts and sentences associated with a disorder are seen to operate at a level of their own, one that cannot be reduced to description as, for example, the neural impulses of a brain state. My view is different. I think recognition of the abstract nature of mental/linguistic events does not diminish the

importance of the neural impulses that, at another level, form the infrastructure *for* those events.

A more complete approach to modeling requires recognition that life is irreducibly structured in many abstract hierarchical levels, as explicated a half-century ago by Michael Polanyi (1968). In this perspective, the levels of function must be seen as real, just as Borsboom et al. argue that schizophrenia and other conditions are indeed "real patterns" (Dennett 1991). The target article, however, does not reflect the irreducible natural hierarchical levels needed to characterize the mind's underpinnings. I agree that the content of paranoid thoughts (symptom 1) cannot be reduced to the neural impulses associated with them. Neural impulses corresponding to paranoid thoughts may be driven, at least in part, by a state such as anxiety (symptom 2), which is also not reducible to the neural impulses upon which anxiety is grounded. Further, anxiety may be driven by hormonal imbalances (symptom 3). The three symptoms are related hierarchically, but the explication can go deeper. Hormone imbalance is surely founded on a system of gene expression driving imbalanced hormone production (symptom 4). Gene expression similarly cannot be reduced to genetic makeup (symptom 5), but genetic makeup clearly limits gene expression. The implied hierarchy outlined here has five naturally ordered levels, but many more are surely involved. And yet each level can be driven by levels above it. Paranoid thoughts can feed anxiety, which can feed hormone imbalance, and so forth. The target article's reasoning, represented in Figure 2, includes no account of hierarchical relations among levels of function, but instead treats brain states, symptoms, and environmental influences at a single level.

The approach has the advantage of emphasizing multiple realizability, the idea that there is no one-to-one mapping between brain states and symptoms, and no single "common cause" for symptoms. The authors persuasively point out that much effort has been wasted in pursuit of oversimplified expectations regarding roots of mental disorders. Still, the suggested alternative is itself oversimplified, in my opinion. An important deficiency is that the terms *symptom* and *brain state* in the article are undifferentiated with regard to level of function. The terminology imposes an arbitrary boundary among levels, only lower levels of function being treated as biological. However, the more fundamental issue is that failure to directly acknowledge the hierarchy of levels squanders the opportunity to reflect the mind's hierarchical structure and the origin of the mind in evolution and development. A key feature of any model of mental disorders is its ability to predict, across development, the risk of mental disorder, and this implies a critical need to represent evolved and developmental foundations.

Natural selection and developmental processes are organized around functions of life – for example, the tendency to experience states such as fear. Such functions are abstract, but are nonetheless real because they constitute organizing principles, subject to selection pressures, corresponding to complex neural events. Furthermore, all of the levels of function considered here have deeply conserved features, just as it has been dramatically demonstrated that organization of life includes foundations in unicellular life that are reflected as conserved operations and characteristics in all multicellular forms (Carroll 2005; Grant 2015; Hills 2006). Psychiatric conditions are grounded in the biology of humans and that of ancient life forms.

By suggesting the primary focus of psychopathology should be symptom-symptom networks, the target article radically de-emphasizes origins. It presents an undifferentiated list of symptoms and their interactions, with no explanation for how symptoms arise, why they persist, or why they are heritable. Further, it offers no explanation for why psychotic symptoms arise predominantly in post-adolescence (Hare et al. 2010; Kessler et al. 2007), but reliably earlier in men than women (Skene et al. 2017). In addition, it tends to isolate individuals diagnosed with disorders from the neurodiversity of the human population of which they are a part, because it refers to behavior patterns as symptoms without indicating how one should treat normal behavior that resembles the symptoms. This isolation is undesirable, given the tendency for psychiatric disorders to occur at higher-than-chance levels in proband families and for the same families to include higher-than-chance occurrences of other psychiatric conditions (Clementz et al. 2016). Furthermore, unaffected members of these families typically show subclinical patterns. The target article's insistence on the term *symptoms* in its modeling creates an arbitrary boundary between behavior in affected and unaffected individuals.

In a dynamic systems evo-devo approach (Gottlieb 2002; Müller & Newman 2003; Newman & Müller 2000), abstract levels of biological function are primary objects of modeling, with focus on interactions across the hierarchy. This approach facilitates prediction of risk for disorders, by focusing on evo-devo pathways. My colleagues and I have focused on the origins of language and ways in which a hierarchical scheme is required in order to characterize evo-devo foundations both of language itself (Oller et al. 2016) and of divergences from typical development (Oller et al. 2010; Patten et al. 2014; Warlaumont et al. 2014).

In the attempt to combat oversimplified portrayals of psychiatric conditions, Borsboom et al. argue for isolating thoughts as sources of disorder from the biological foundations that yield thoughts. I agree that thoughts are not merely neural impulses. Still, it is preferable for modeling to represent thoughts explicitly in relation to the developmental and evolutionary foundations that make them possible.

Brain networks require a network-conscious psychopathological approach

Achille Pasqualotto[a,b]

[a]Faculty of Arts and Social Sciences, Sabanci University, Tuzla 34956, Istanbul, Turkey and [b]Division of Psychology, School of Social Sciences, Nanyang Technological University, 637332 Singapore.
achille@ntu.edu.sg

doi:10.1017/S0140525X18001115, e20

Abstract

In experimental psychology and neuroscience, technological advances and multisensory research have contributed to gradually dismiss a version of reductionism. Empirical results no longer support a brain model in which distinct "modules" perform discrete functions, but rather, a brain of partially overlapping networks. A similarly changed brain model is extending to psychopathology and clinical psychology, and partly accounts for the problems of reductionism.

In the last few decades, the fields of experimental psychology and neuroscience have gradually moved away from oversimplification of multisensory processing (Ghazanfar & Schroeder 2006; Stein 2012). In particular, we have witnessed the slow demise of the concept that external inputs are processed by distinct modules and only later associated (Fodor 1985; Kanwisher et al. 1997; Luria 2012), and the affirmation of the concept that external inputs are processed by partially overlapping networks (Felleman & Van Essen 1991; Lloyd 2000; Macaluso & Driver 2005; Pasqualotto 2016; Pasqualotto et al. 2016). Initially, brain functions were investigated using one sensory modality isolated from the others, which only later admitted more than one sensory input. The latter approach better suits what happens in the *real* brain, where sensory inputs conveyed by different sensory modalities are processed *in parallel* and influence one another (Pasqualotto & Proulx 2015; Pasqualotto et al. 2013; Stein 2012; Wan et al. 2015). Moreover, different brain processes start to influence one another during the early phases of processing (Convento et al. 2013; De Meo et al. 2015; Foxe & Schroeder 2005). For example, touch affects vision during the early phase of visual processing (Convento et al. 2013). Progressively, research has showed that the clusters of brain areas (or networks) that were thought to be unisensory are actually multisensory (Kauffman et al. 2000; Murray et al. 2005; Zangaladze et al. 1999). For example, parts of the visual network are active during tactile tasks (Zangaladze et al. 1999). These findings have triggered theories such as "neural reuse" (Anderson 2010) and "metamodal brain" (Pascual-Leone & Hamilton 2001).

Together with advances in empirical research, improved technology has provided better understanding of partially overlapping brain networks. For example, improvements in fiber-tracking tools have shown the physical interconnections of different brain networks (Beer et al. 2011; Le Bihan 2003; Uesaki et al. 2018). Likewise, the development of multi-voxel pattern analysis (MVPA) has allowed us to understand the time course of their interactions (Furlan et al. 2013; Norman et al. 2006; Woolgar et al. 2016).

Borsboom et al. present convincing evidence that this novel paradigm is infiltrating *clinical psychology* too. Just as oversimplified models in experimental psychology and neuroscience were not able to convincingly depict sensory processing, it seems that similar mistakes in clinical psychology did not allow the adequate depiction and treatment of the pathological brain (Fornito et al. 2015). Brain complexity itself, and not network analysis or reductionism, has made earlier approaches not able to study and/or treat the pathological brain – a complexity that was first revealed between experimental psychology and neuroscience. The initial approach could be likened to looking at a grand panorama through a narrow tube; it would be very difficult to understand and appreciate what we are looking at (the cost of reducing the field of view was actually investigated by Loomis et al. 1991).

Although the growing suspicion of reductionism in psychopathology may seem a recent phenomenon, the systemic approach is an early example of non-reductionist psychotherapies (Johnstone & Dallos 2013; Loriedo 2005; Tomm 1984). This is not limited to the patient affected by a psychopathology, but involves his or her family (or relevant persons) as well. Instead of exclusively focusing on the patient (reductionist approach), the systemic approach has the advantage of involving the closest people to the patient; thus, it attempts to treat the patient *and* to correct the environment where he or she lives. A more recent example of a non-reductionist approach to psychopathology is provided by the use of acting to counter the symptoms of dyslexia (Whitfield 2009; 2016; 2017). In particular, this approach is aimed to bypass the reading impairment using (non-verbal) visual imagery and action (acting) to remember a difficult-to-access text.

A recent development within the field of neuroscience may further speed the appreciation of brain complexity. In fact, apart from the gradual acceptance that the brain functions as a network or networks (Anderson 2010; Bressler & Menon 2010; Fair et al. 2009), neuroscience has recently started to consider the role of social interactions on brain functioning – social neuroscience (Gallese 2008; Immordino-Yang & Damasio 2007; Steinberg 2008). That is, starting from considering the brain as a set of individual modules, we moved to a model of the brain with partially overlapping networks and now consider the interaction among the brain and other individuals' brains. Social neuroscience has been developing along two main research lines: one investigating the effect of social emotions on brain functioning (Kawamichi et al. 2015; Kitada et al. 2010; Lederman et al. 2007), and one investigating the effect of cultural norms and attributes on brain functioning (Chiao 2009; Han et al. 2013; Kitayama & Park 2010). To close the circle, this development within neuroscience substantiates the use of socio-cultural aspects in psychopathology (Hacking 1999; Murphy 2005).

Reductionism – simplified and scientific

Leonid Perlovsky

Department of Psychology, Northeastern University, Boston, MA 02115.
lperl@rcn.com
www.leonid-perlovsky.com/

doi:10.1017/S0140525X18001127, e21

Abstract

In this commentary on Borsboom et al.'s target article, I address an inadequate, simplified use of the idea of "reductionism" in clinical psychology and psychiatry. This is important because reductionism is a fundamental methodology of science. Explaining mental states and processes in terms of biological and brain states and processes is fundamental for the science of psychology. I also briefly address a fundamental methodology of the goal of psychology as a hard science.

The explanation of mental states and processes in terms of biological and brain states and processes is fundamental for the science of psychology. This reductionism, however, can be complex; a specific recognized brain state may not correspond to a specific recognized mental state. As correctly pointed out by Borsboom et al. in the target article, for claiming "reductionism" it is insufficient to identify neural correlates of psychiatric conditions. Finding such correlates could be a first step in a long and complicated process of scientific research proving causal relations. However, the fact that a road to explanatory reductionism might be complex does not mean that it "is out of the question in network models" (target article, sect. 4, para. 3). This discussion is necessary because psychology does not yet reach the status

of a hard science, and the authors of the target article make excellent examples of this; still, it is not fully acknowledged. Nevertheless, the hard science of psychology must be the goal. Steps toward this goal are taken by physics of mind (Perlovsky 2006; 2010a; 2010b; 2016; 2017a; 2017b; Schoeller et al. 2018), a new science that turns psychology into a hard science. To the credit of physics of mind are the elucidating of psychological mechanisms that have not been understood for hundreds of years, including, for example, why musical emotions strongly affect us, a question that Darwin called the greatest mystery. I agree with the authors that demanding explanatory reductionism as a condition for grant awards might be premature; still, I argue that we must maintain a possibility and a goal of physics of mind – the hard science of psychology – and eventually, of clinical psychology and psychiatry.

Borsboom et al. ask: "What is the evidence for uniquely biological explanations of mental disorders?" (sect. 1, para. 3). They cite a 2013 *Nature* editorial according to which, "despite decades of work, the genetic, metabolic and cellular signatures of almost all mental syndromes remain largely a mystery" (Adam 2013, p. 417). Uncovering biological background did not lead to reductive explanations of mental disorders. Borsboom et al. then go on to say: "Given this absence of compelling evidence, it seems sensible to entertain the possibility that explanatory reductionism is wrong" (sect. 1, para. 4). I argue here against this premise of the target article.

The authors see an ideal "reductive explanation in the history of science" in "the reconstruction of the ideal gas laws in terms of statistical mechanics" (sect. 2, para. 1). Using an example from physics, I agree, is "the best," because physics is the paramount prototype of science, the first, and I would argue the only "hard science" is physics. Still, I would argue against such a straightforward comparison of physics and psychology or psychiatry. The reason is that the ideal gas and the human brain are incomparable in complexity. The ideal gas is one of the simplest natural systems, whereas the brain is among most complex.

For many years it has not been considered possible to develop a brain-mind theory with the same accuracy as physics describes physical systems. Nevertheless, recently this step has been undertaken. For developing the physics of mind, it is essential to establish the fundamental aspects that sets physics aside from other branches of science. This fundamental methodology specific to physics is unique for all areas of physics and is not used in any other area of science. The three specific aspects of the methodology of physics are: (1) identifying a few fundamental laws and their mathematical formulation, which serve as a basis for the considered area of physics; (2) mathematically formulating the theory of this area of physics based on these fundamental laws and making experimentally verifiable unexpected predictions; and (3) the experimental verifications of predictions.

Using this methodology of the physics of mind, it was possible to solve most of the complicated problems in psychology which had remained unsolved for decades and even millennia, such as: the interactions between cognition and language; the cognitive mechanisms of emotions of the beautiful; and the cognitive mechanisms of the emotions evoked by music (which Darwin considered a mystery, a most complex problem in human evolution). This proves that explanatory reductionism is capable of explaining the most complicated problems in psychology, even when network structures are involved (cognition, language, beauty, music). The proper methodology of physics of mind has been used in these examples; an essential step is identifying appropriate fundamental laws for each problem. This conclusion points us in the right direction for further research. But it is far from a ready-made solution for every problem.

Elimination, not reduction: Lessons from the Research Domain Criteria (RDoC) and multiple realisation

Tuomas K. Pernu[a,b]

[a]Department of Philosophy, King's College London, London, WC2R 2LS, United Kingdom and [b]Molecular and Integrative Biosciences Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, 00014 Helsinki, Finland.
tuomas.pernu@kcl.ac.uk http://www.tuomaspernu.london

doi:10.1017/S0140525X18001139, e22

Abstract

The thesis of multiple realisation that Borsboom et al. are relying on should not be taken for granted. In dissolving the apparent multiple realisation, the reductionist research strategies in psychopathology research (the Research Domain Criteria [RDoC] framework, in particular) are bound to lead to eliminativism rather than reductionism. Therefore, Borsboom et al. seem to be aiming at a wrong target.

Borsboom et al. aim to show that reductive research strategies are misguided in the context of psychopathology research. More specifically, they claim that adopting network models as an alternative framework for the analysis of mental disorders will show how the reductive aspirations of the traditional research are ill-founded. It seems, however, that they have misconstrued their target: The ultimate aim of neuroscientifically based research (or more generally, physiologically based research) on psychopathology is not to reduce mental disorders to neural phenomena, but to eliminate the current notions of mental disorder altogether by changing the nosological practices in fundamental ways.

The focus of the analysis that Borsboom et al. are offering is on a DSM-based nosology. The DSM defines mental disorders in a symptom-centred way. There is nothing inherently wrong with this: After all, it is from symptoms that all clinical work starts off, and the DSM can therefore be a useful diagnostic tool for clinicians. But Borsboom et al. are not interested in clinical practices in psychiatry, at least not primarily; rather, they are interested in psychopathology research – the question of what mental disorders really are, and how we should conduct research on defining their true nature. Symptoms, however, are mere signs, marks of the underlying disorder or illness; and even in clinical practice, the ultimate aim is not just to remove the symptoms, but to cure the physiological condition they stem from.

Therefore, those who are explicitly proposing reductionist research on psychopathology tend to stress that basing our scientific understanding of mental disorders on the DSM is ill-founded. In particular, the Research Domain Criteria (RDoC) project ultimately aims to replace the DSM-based classifications of mental disorders. The main thrust of this project is in the conviction that the DSM nosology is invalid: that it clusters together disorders that are

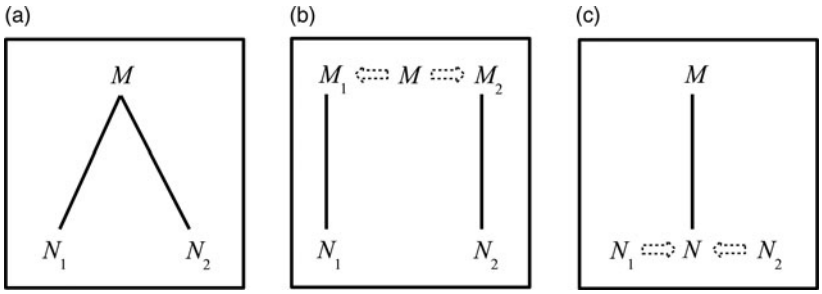

Figure 1 (Pernu). Figure 1A represents the multiple realisation hypothesis. Figure 1B represents kind splitting. Figure 1C represents unification or merging. (Source: Pernu, 2018)

symptomatically similar, but aetiologically and physiologically different (cf. Cuthbert 2014; Cuthbert & Kozak 2013; First 2012; Insel et al. 2010). Therefore, the idea is that we should aim to abandon the superficial DSM classifications and replace them with more valid classifications based on physiological aetiologies. Thus, this agenda is eliminativist, not simply reductionist, and Borsboom et al. seem to be aiming at a wrong target.

Philosophically, the tension between the DSM and the RDoC is easy to appreciate. The core of the issue is whether mental states – mental disorder types, in this case – are identical with their physiological realisers. In current philosophy, this translates into the question as to whether mental states are multiply realised (Fodor 1974; Putnam 1967). This issue is also at the heart of the analysis Borsboom et al. are offering. However, Borsboom et al. simply take for granted that the answer to this question is positive, thus ignoring the fact that there are a number of reasons to be skeptical of multiple realisation (e.g., Bechtel & Mundale 1999; Bickle 1998; 2003; Polger & Shapiro 2016; Shapiro 2000). As illustrated in Figure 1, the apparent cases of multiple realisation (Fig. 1A) have the tendency to become dissolved either by kind splitting (Fig. 1B) or by realiser merging (Fig. 1C). In the former case, the purportedly multiply realised mental state or function (M) splits into two (or more) separate entities (N_1 and N_2) as it is understood that the mental state or function does not constitute a single, unified psychological or neural entity (e.g., the way that "memory" or "attention" split to several different psychological and neural functions). In the latter case, the purportedly multiply realised mental state or function is identified with a single, unified neural state or function (N) as it is understood that the different realisers are actually physiologically the same (e.g., the way that intentions to grasp objects can be identified with the average neural activity of specific neural ensembles).

The RDoC framework can now be seen to aim at kind splitting: The apparently (symptomatically) homogeneous clusters of mental disorders (classified in terms of the DSM) can be predicted to split into new, homogeneous sub-clusters, each aligned with their physiological constitution (cf. Pernu 2019). Therefore, mental disorders would be multiply realised no more. The result is not reduction (as in Fig. 1C), but elimination: Our current understanding of mental disorders – "folk psychiatry" – will be fundamentally transformed, and the symptomatically defined notions of mental disorders (M) will give way to new notions, aligned with their neural-level realisers (M_1 and M_2).

In principle, there is a more forceful argument on offer for challenging the reductionists (or eliminativists). One could claim that different mental states – different mental disorders in particular – could be realised by the same physiological states. In other words, one could claim that it is not multiple realisation, but rather "multiple realisation in reverse" that the non-reductivists should be focusing on. And, in fact, in some places in the target article, Borsboom et al. point to this sort of an analysis. Note that this line of thought does not have to be in any way particularly controversial: Neural plasticity and reuse would already indicate that the same neural basis could give rise to different mental functions (cf. e.g., Anderson 2010).

If mental disorders would indeed be multiple realised in reverse, that would undermine the reductionist research strategies in a quite straightforward way: We would not be able to read off mental disorders from biomarkers, for any of them could ground different mental disorders. However, not only do Borsboom et al. fail to focus their analysis on this issue, one can immediately point to a fundamental philosophical problem: Multiple realisation in reverse would breach the core idea of nonreductive physicalism – namely, the idea that the physical basis is sufficient to exhaustively fix all of the higher levels of reality (mental disorders among them), that is, the idea of mind-body supervenience. Basing the analysis on the thesis of multiple realisation in reverse (rather than on the traditional thesis of multiple realisation) would therefore be bound to amount to a Pyrrhic victory.

Acknowledgments. I would like to thank Mr. Peter Cave, Dr. Nadine Elzein, and the participants of the Philosophy & Medicine Reading Group at King's College London for inspiring discussions on the topic. This work has been financially supported by the Emil Aaltonen Foundation.

Brain networks for emotion and cognition: Implications and tools for understanding mental disorders and pathophysiology

Luiz Pessoa

Department of Psychology and Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742.
pessoa@umd.edu http://www.lce.umd.edu/

doi:10.1017/S0140525X18001140, e23

Abstract

Understanding how structure maps to function in the brain in terms of large-scale networks is critical to elucidating the brain basis of mental phenomena and mental disorders. Given that this mapping is many-to-many, I argue that researchers need to shift to a multivariate brain and behavior characterization to fully unravel the contributions of brain processes to typical and atypical function.

How is function mapped onto the brain? Are functions implemented by single regions, and do single regions perform unique functions? The characterization and understanding of this structure-function mapping is essential for advancing our understanding about how complex mental functions and dysfunctions are related to the brain.

Brain regions participate in many functions, and many functions are carried out by multiple regions (Lindquist & Barrett 2012; Pessoa 2013). For instance, the dorsal-medial prefrontal cortex (PFC) is important for multiple cognitive operations, as well as for emotional processing (one-to-*many* mapping). Conversely, both frontal and parietal regions participate in attentional and executive processes, illustrating the situation of multiple regions carrying out a related function (*many*-to-one mapping). More generally, the mapping between structure and function is both *pluripotent* (one-to-many) and *degenerate* (many-to-one). Pluripotentiality means that the same structural configuration can perform multiple functions. Degeneracy refers to the ability of structurally different elements to perform the same function, yield the same output, or complete a task (Edelman & Gally 2001). To the extent that pluripotentiality and degeneracy hold, the combination of the two indicates that there are no "necessary and sufficient" brain regions.

An alternative approach conceptualizes mental functions in terms of brain networks. The *network itself is the unit*, not the region, and processes that support behavior are implemented via the interaction of multiple areas, which are dynamically recruited into multi-region assemblies. Does a network account solve the many-to-many problem outlined above? As argued elsewhere, the attempt to map structure to function in a one-to-one manner in terms of networks will be fraught with similar difficulties as the one based on brain regions – the problem is essentially passed to a different level (Pessoa 2014). Thus, two distinct networks may generate similar behavioral profiles (*many*-to-one), and a given network will participate in several behaviors (one-to-*many*). Broadly speaking, a network's operation will depend on several more global variables, namely an extended context that includes the state of several neurotransmitter systems, arousal, slow-wave potentials, and so on. In other words, a network that is solely defined as a "collection of regions" is insufficient to eliminate the many-to-many problem. What if we extend the concept of a network with these additional variables? Cacioppo and Tassinary (1990) propose that psychological events can be mapped onto physiological ones in a more regular manner by considering a spatiotemporal pattern of "physiological events." The notion of a network can then be extended to incorporate other physiological events, for instance, the state of a given neurotransmitter. How extensive does this state need to be? Clearly, the usefulness of this strategy in reducing the difficulties entailed by many-to-many mappings will depend on how broad the context must be.

In a manner that addresses Borsboom et al., we can ask: Are there specialized brain circuits for emotion? In an important sense the answer is "no," as the very boundary between emotion and the "rest of the brain" is ill defined. But how can a researcher interested in typical and atypical behaviors proceed, then? From the standpoint of studying specific tasks or conditions, distributed *activation fingerprints* (Anderson et al. 2013; Passingham et al. 2002; Pessoa 2014; 2017) provide summaries of evoked responses or states (see Figure 1A, left). Further insight can be obtained by studying multiple related tasks/conditions, and determining a multivariate fingerprint that highlights the relative commonality of activation across regions; for instance, that regions R_A and R_B tend to (but do not always) participate together across tasks/conditions (see Figure 1A, right).

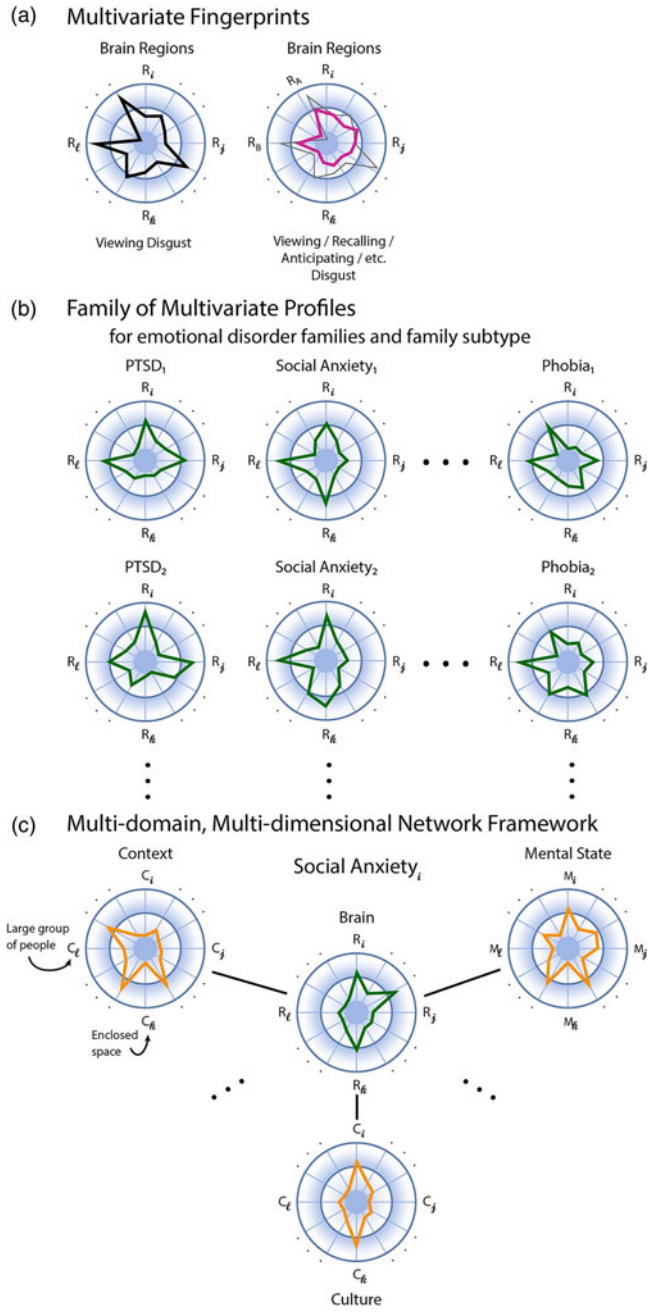

Figure 1 (Pessoa). Distributed characterization of brain-behavior mapping. **(A)** Left: The polar plot shows the distributed pattern of activation across regions R during an emotion task, such as viewing pictures eliciting disgust. **(A)** Right: Multi-region pattern of activation across tasks. The profile in pink represents activations that are (relatively) common across tasks (the gray outline is the same profile indicated at left). **(B)** Multivariate profiles can be applied to characterize patterns of brain activation that are associated with emotional disorders and their subtypes. **(C)** Brain profiles (in this case, for social anxiety) can be considered jointly with multivariate profiles that summarize contextual, mental state, social, cultural, and other factors.

Addressing the challenges raised by Borsboom et al., it is possible to advance our understanding of brain/mental disorders and psychopathology by building on the ideas presented previously. Consider anxiety, for example. Anxiety can be viewed as a *family* of related disorders. Let's suppose that the anxiety family

comprises a dozen relatively stable subtypes (related to PTSD, social anxiety, phobias, etc., but not only to these; the exact number is also not important). For each subtype, a multivariate fingerprint can be used to characterize patterns of brain activation that are associated with emotional disorders during a task or across conditions (Figure 1B). The fingerprints are then summary statements that should also include variability information around prototypical cases (see Figure 1A, right). The proposed framework allows disorder families to be heterogeneous themselves, so it is possible to conceptualize each subtype as a family, too – for example, the PTSD *subfamily*.

It is also possible to link mental disorder families and subtype families to the expanded scheme presented by Borsboom et al. For example, each of the individual fingerprints in Figure 1B can be linked to mental state and to contextual, social, cultural, and other factors, each of which also can be summarized by a fingerprint that summarizes a network of relationships within their own domain (Figure 1C). For example, the "context" factor can refer to scenarios such as being in a large group of people, being in an enclosed space, and so forth. This general framework now provides a concrete methodology to study mental disorders in a way that acknowledges the importance of brain processing, while including other domains that are significant for advancing the understanding of the multivariate and multi-domain aspects of the disorders. Note that the approach can be pursued concretely with existing network science techniques to study multi-relational networks, for example (Mucha et al. 2010). (Due to the need for brevity, I have not discussed issues of network dynamics and overlap, but they are also important; see Pessoa 2018.) Notably, in order to fruitfully apply the approach, it is not necessary to resolve the vexing problems related to the causal status of mental states, in particular (but see the "enactive" approach that informs these questions; e.g., Thompson 2007; Varela et al. 1991). Taken together, the framework summarized here has the potential to address the important issues raised by Borsboom et al. and to help elucidate how the brain contributes to mental function and dysfunction.

Acknowledgments. I am grateful to the National Institute of Mental Health for research support (R01 MH071589 and R01 MH112517). I also thank Christian Meyer for assistance with the figure.

Taking an engineer's view: Implications of network analysis for computational psychiatry

A. David Redish,[a] Rebecca Kazinka,[b] and Alexander B. Herman[c]

[a]Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455; [b]Department of Psychology, University of Minnesota, Minneapolis, MN 55455; and [c]Department of Psychiatry, University of Minnesota, Minneapolis, MN 55455
redish@umn.edu kazin003@umn.edu herma686@umn.edu

doi:10.1017/S0140525X18001152, e24

Abstract

An engineer's viewpoint on psychiatry asks: *What are the failure modes that underlie psychiatric dysfunction?* And: *How can we modify the system?* Psychiatry has made great strides in understanding and treating disorders using biology; however, failure modes and modification access points can also exist extrinsically in environmental interactions. The network analysis suggested by Borsboom et al. in the target article provides a new viewpoint that should be incorporated into current theoretical constructs, not placed in opposition to them.

Borsboom et al. challenge the notion that the brain should occupy a privileged position in mental health nosology and science, contending instead that symptom networks reflect the best units of analysis. However, psychiatry has made great strides in understanding and treating disorders using biology, and it is naïve to assume that because the model does not completely account for the full complexity, it is therefore useless.

Because all behavior arises from brain function, neurobiology is obviously critical for understanding psychiatric phenomena and not simply an example of "local reduction" of symptom networks. However, we think that the authors do have an important point that has not been incorporated well into current psychiatric reasoning: namely, that the trajectory of interactions with the external environment contains consequence chains that provide additional access points for treatment.

It is important to remember that psychiatric problems are not simply social constructs, but lead to real devastating consequences. For example, a patient with obsessive-compulsive disorder (OCD) unable to stop compulsive hand-washing is damaging their skin, leading to an increased risk of infection (Swedo et al. 1989) – certainly not something one would want anywhere, especially in a plague situation.

To help patients overcome their difficulties, we take an engineer's point of view, which asks two questions: (1) What are the failure modes that underlie psychiatric dysfunction? and (2) How can we modify the system?

The concept of a failure mode comes from reliability engineering – it recognizes that the structure of a process has specific ways in which it can fail. If we understood how environmental and neurobiological effects lead to behavior, then we could identify how this interaction can fail (MacDonald et al. 2016; Redish 2013; Redish et al. 2008). We agree that it is unrealistic to think that only the brain impacts disease – these failure modes may also be occurring outside of the individual's brain, that is, in the extended world.

Neurobiological effects on behavior have to be understood computationally, as information processing (Redish 2013). While these definitely have environmental interactions, the way to understand this is through quantitative theory. We believe that the network analysis that the authors propose here can provide an important contribution to quantifying the brain-environment interaction in computational models. Once you go to these theoretical models, the problem of reductionism falls away. The question is not whether an emergent phenomenon exists at a lower level, but rather how lower-level effects combine to produce the emergent phenomenon. This more-nuanced scientific discussion can be seen in the emergent phenomenon of traffic: Traffic on city streets is something that emerges only from structure. However, we can derive very accurate models of traffic patterns from understanding the underlying phenomena (physical structures of cars, timing of traffic lights, reaction times of typical drivers, etc.) (Seibold 2015). The fact that traffic is an emergent

network phenomenon does not mean that we can't reduce it to a complex interaction of other parts. It just means that we need to understand both the parts and their interaction.

The target article provides an important insight that has not been included in current theoretical conceptualizations: that some of these failure modes arise not from neurobiological failures, but from environmental failures that neurobiological systems are unable to cope with. Thus, for example, sleeplessness causes anxiety (Ohayon & Roth 2003), but anxiety can cause sleeplessness (Marcks et al. 2010), which could lead to a spiraling dysfunction. Intrusive memories in posttraumatic stress disorder (PTSD) are episodic, not semantic (Shay 1994; Talarico & Rubin 2003), and may arise from insufficient neural memory consolidation during sleep (McClelland et al. 1995; Payne & Nadel 2004; Rasch & Born 2013; Redish 2013). Even if normal consolidation processing were intact, a patient with PTSD unable to sleep might be unable to consolidate memory due to the lack of sleep and would remain subject to episodic flashbacks. If this scenario were true, it would mean that concentrating research looking for dysfunction in the neural system's underlying consolidation may be less fruitful (practically) than finding ways to facilitate an intact consolidation process.

Similarly, one could imagine scenarios in which treatment can use intact neural systems to bypass dysfunctional failure modes through modification of environmental components. For example, we have proposed that contingency management (a behavioral treatment in which addicts are rewarded monetarily for not taking drugs [Petry 2011]) works, in part, by shifting the decision process from a dysfunctional habit-based system into a non-dysfunctional deliberative system (Regier & Redish 2015). If this hypothesis is correct, then this treatment is bypassing, not curing, the brain dysfunction. Thinking about the interaction of brain and world provides alternative controls for treatment. Importantly, however, this proposal requires understanding the neuroscience underlying the information processing because it implies that we need to test for the intactness of the deliberative neural system before assigning someone to contingency management.

While extrinsic causal chains are discussed in the psychiatric literature, as in the widely taught "biopsychosocial model" (Frances 2014), they remain non-computational word-theories, which make them difficult to implement practically. Network analysis provides a mathematical toolbox to study these brain-behavior interactions, whether the causal chains arise from brain or environmental dysfunction. Importantly, network analysis is not atheoretical, and the network analysis requires hypothesized constructs. Network analysis applied to a taxonomy of DSM-based symptoms will produce one answer (American Psychiatric Association 2013; Borsboom et al. 2011), while a network analysis run over the Research Domain Criteria (RDoC; Flagel et al. 2016; Insel 2014; National Institute of Mental Health 2018) or decision-making constructs (Rangel et al. 2008; Redish 2013; Redish et al. 2008) will produce another. It will require collaborations with experimental and clinical psychiatrists to determine which of these taxonomies provide better explanation of the data and whether these tools can aid in practical psychiatric treatment.

Psychiatry has made tremendous progress over the last century, and computational psychiatry has opened up new ways of understanding the interactions of brain and behavior over the last decade (Huys et al. 2016; Redish & Gordon 2016). We think that network analysis has a part to play, but it needs to be incorporated into the current theoretical constructs, not placed in opposition to them.

Special, radical, failure of reduction in psychiatry

Don Ross[a,b,c]

[a]Department of Philosophy, University College Cork, Cork, T12 AW89, Ireland; [b]School of Economics, University of Cape Town, Private bag, Rondebosch 7701, South Africa; and [c]Center for Economic Analysis of Risk, J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30303.
don.ross931@gmail.com http://uct.academia.edu/DonRoss

doi:10.1017/S0140525X18001164, e25

Abstract

Use of network models to identify causal structure typically blocks reduction across the sciences. Entanglement of mental processes with environmental and intentional relationships, as Borsboom et al. argue, makes reduction of psychology to neuroscience particularly implausible. However, in psychiatry, a mental disorder can involve no brain disorder at all, even when the former crucially depends on aspects of brain structure. Gambling addiction constitutes an example.

The weight of theory and clinical evidence on mental disorders strongly supports the general perspective defended by Borsboom et al. in the target article. Indeed, increasing numbers of philosophers of science regard it as probable that most of the causal structure of the world, as identified across the sciences including fundamental physics, is best modeled in terms of networks. This is an important part of standard explanations for the fact that most scientific progress does not involve reduction (Horst 2007). Scientific progress generally proliferates types of processes rather than collapses them (Ladyman & Ross 2007), and the greatest challenge for most modelers is not *isolating* interrelated systems, but carefully *entangling* them in a way that still allows for rigorous estimation of parameters and causal effects in real empirical data.

The task that Borsboom et al. take on is explaining how and why network modeling undermines reduction in the specific case of psychiatry. Mental disorders, they argue, should seldom, if ever, be identified with brain disorders, though they recognize that the latter are often aspects of the former. However, they slightly *under*cook the radicalism of their challenge to reductionism as a *clinical* strategy. This is because their examples are limited to cases where the expression and perseverance of a complex brain state under a certain kind of mental characterization require reference to intentional patterns, which by their nature involve reference to environmental conditions and behavioral responses. But these are among the factors that block reduction of minds to brains *generally;* they fail to isolate particular obstacles to reduction that arise for *disorders*.

In each case that Borsboom et al. consider, one or more brain disorders are at least involved, even if not the "essence" of what makes the phenomena kinds of *mental pathologies*. To illustrate the full extent to which reductionism is misguided as a specifically psychiatric methodology, it is helpful to focus on a case in which the brain's contribution to a mental disorder results from its working just as natural selection "designed" it to do, but in an environment where this produces mental pathology. Gambling addiction is just such an instance.

Let us set the stage as conservatively as possible. Suppose that we apply the concept of "addiction" stringently, so that it refers, where gamblers are concerned, only to people who experience aversive somatic withdrawal symptoms when not gambling, to the extent that normal cognitive and emotional functioning in non-gambling settings is crowded out. This conceptual stringency might on the face of it seem to stack the odds in favour of a reductionistic account. It separates people who have a particular physiological dependency from the much larger population of problem gamblers who struggle with a range of social coping and self-management issues, and who lie along a smooth continuum with people for whom gambling is relatively low-cost recreation. The population of "true addicts" can be statistically distinguished from the wider group of "mere" problem gamblers (Kincaid et al. 2013), and such separation can be legitimately motivated by the suggestion that the former, but not the latter, are appropriately treated (in part) by neuropharmacological therapies (Ross et al. 2008). Is this not an exemplary instance of a reductionist's response to a critique such as Borsboom et al.'s, wherein those suffering from a brain disorder are systematically carved out from a class of behaviorally related cases where social forces mimic the "real disease"?

As we add known causal details, the reductionist path seems at first to pay dividends. Addictive forms of gambling involve sequences of statistically independent events that sometimes yield rewards and can be generated by simple, stereotyped actions on the part of the gambler. The addictive syndrome is made possible by a conjunction of three facts, which are indeed about neural mechanisms and architecture:

1. Allocation of cognitive-emotional attention and reward prediction are approximately fused in the same neural learning circuit based in the ventral striatum.
2. The circuit in question operates a specific form of Rescorla-Wagner learning algorithm that cannot settle on a model of genuine randomness, preventing it from learning that there is nothing to be learned from repeatedly playing things like slot machines.
3. Once the wider system in which the striatal learning mechanism is embedded discovers an action pattern that reliably delivers information relevant to trying to predict the reward sequence, it automatically cues motor preparation for harvesting activity around the focus of attention. Somatic cravings are the subjective experience of such motor preparation in the absence of opportunity for action. The cravings cue thoughts about gambling, and the thoughts sustain attention and motor preparation; so there is self-sustaining feedback.

This story seems transparently to be mainly about the brain. But for medical purposes – that is, as an account of a *disorder* – it is not reductive *at all*. This is because every component of the causal network that lies within the cranium is working just as evolution selected it to do. An easily accessible learning opportunity is presented, and the appropriate neural network tries to take advantage of it in the way that has garnered sustaining reward streams for active vertebrates for hundreds of millions of years. Alas, cunning manipulators have engineered an environment in which this "proper functioning" (Millikan 1984) leads to emotional, cognitive, behavioral, and social disaster for people. Since the introduction of sophisticated digital technology into slot machine design, the proportion of casino revenues derived from addicts has exploded (Schüll 2012). There is no reason to hypothesize that any endogenous vulnerability in brains has recently increased; the entire cause of this psychiatric pandemic lies in malicious environmental engineering, plus regulatory failure. So we have an alarmingly widespread mental disorder, crucially reliant on neural properties of a causal network – and *no brain disorder at all*. There is not merely a failure of reduction but, contra a concession Borsboom et al. allow, even of supervenience, *with respect to what is disordered* – the mind (and society), but not the brain.

Intentional content in psychopathologies requires an expanded interpretivism

Marc Slors,[a] Jolien C. Francken,[b,c] and Derek Strijbos[a,d]

[a]Faculty of Philosophy, Theology and Religious Studies, Radboud University Nijmegen, 6500 HD, Nijmegen, The Netherlands; [b]Department of Psychology, University of Amsterdam, 1018 WT, Amsterdam, The Netherlands; [c]Amsterdam Brain and Cognition, University of Amsterdam, 1001 NK, Amsterdam, The Netherlands; and [d]Center for Development Disorders, Dimence Mental Health Institute, 7416 SB, Deventer, The Netherlands.
m.slors@ftr.ru.nl j.c.francken@uva.nl d.strijbos@ftr.ru.nl
https://radboud.academia.edu/MarcSlors/ www.jolienfrancken.com
https://radboud.academia.edu/DerekStrijbos/

doi:10.1017/S0140525X18001176, e26

Abstract

We argue that the explanatory role of intentional content in connecting symptoms in a network approach to psychopathology hinges neither on causality nor on rationality. Instead, we argue that it hinges on a pluralistic body of practical and clinical know-how. Incorporating this practical approach to intentional state ascription in psychopathological cases expands and improves traditional interpretivism.

One of the novel features of the network approach to psychopathology defended in the target article by Borsboom et al. is their explicit recognition of an explanatory role played by intentional content. According to the authors,

> 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). (sect. 5, para. 3)

We are sympathetic to the proposed analysis, and we think that the explicit inclusion of intentional content strengthens earlier defences of the view. But what, exactly, does "sense making" and "understanding" mean here? The target article suggests answers to this question in terms of (1) causality and (2) rationality – intentional states are supposed to cause and/or rationalize behaviour associated with specific psychiatric symptoms. Both options are associated with an interpretivist view of intentional content that, as indicated in the target article, we share with the authors (Francken & Slors 2014; 2018). Nevertheless, we will

argue that both options are problematic and sketch an alternative view on what "sense-making" involves.

Interpretivism Dennett-style (Dennett 1987) – the version of interpretivism that is closest to what the authors have in mind – rejects the idea that intentional states are concrete items with straightforward causal powers (Slors 2007). From an interpretivist point of view, intentional states are abstract concepts that describe and identify specific *patterns* in human behaviour. Exploiting these patterns for explanatory purposes *does* involve exploiting the causal processes that underlie them (so interpretivism does not rule out "mental causation" altogether; see Eronen 2017). But interpretivists emphasize that we cannot identify these patterns – and hence exploit their causal underpinnings for explanatory purposes – other than by using the mentalistic idiom of "beliefs," "desires," "fear," "hopes," and so forth. Hence, the explanatory value of ascribing intentional states does not hinge solely on the causal relevance of intentional contents.

On the standard interpretivist picture, the explanatory value of ascribing intentional states hinges on the fact that these states "rationalize" behaviour: they provide *reasons* for actions, which are intelligible against the background of real or assumed rationality. The problem, however, is that this background seems to be lacking, at least in part, in the case of psychopathologies. Pathological intentional states such as delusions do not conform to the holistic constraints of our folk-psychology. Having the delusion that you are being spied on by the CIA will explain actions such as drawing the curtains. But it is not rationally connected with the agent's perceptions, other thoughts, motivations, and behaviour in the same way that a regular belief would be. Somewhere in the network of relations that delusions have with other intentional states, irrationality slips in. This is why many philosophers – certainly those of an interpretivist bent – would emphasize that delusions are not beliefs (Gerrans 2013; Radden 2011). They do not rationalise behaviour as beliefs do.

The explanatory role of intentional content in network analyses of psychopathologies hinges neither on causality nor on unqualified rationality, then. What does it hinge on? Part of the answer here is *bounded* rationality. The connection between the delusion that one is being spied upon by the CIA and the action of drawing the curtains can be seen as "rational," but only as long as the irrational inconsistencies within the larger context of the individual patient (or, by generalization, a particular population) are deliberately disregarded. To interpret pathological reasoning is to draw islands of reason in a sea of confusion, so to speak. Knowing how to demarcate these islands, however, requires familiarization with the mental status and history of the patient. This calls upon skills that do *not* take rationality as the main criterion for ascribing intentional content, such as empathic understanding (by means of simulation (Gordon 1996), phenomenological analysis (Ratcliffe 2015), and the targeted use of diagnostic tests and practice-based, theory-informed clinical scripts of the unfolding of psychopathological phenomena. Clinical understanding of the connection between symptoms, like everyday folk psychology, is essentially *pluralistic* (Andrews 2012).

We suggest that this whole body of social skills and practical knowledge – rather than rationality – provides the necessary heuristics for network analyses of psychopathologies at the intentional level of description. Within a particular network, targeted application of the criterion of rationality may help draw connections between specific symptoms, but only against this pluralistic background.

In short, we claim that the network view of psychopathologies defended in the target article is not supported by classic interpretivist views on intentional state ascription. We do not take this to be a criticism of the network view or to undermine its argument against reductionism in psychopathology research, however. Rather, we suggest that the fruitful use of intentional contents in forging connections between symptoms in a network analysis of psychopathologies, as argued for in the target article, points to a necessary expansion and improvement of interpretivism and its non-reductionist implications. Ascription of intentional states does not merely make behaviour intelligible and predictable by depicting the agent as being fully rational. Specifically in pathological cases, we use a range of further, non-rationalizing interpretive strategies, including simulation, empathizing, and using clinical knowledge and experience.

Why not be pluralists about explanatory reduction?

Kathryn Tabb

Department of Philosophy, Columbia University, New York, NY 10027.
kct2121@columbia.edu www.kathryntabb.com

doi:10.1017/S0140525X18002054, e27

Abstract

Borsboom et al. convincingly argue that, from their symptom network perspective, mental disorders cannot be reduced to brain disorders. While granting that network structures exist, I respond that there is no reason to think they are the only psychiatric phenomena worth explaining. From a pluralist perspective, what is required is not a full-scale rejection of explanatory reductionism but a critical attention to the circumstances of its application.

Borsboom et al. have done an important service in drawing our attention to a problematic ideology espoused by some of psychiatry's most influential voices. In one sense the slogan "mental disorders are brain disorders" is unprovocative, since there are few Cartesians around anymore to defend the existence of "pure" mental substance. But the slogan is usually understood to declare something more radical: that mental disorders are to be explained in purely physical (or biological) terms. There is a rhetorical sleight of hand here, from a claim about constitution to a claim about explanation. The twenty-first century penchant for physicalism over dualism is being manipulated to fuel public support for a set of research strategies that the authors group under the banner of "explanatory reductionism." Borsboom et al. aim to show that if, as they believe, mental disorders are best understood as causal networks of symptoms, then explanatory reductionism must be misguided.

In the philosophical literature, reductive explanations have traditionally been understood as explaining entities at a higher hierarchical level in terms of the interaction of constituent lower-level entities (Kaiser 2015; Kauffman 1971; Sarkar 1992). Such explanations can be partial – what Schaffner (2006; 2016) calls "patchy" – allowing reductionist methods to provide "creeping" but "not sweeping" accounts of the relevant phenomena. Borsboom et al. repeatedly acknowledge the value of such patchy reductions for

psychiatry (e.g., see sections 4 and 7 in the target article). Nonetheless, they claim that these successes do not vindicate the reductionist program, because "they have not been translated into convincing reductive explanations of mental disorders through central pathogenic pathways rooted in neurobiology, as many had expected" (sect. 1, para. 3).

This representation of the explanatory reductionist as insisting "that it is possible in principle to identify a common pathogenic pathway at the level of the brain" (sect. 2, para. 7) is narrower than the philosophical use of the term, and is, in 2018, a bit of a straw man. Certainly there was optimism throughout the twentieth century that the biological essences of psychiatric disorders would be uncovered, to transformative effect. But the consensus now is that conditions such as schizophrenia and depression are profoundly polygenic, and are not caused by single biological abnormalities (Kapur et al. 2012; Kendler 2013). Indeed, the leadership of the National Institute of Mental Health (NIMH), whom the authors rightly target as exemplifying explanatory reductionism, has emphasized that psychiatric disorders lack this sort of essential structure (Tabb 2015). Contrary to being motivated by the search for a common pathogenic pathway, as Borsboom et al. suggest (sect. 2), precision medicine's hunt for biomarkers aims to re-stratify the patient population in order to deal with the multiple realizability of diagnostic categories at the genetic and neuroscientific levels.

The relevant foil for patchy reductionism, then, is not a naive explanatory reductionism that assumes a simple causal story, but rather a *triumphalist* explanatory reductionism (hereafter, triumphalism) that maintains that the basic sciences will, in the end, provide all-purpose answers. When the authors argue that "mental disorders cannot be explained in terms of neural mechanisms" (sect. 1, para. 5), I take them to mean that triumphalism is misguided, because mental disorders cannot be *entirely* explained in terms of neural mechanisms. Reading the claim as a rejection of reductionism generally would put it at odds with the authors' other statements about the value of patchy neuroscientific reductions. In any event, it seems that this sort of triumphalist overreach, rather than a focus on common pathways, makes the "mental disorders are brain disorders" slogan so frustrating, and its effects so pernicious.

For advocates of symptom network models, the costs of painting all reductionists with the same dismissive brush are relatively low. After all, their models concern non-hierarchical causal relationships, so they need not engage with reductionism of *any* stripe. Borsboom et al.'s conditional framing (see their Note 4 in the target article) represents just this kind of reasoning: *If* the network model is correct, *then* explanatory reductionism can and should be rejected. But their argument also allows for another conditional: *If* there exist other sorts of psychiatric causes besides symptoms which *do* act as underlying mechanisms, which *do not* include intentional features, and which *are not* socially constructed, then these features will remain ripe for reductive treatment. In other words, if one thinks that network relations do not provide the full causal story, one only has reason to reject sweeping reductions on the grounds the authors give, but can still pursue creeping ones.

The authors conclude by advocating for holism, "the research strategy that is focused on the interaction between parts rather than on their individual realization" (sect. 8, para. 6). Given this, they might allow the claim that mental disorders are in some sense brain disorders *too*, or *in part*, insofar as there is room for patchy reductive explanations even once the triumphalist ideology has been rejected. There is, however, another way to see symptom networks as only part of the causal story; not in terms of holism, but in terms of *pluralism*. The pluralist would take the models represented in Borsboom et al.'s Figure 1 more seriously than the authors do, assuming a heterogeneity of mental disorders in which some cases might be best understood at lower levels (Figs. 1a and 1b), some at higher levels (Fig. 2), and some at multiple levels (Figs. 1c and 1d). In contrast with the holist's, this conclusion is more modest; the pluralist only can be sure that not *all* mental disorders are *exclusively* brain disorders. But this is enough to logically entail a rejection of triumphalism. And it leads to a generative question the holist won't pursue: Which psychiatric cases admit of reductionist approaches, and which don't? Borsboom et al. provide the beginnings of this sort of rubric when they provide three circumstances under which mental phenomena will resist reductive explanation – but there is much more work to be done.

Just as valuable as the authors' analysis of the canny sloganeering of the triumphalist is their own insurrectionist response. Their positive program might be captured by the alternative motto of "Symptomatology first!" As the authors importantly note, "to know which physical processes to investigate […], 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 in which we are ultimately interested" (sect. 7, para. 8). Basic sciences such as genetics and neuroscience will be of use for psychiatry only if they provide medical, rather than merely scientific, explanations. In this sense, Borsboom et al. are absolutely right to insist that we will always need psychopathology before reduction.

Problem behavior in autism spectrum disorders: A paradigmatic self-organized perspective of network structures

Lucio Tonello,[a,b] Luca Giacobbi,[a] Alberto Pettenon,[a] Alessandro Scuotto,[a] Massimo Cocchi,[c] Fabio Gabrielli,[b] and Glenda Cappello[a,b]

[a]"Ca'Leido" Autism Center, 61-31030 Altivole (TV), Italy; [b]Państwowa Wyższa Szkoła Techniczno-Ekonomiczna (PWSTE), 37-500 Jarosław, Poland; and [c]Scuola di Agraria e Medicina Veterinaria, University of Bologna, Ozzano dell'Emilia, 40064 Bologna (BO), Italy.
luciotonello@gmail.com luca.giacobbi@ca-leido.it
alberto.pettenon@ca-leido.it alessandro.scuotto@tin.it
massimo.cocchi@unibo.it fabio.gabrielli@pwste.edu.pl
glenda.cappello@pwste.edu.pl www.pwste.edu.pl www.ca-leido.it

doi:10.1017/S0140525X18001188, e28

Abstract

Autism spectrum disorder (ASD) subjects can present temporary behaviors of acute agitation and aggressiveness, named *problem behaviors*. They have been shown to be consistent with the self-organized criticality (SOC), a model wherein occasionally occurring "catastrophic events" are necessary in order to

maintain a self-organized "critical equilibrium." The SOC can represent the psychopathology network structures and additionally suggests that they can be considered as self-organized systems.

Autism spectrum disorders (ASD) represent a set of neurodevelopmental lifelong disorders characterized by deficits of social interactions, impairments in communication skills, as well as excessively repetitive and stereotyped patterns of behavior, interests, and activities (American Psychiatric Association 2013; World Health Organization 2010).

In particular, subjects with ASD can present temporary behaviors of acute agitation and aggressiveness. In fact, they can display momentary maladaptive and challenging behaviors as oppositional or disturbing activities as well as more severe behaviors such as self-injuries or aggressions. These multiple agitation forms are often called, for short, *problem behaviors* (or *crisis behaviors*) (Horner 2010; Stark et al. 2015).

The problem behavior has been defined as the "tip of the iceberg" because it seems to be the result of many different underlying variables interplaying in a complex way (Mesibov et al. 2004). Interestingly, in a very recent experimental work (Tonello et al. 2018), ASD problem-behavior dynamics have been shown to fit the self-organized criticality (SOC) model.

The SOC is commonly used to describe natural phenomena and systems such as earthquakes, riverbank failures, and landslides, where a type of "catastrophic event is necessary in order to maintain a "critical equilibrium" (Bak 1997). As a SOC toy model, suppose we drop some grains of sand on a small, round table: Grain upon grain, a sand-pile forms in the shape of a cone. It keeps growing until the table surface is completely covered, and over it, a defined cone shape is reached. At this point, adding new grains won't change the pile shape any more. In fact, adding grain after grain will cause, suddenly and occasionally, landslides on the pile surface so that the target shape is maintained. So, the pile organizes itself (as a definite cone) and maintains its critical equilibrium through "catastrophic events" (i.e., landslides).

According to the SOC model, "stressors" for an ASD subject would act as the grains of sand. As grains drop on the pile, so stressors occur in her or his life (Stark et al. 2015). As they interact (in a complex way) in the sand-pile, so they interplay within a subject. Similarly to a landslide in a sand-pile, they cause a form of crisis behavior in an ASD subject. Namely, problem behavior is a type of symptom resulting from a complex network of interacting stressors.

Notably, scientific literature can identify stressors as internal (e.g., a feeling of hunger), as well as external (e.g., noise from a car's horn) (Mesibov et al. 2004). Now, it is reasonable to consider such stressors as being related to "intentional information" making sense in Davidson and Dennett interpretivism (Davidson 1984; Dennett 1987), just as suggested by Borsboom et al. in the target article. As described in the network models, a symptom (e.g., a type of communication ability impairment) can interact with stressors (e.g., feeling of hunger) while another symptom (e.g., hyper-reactivity to sounds) can interact with other stressors (e.g., a car's honking horn). This complex interplay of symptoms and stressors could lead to, for example, a challenging behavior, a peculiar symptom within the family of what has been called problem behaviors (in turn, feeding the social impairment, another fundamental ASD feature, thus taking place within the network).

Therefore, in our opinion, the network model proposed by Borsboom et al. seems to be highly consistent with the body of published literature regarding ASD problem behaviors and, particularly, with the recently proposed SOC model.

Interestingly, the SOC model could suggest new possible insights. In fact, according to Borsboom et al., "'to suffer from a disorder' means 'to be trapped in the stable state of a self-sustaining symptom network'" (sect. 3, para. 3). The SOC perspective could additionally suggest the network as a self-organized system, in which symptoms are "necessary" events that maintain a type of critical equilibrium (Ramos Sassi Piqueira 2011). For instance, as for ASD, can we think of panic attacks as "landslides" in the "sand-pile" of panic disorders, or of sad moods as landslides in the sand-pile of major depression?

Hence, a pathological state would not be just a "stable state" but a "critical state of equilibrium" of a self-organized complex interaction of stressors and symptoms, just like grains in a sand-pile.

The biology of mental disorders: What are we talking about?

Alfonso Troisi

Department of Systems Medicine, University of Rome Tor Vergata, 00133 Rome, Italy.
alfonso.troisi@uniroma2.it
http://medschool.uniroma2.it/2016/05/31/alfonso-troisi/

doi:10.1017/S0140525X1800119X, e29

Abstract

After the Darwinian revolution, biology is not only the study of the operation of structural elements (functional biology), but also the study of adaption and phylogenetic history (evolutionary biology). From an evolutionary perspective, the biology of mental disorders is not just "neurobiology and genetic constitution" but also adaptive reactions to adverse situations. Evolutionary explanations of mental disorders are biological and non-reductionist.

Borsboom et al. contrast the symptom network model with biological explanations of psychopathology and conclude that mental disorders cannot be explained in terms of biology. In fact, their conclusion is based on a biased and flawed description of the biological study of human mind and behavior. They reduce biological explanations of mental symptoms and disorders to abnormal neurobiological mechanisms and defective brain circuits. Even though such a reductionist model is still frequently encountered in the research and clinical literature, it is not illustrative of how biological explanations can improve our understanding of the origin of mental disorders. After the Darwinian revolution, biology is not only the study of the operation and interaction of structural elements, from molecules up to organs and whole individuals (functional biology). Modern biology is also the study of adaption and phylogenetic history (evolutionary biology) (Mayr 1982). When applied to psychiatry, the evolutionary approach shows that the biology of mental disorders is not just "neurobiology and genetic constitution" but also the study of evolved reactions to adverse

environmental circumstances, including adaptive symptoms and calibrated life history strategies (Troisi 2017).

Keller and Nesse (2006) introduced and tested a new framework for understanding the adaptive significance of depressive symptoms. Their hypothesis (the "situation-symptom congruence" hypothesis) predicts that, if different depressive symptoms serve different evolved functions, then different events that precipitate a depressive episode should give rise to different symptom patterns that increase the ability to cope with the adaptive challenges specific to each situation. The hypothesis was tested by asking 445 participants to identify depressive symptoms that followed a recent adverse situation. Guilt, rumination, fatigue, and pessimism were prominent following failed efforts; crying, sadness, and desire for social support were prominent following social losses. These significant differences were replicated in an experiment in which 113 students were randomly assigned to visualize a major failure or the death of a loved one. The results of the study confirmed the prediction that symptoms eliciting comfort (e.g., crying) should be especially prominent when social bonds are threatened, lacking, or lost, whereas symptoms dissuading the individual from pursuing current and potential goals (e.g., pessimism and fatigue) should arise when the environment is unpropitious and future efforts are unlikely to succeed. Strong support for the evolutionary hypothesis came subsequently from a study of 4,856 individuals who experienced different patterns of depressive symptoms associated with nine categories of adverse life events (Keller et al. 2007). These findings offer a biological explanation of the origin of depressive symptoms without implying defective brain circuits and argue against "reductive models that suggest that neural and molecular levels are the only ones at which we will find true explanations for the phenomenon of clinical depression" (Keller et al. 2007, p. 1528).

Environment (and especially social environment) is a crucial variable in evolutionary explanations of mental disorders, which are at the same time biological and non-reductionist. The concept of developmental plasticity is based on the understanding that the phenotype and genotype do not have a fixed relationship and that the phenotypic attributes of individuals are affected by developmental processes. The relationship between environmental influences and the consequential phenotypic change may have directional components of adaptive value (Belsky 2016).

Life history theory is a mid-level evolutionary framework that explains individual differences in various correlated behaviors and outcomes such as mating strategies, risky behaviors, reproductive development, and health. These phenotypic variables are conceptualized as indicators of individual differences along a fast-slow life history continuum. Individuals adopting a fast strategy (that theoretically is most adaptive under harsh and unpredictable environmental conditions) employ short-term mating tactics, engage in risky behaviors, are less future oriented, and devote less time to their offspring (Chua et al. 2017). Consistent with the notion that evolution is incapable of forward thinking, life history theory offers a biological explanation of how early social experiences provide the cues that calibrate individuals' systems for adapting to their future social environments. Adverse early experiences (e.g., childhood maltreatment) tend to regulate development toward a so-called fast life history characterized by behaviors that would maximize fitness when life is short and the future cannot be controlled or predicted, including heightened risk-taking, early reproduction, and decreased prosocial traits. The life history theory has been especially productive in helping to understand unstable attachments, social antagonism, sexual promiscuity, and early reproduction in some individuals not simply as behavioral disorders, but as a natural outcome of context-sensitive developmental mechanisms that are adaptive for genes, even if not for individual and social well-being (Del Giudice et al. 2011). Several personality disorders, as currently described by psychiatric classifications, seem to reflect "fast life history strategies" caused by a combination of genetic characteristics and early unpredictable or adverse experiences (Brüne 2016). An important implication of life history theory is that the biological prevention of psychiatric disorders should target social risk factors like poverty, maltreatment, and affective deprivation.

Mental health professionals continue to employ a mind-brain dichotomy when reasoning about clinical cases. In clinical discourse, references to "mind" and "brain" have become a form of code for different ways to think about the etiology of psychiatric disorders and their treatment. The etiology and pathogenesis of "biological" disorders would depend mainly on genetic predisposition and neural dysfunction, whereas environmental factors and interpersonal problems would be the main causal factors of "psychological" disorders. The divide between mind and brain fades away if one reasons in terms of adaptive function and employs the concept of behavioral systems. A behavioral system (e.g., the attachment system) can be defined as an integrated group of functionally related components consisting of specific psychological processes, physiological mechanisms, anatomical structures, and genetic influences. No component has an intrinsic priority over the others, and malfunctioning can originate anywhere in the system. When this happens, malfunctioning propagates to all of the components. From such a perspective, it does not make any sense to say that a psychiatric disorder is a "brain disorder" or a "mind disorder."

What's in a model? Network models as tools instead of representations of what psychiatric disorders really are

Hanna M. van Loo[a] and Jan-Willem Romeijn[b]

[a]Department of Psychiatry, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands and [b]Faculty of Philosophy, University of Groningen, 9712 GL Groningen, The Netherlands.
h.van.loo@umcg.nl j.w.romeijn@rug.nl
http://www.rug.nl/staff/h.m.van.loo/ http://www.philos.rug.nl/~romeyn/

doi:10.1017/S0140525X18001206, e30

Abstract

Network models block reductionism about psychiatric disorders only if models are interpreted in a realist manner – that is, taken to represent "what psychiatric disorders really are." A flexible and more instrumentalist view of models is needed to improve our understanding of the heterogeneity and multifactorial character of psychiatric disorders.

This commentary addresses Borsboom et al.'s claim that "if mental disorders are indeed networks of causally related symptoms, reductionist accounts cannot achieve [success]" (target article, Abstract). Borsboom et al. pose that psychiatric disorders are, in essence, networks of symptoms rather than brain states, and that this rules out reductionism. This claim rests on a particular

realist interpretation of disease models, namely that models represent what diseases, in essence, are. We take issue with this interpretation. Models should not be understood as representing the true nature of psychiatric disorders, but as tools to improve our understanding of different aspects of these disorders. A realist interpretation of models might be detrimental for progress in psychiatry, as it may invite exclusivity in the use of models.

Before we engage in argument, we want to align with Borsboom et al. on their anti-reductionist and multifactorial conception of mental disorders. We share many of their views: An exclusive focus on brain processes will hamper progress in psychiatry, because most psychiatric disorders result from interacting biological, psychological, and environmental factors (Kendler 2014). Research focusing on all of these levels will thus benefit psychiatric science more than an exclusive focus on underlying brain processes. In addition, many psychiatric disorders describe a heterogeneous group of patients in terms of symptoms, etiology, and course of illness (Van Loo et al. 2012). To deal with the heterogeneous and multifactorial nature of psychiatric disorders, psychiatry needs flexibility in its research methods. That is why we are critical of Borsboom et al.'s line of reasoning. The implicit realism in their arguments might invite another kind of exclusivity, namely in the use of network models. We believe this might have similar adverse effects on progress in psychiatric science.

Borsboom et al. use a realist view about models

In their first argument against reductionism, Borsboom et al. move from two observations (psychiatric symptoms are highly correlated, and no single biological cause has been found to explain these relations) and a common-sense notion of causality (it seems natural that, e.g., insomnia and fatigue are directly causally related) to a preference of network models over latent variable models. They then argue that if psychiatric disorders are indeed symptom networks, reductionism will fail, because in symptom networks there is no common cause. However, this argument only works if we adopt realism about the models.

First, concerning latent variable models in psychiatry, the authors are right to point to a strong association between such models and reductionism. They invite a physicalist and causal interpretation: The latent variable gets the role of cause – hence the terminology of "common cause models" – and then this cause is imagined to have some physical realizer, for example, a neurobiological process. However, the association between latent variable models and reductionism is by no means a conceptual necessity (Bringmann & Eronen 2018). It is perfectly possible to use latent variable models without interpreting the latent as pointing to an entity in the world, let alone as cause. Moreover, even if we interpret the latent as causal, we can still avoid physicalism if we drop the specific physical realization of the latent cause. Thus, the so-called common-cause models are inviting reductionism only if we give them a particular realist interpretation.

Second, consider the claim that network models make physicalist reductionism unfeasible. As the authors briefly indicate themselves, the use of network models is, in principle, compatible with a reductionist viewpoint, but if symptom networks are taken to be realist representations, this kind of "network reduction" becomes rather implausible (see target article, sect. 4). However, we could employ network models in a more pragmatic manner, namely as instruments of prediction and control, without committing to the idea that they provide a picture of the world. Then it does not say much that network models lack a representation of biological causes: a complete picture is not expected. Anti-reductionism follows from network models only if those models are interpreted as providing a realist account.

Summing up, latent variable models are reductionist, and network models are anti-reductionist, *only* if we give those models a realist reading. Rather than maintaining realism in the service of an argument against reductionism, we would do better to reconsider this realism itself.

Alternative view on status and function of models

So how should we think of the status of models? Following a well-developed line of thinking in the philosophy of science, we think that models function as instruments of investigation instead of being exact representations of reality (Morgan & Morrison 1999). Statistical models, but also other models – the Bohr model of the atom, the double-helix model of DNA, general equilibrium models of markets – typically involve aspects of both theory and data. A model is an autonomous tool, because it is partially independent of both of them. Precisely because models are partly independent of both, they can be used as instruments of exploration in both domains. As such, models play a mediating role, helping us connect empirical facts about disorders to theoretical accounts of them, thereby performing a variety of tasks.

In the context of psychiatry, this means that psychiatric disorders must not be thought of as being symptom networks, nor as being brain disorders, but that different aspects of psychiatric disorders can be investigated using different types of models. This ties in with our general idea that our attitude towards models must be guided by what will bring psychiatric science further (Van Loo & Romeijn 2015). For example, network models are well-suited to investigate complex dynamic interactions among multiple variables in the development of psychopathology (Wichers et al. 2017). Latent variable techniques instead may be used to achieve data reduction, and to explore underlying latent factors (Bringmann & Eronen 2018). Thinking in terms of a common cause led to the recent discovery of anti-NMDA receptor encephalitis – a rare auto-immune disorder that leads to psychotic symptoms and agitation (Dalmau et al. 2007). An instrumentalist view of models offers more flexibility for switches between and divergence among research strategies, and this is needed to improve our understanding of the heterogeneity and multifactorial character of psychiatric disorders.

Acknowledgments. In the preparation of this commentary, we have benefited greatly from discussions with Drs. Markus Eronen and Laura Bringmann, Faculty of Behavioural and Social Sciences, University of Groningen.

Families of network structures – we need both phenomenal and explanatory models

Tony Ward and Ronald Fischer

School of Psychology, Victoria University of Wellington, PO Box 600, Wellington, New Zealand.
Tony.Ward@vuw.ac.nz Ronald.Fischer@vuw.ac.nz
https://mindcultureevolution.com/

doi:10.1017/S0140525X1800122X, e31

Abstract

Symptom network models (SNWMs) play an important role in identifying but not explaining *patterns* of symptoms. We discuss underlying assumptions of SNWMs and argue that they represent phenomenal models, best suited to detecting patterns among symptoms. SNWMs need to be supplemented with mechanistic models that provide constitutive and etiological explanations of each symptom (network nodes) *once* relevant patterns have been identified.

The target article by Boorsboom et al. infuses new ideas into a longstanding debate on the definition, classification, and explanation of mental disorders. It has the potential to create novel ways of mapping symptom structures, understanding comorbidity, and tracking the emergence and development of symptoms. We argue that, while the symptom network model (SNWM) can play an important role in identifying *patterns* of symptoms, it offers less when it comes to their explanation.

In their article, Borsboom et al. present five, overlapping, theoretical and empirical arguments critiquing the neurocentric, reductionist approach to defining and explaining mental disorders, and propose that a SNWM approach should be adopted in its place. In essence, they argue that:

1. Mental disorders are not disease entities, but are best understood as dynamic networks of symptoms that trigger each other rather than as "things."
2. Contrary to the claim of neurobiological conceptions of mental disorders, the mind cannot be reduced to the brain (e.g., "mental disorders are not brain disorders at all," sect. 1, para. 5).
3. Understanding the intentionality of symptoms (what symptoms mean at a phenomenal level) is key.
4. External-contextual factors are important causes of mental disorders and initially activate symptom networks that then become self-sustaining.
5. In light of the above arguments, multilevel explanatory strategies that cover neurobiological, phenomenological, and social-cultural domains are required in order to describe and explain symptom networks.

Borsboom et al. conclude that SNWMs can deal with these challenges and define, classify, and explain mental disorders more adequately than reductionist neurobiological theories.

A first critical point is that these arguments are only loosely aligned, and although collectively seeming to provide a powerful challenge to neurological conceptions of mental disorders, do not in fact do so. The reason for this is that they individually fail to provide good reasons for adopting a symptom network approach to explaining psychopathology. In brief:

(a) The primary causes of mental disorders could be biological in nature without being disease entities.
(b) Reductionism is a metaphysical argument, not an epistemological (explanatory) one. Therefore, it is not seriously considered by contemporary cognitive neuroscientists, who agree that personal and subpersonal psychological levels of explanation provide unique and valuable insights into the mind (Eliasmith 2013).
(c) Privileging the intentionality of symptoms on *a priori* grounds is to commit oneself to folk psychological conceptions of the mind and risk begging important theoretical questions.
(d) Multilevel explanations can be developed within the context of a view of the mind as a material system. In other words, a strategy of explanatory pluralism (for pragmatic reasons) can be neutral with respect to the nature of the mind from a metaphysical perspective.

In our view, the failure to be clear about the limits of these distinct arguments means that there is a temptation to structure the debate between neurobiological and intentional level explanations as one between reductionist versus non-reductionist approaches. This is not helpful and represents a false dichotomy.

A second critical point is that the SNWM approach is best viewed as simply one phase of scientific inquiry, and requires supplementing by other types of methodological and theoretical models. There are a number of different aims evident in scientific practice, including prediction, description, classification, and explanation – both causal and compositional (Craver & Kaplan, in press). Different kinds of models are constructed in service of these aims. *Phenomenal* models are essentially descriptive in nature and aim to capture patterns such as co-occurring events or mental states and behavior (Hochstein 2016). *Explanatory* models set out to depict the components, their relationships, and organisation in mechanisms that underlie phenomena (Craver & Kaplan, in press). For example, a phenomenal model of depression seeks to describe the symptoms associated with the syndrome and their temporal relationship to one another, while explanatory, mechanistic models provide insight into their underlying constituents, processes, and organization (e.g., serotonin dysregulation, neural activation patterns, attentional biases, negative self-evaluations; Beck & Bredemeier 2016).

In our view, the symptom network model is usefully construed as a phenomenal model and therefore is best suited to detecting patterns among symptoms, as opposed to representing the mechanisms that constitute them (Craver & Kaplan, in press; Hochstein 2016). Identifying the covariance between symptoms and the order in which they appear is an important scientific task and provides a description of phenomena that can subsequently be a focus of explanation. Symptom network models have a crucial role to play in the scientific inquiry process, but a limited one.

Thinking of the SNWM in these terms allows room for developing constitutive and etiological explanations of each symptom (nodes in the network) *once* the relevant patterns have been identified. Symptoms are complex structures constituted by entities and processes at different levels of organization. Thus, it is reasonable to construct explanatory models, each directed at different aspects of symptoms, and collectively providing an overall explanation of them (i.e., phenomena). These mechanistic models shed light on the way different symptoms exert their effects and influence each other.

A final critical point is that Boorsboom et al. overlook the plurality of models required to engage in research. Even from a SNWM perspective, a *family* of network models is needed to identify and model relationships among symptoms, each with a different goal: modeling symptom structure within mental disorders, symptoms clusters (irrespective of diagnosis), temporal shifts (e.g., progression within individuals), and so on. Networks of explanatory and descriptive models are necessary

to provide a comprehensive explanation of symptoms and their grouping into disorders, each focusing on different types of factors and processes.

Authors' Response

Reductionism in retreat

Denny Borsboom,[a] Angélique O. J. Cramer,[b] and Annemarie Kalis[c]

[a]Department of Psychology, University of Amsterdam, 1018 WT Amsterdam, The Netherlands; [b]Department of Methodology and Statistics, Tilburg University, 5000 LE Tilburg, The Netherlands; and [c]Department of Philosophy and Religious Studies, Utrecht University, 3512 BL Utrecht, The Netherlands.
d.borsboom@uva.nl aoj.cramer@gmail.com A.Kalis@uu.nl
https://www.dennyborsboom.com https://www.aojcramer.com
https://www.uu.nl/staff/AKalis

doi:10.1017/S0140525X18002091, e32

Abstract

We address the commentaries on our target article in terms of four major themes. First, we note that virtually all commentators agree that mental disorders are not brain disorders in the common interpretation of these terms, and establish the consensus that explanatory reductionism is not a viable thesis. Second, we address criticisms to the effect that our article was misdirected or aimed at a straw man; we argue that this is unlikely, given the widespread communication of reductionist slogans in psychopathology research and society. Third, we tackle the question of whether intentionality, extended systems, and multiple realizability are as problematic as claimed in the target article, and we present a number of nuances and extensions with respect to our article. Fourth, we discuss the question of how the network approach should incorporate biological factors, given that wholesale reductionism is an unlikely option.

R1. Introduction

Consider the following thought experiment. Imagine that you lived in a world – call it *Reductionia* – in which compelling evidence existed for the thesis that mental disorders are, in fact, brain disorders. Reductionia would be much like our world, in the sense that the literature would be equally rife with statements affirming the neural basis of mental disorders; however, in contrast to our world, Reductionia would also be rife with strong evidence for this reductionist thesis. For instance, in Reductionia, the etiology of mental disorders would be understood in terms of broken brain circuits and cerebral deficits, treatment would be directed at deviant neural pathways, and genetic profiling would inform treatment with precision drugs that effectively relieve suffering from mental disorders.

Now suppose that a number of scholars in Reductionia wrote a paper in an influential journal – call it the *Sciences of Brain and Behavior* (SBB) – and argued for the thesis that mental disorders are not brain disorders at all. Suppose further that, miraculously, this paper got through peer-review at SBB, and that dozens of scholars were allowed to comment on it.

We would like you take a moment to imagine what these comments would look like. We think, and we expect you will agree, that the SBB commentators in Reductionia would most likely tear the paper apart by presenting scientific evidence in support of the thesis that mental disorders are brain disorders. Responses might have been comparable to those that would befall an author who, in our actual world, published a paper saying that, for example, Parkinson's is not a brain disease or that Down's syndrome is not due to a genetic condition. Commentators would simply point to the evidence and chastise the authors for their ignorance of the literature.

Now compare the SBB commentaries in Reductionia to the BBS commentaries to our target article in the actual world. The difference is astonishing. None of the commentators appears able to point to convincing evidence that, generically speaking, mental disorders are brain disorders; in fact, it seems that most commentators do not even bother. This brings us to the first important conclusion of this response to commentaries: *The thesis that mental disorders are brain disorders enjoys no appreciable support.*

Importantly, this establishes that the reductionist position on mental disorders as brain disorders does not represent a scientifically justified conclusion, as is often supposed in the popular and scientific literatures, but instead is a hypothesis. Our target article dealt with the question of how likely this hypothesis is, given the assumption that network theories of mental disorders are broadly correct. Thus, as pointed out by some commentators (e.g., **Tabb**), the main argument defended in our article is a conditional one: *If* mental disorders are causal networks of symptoms (as we have argued elsewhere), there are strong reasons to believe that reductionist explanations of mental disorders are blocked. Of course, this does not mean that explanatory reductionism would be vindicated if the network theory might turn out incorrect; there are many other reasons to be skeptical about reductionism (**Eronen & Bringmann**) and, especially, about what Tabb denotes as the *triumphalist* variant of the position. Our paper argued that network theory is sufficient to block explanatory reductionism, but not that it is necessary. Also, a failure of explanatory reductionism does not imply that certain symptoms or functional relations between them could not receive a partial reductive analysis (**Tabb**; **Hur, Tilman, Fox, & Hackman** [**Hur et al.**]; **Ward & Fischer**) or that some conditions of psychopathology could not have a physical cause, as **Van Loo & Romeijn** point out. It does mean, however, that we should considerably tone down our expectations on what we can reasonably expect to learn from biological approaches of psychopathology.

How plausible is the premise in the conditional argument presented – that is, the idea that psychopathology should be approached from a complex systems perspective, using network models, dynamical systems, and associated techniques? Several commentators embrace our suggestion that it does indeed make sense to think of mental disorders as networks, and also agree with us that, from such an approach, it follows that mental disorders cannot be brain diseases. For example, **Ioannidis** argues that this "new narrative" can explain why pharmaceutical treatments have only limited effects, and that network approaches suggest alternative forms of clinical research focusing not on etiological pathways but on a variety of pragmatically relevant treatment outcomes such as quality of life, relationships, and professional success. According to Ioannidis, we need to think of mental disorder

"as an evolving, ever-changing challenge for society-at-large." To this, **Baran** adds that reductionist explanations in psychiatry heavily rely on animal "models" of mental disorders – models which cannot do justice to the "massively multifactorial system networks which go awry in mental disorders." **McNally**, stating that the network approach has "dramatically changed" the landscape of models of psychopathology, even suggests that biology itself might move into an anti-reductionist direction in response to the "network takeover." Moreover, in support of our argument, several commentators emphasize that network approaches are highly consistent with recent insights on different kinds of disorders. For instance, **Hens, Evers, & Wagemans (Hens et al.)** and **Tonello, Giabocci, Pettenon, Scuotto, Cocchi, Gabrielli, & Cappello (Tonello et al.)** both defend the view that network models might offer a valuable analysis of autism spectrum disorders (see also Deserno et al. 2017; 2018); **Hyland** thinks it could also be fruitfully applied to certain complex functional disorders such as *irritable bowel syndrome* and *chronic fatigue syndrome*. **Field, Heather, & Wiers (Field et al.)** emphasize how the concept of addiction fundamentally depends on the cultural and historical context, and argue that our approach offers an opportunity to fundamentally rethink the "brain disease model" of substance abuse, a position also supported by **Ross'** arguments.

Some commentators even claim that our objections to reductionism aren't radical enough. For example, **Ross** thinks that we still give reductionist approaches too much credit: He convincingly argues that, in some instances, psychopathology does not require any kind of biological abnormality to be present. The argument is that, in addictive disorders, the relevant brain mechanisms fulfill precisely the characteristic functions they were selected for in our evolutionary past, but these neural systems are "hijacked" by our current environment of roulette tables and gambling machines. Ross' example is well taken and matches many of the symptom-symptom relations seen in network analyses of mental disorders (e.g. insomnia → fatigue, anxiety → avoidance) which typically come across as prosaic precisely *because* they reflect normal systems in our biological makeup. **Desseilles & Phillips** also hold that the network approach understates, rather than overstates, the complexity of mental health and disorder, as network models do not distinguish between different types of relevant factors and cannot establish "real" causal relations.

Thus, as a first conclusion, even though certainly not all commentators embrace a network approach (see, e.g., **van Loo & Romeijn**; **Eronen & Bringmann**; **O & Brüne**), it is remarkable that none of the commentators, not even our more critical opponents, attempt anything like a spirited defense of the reductionist paradigm in psychiatry. Instead, almost all critics point their arrows at other targets. For example, several authors claim that the explanatory reductionism we criticize is, in fact, a straw man, and they suggest more nuanced positions such as non-reductive materialism should be defended; others suggest that network models merely represent a phenomenological level of analysis. However, it appears that nobody wants to take up the reductionist gauntlet. This, we think, is a useful signal to the world outside of academia because it communicates an important message from the ivory tower to society: Despite the many sources that suggest mental disorders to be brain disorders, reductionism is not a viable scientific position.

R2. Who or what is the straw man?

Several commentators remark that our target article was not properly directed, in the sense that we attack an outdated straw man that nobody believes in anymore. For instance, **Hur et al.** suggest that we "use evidence against extreme reductionism and common causes to devalue clinical and translational neuroscience approaches – effectively throwing the baby out with the bathwater." Similarly, **Troisi** states that the reductionist model as we defined it "is not illustrative of how biological explanations can improve our understanding of the origin of mental disorders." Finally, **Ward & Fisher** argue that reductionism "is not seriously considered by contemporary cognitive neuroscientists, who agree that personal and subpersonal psychological levels of explanation provide unique and valuable insights into the mind." The impression one gets from these remarks is that the actual research in biological approaches to clinical psychology and psychiatry is much more sophisticated than the simple search for a common biological cause of symptomatology, and does in fact already incorporate psychology and environment in integrative ways.

Surely, there exists research for which this picture is accurate, and we would in no way want to suggest that the crude form of reductionism we sketched in our target article is invariably entertained among researchers working on the biological factors involved in mental disorders. Still, it is hard to avoid the impression that neuroscience and genetics, in this respect, are Janus-faced. While the researchers who feature in the current discussion forum insist that nobody really believes in explanatory reductionism, at the very same time many of their colleagues are dominating the media with claims that clearly rest on the acceptance of the thesis that mental disorders are brain disorders in a literal and unqualified sense.

In our target article, we gave the example of NIMH leaders who, as late as 2015 and in full view of the absence of any strong evidence supporting their claim, felt confident declaring that mental disorders are brain disorders (Insel & Cuthbert 2015), and we discussed a number of other influential sources who are on the record as embracing full-fledged versions of explanatory reductionism. We think it is therefore rather remarkable that so many commentators feel confident in stating that this form of reductionism is a straw man. If this is really the case, why is it so easy to find examples of government-funded scientific organizations that communicate the message to the public that mental disorders are brain disorders? The Queensland Brain Institute, to give but one example, simply lists depression as a brain disorder on its website; and the first sentence under the header "What causes depression?" is "The neurobiology of depression is still poorly understood," suggesting that the answer to the question, whatever it may be, must in fact turn on neurobiology.[1] Similarly, in 2017, the Dutch National Institute of Public Health and the Environment published research concluding that one in every four citizens "suffers from a brain disorder."[2] The count, which was widely publicized, indiscriminately included all mental disorders for which data were available. Those who think this type of talk is limited to press releases and promotion material may consider some cutting-edge science. The most recent genetic work on depression, published in *Nature Genetics*, found that 44 genetic loci are significantly correlated with depression measures which, taken together, explain 1.9% of the variance in estimated depression liability (Wray et al. 2018). These results lead the authors of the otherwise excellent study to draw a stunning primary conclusion from the data: namely, that "major depression is a brain disorder."

We are not cherry-picking. It is not at all hard to find these examples, as anyone can verify by searching the Internet for information on mental disorders. The notion that mental disorders are

genetically encoded brain disorders is everywhere around us. It dominates the organization of research, it dominates teaching, and it dominates the media. The central problem is not even that the thesis is necessarily false – as we stated in our target article, in the future we may in fact witness the kinds of breakthroughs that would establish that mental disorders are brain disorders; this is, in our view, spectacularly unlikely but not impossible. The central problem is dogma: The reductionist hypothesis is not treated as a scientific hypothesis, but as an almost trivial fact.

Given that so many examples of outspoken reductionism exist, we feel hesitant to accept that explanatory reductionism is a straw man. Rather, we submit that the idea that mental disorders are brain disorders has somehow become a background assumption of our modern society – an undisputed member of the cabinet of scientifically respectable facts, on par with "the world is round" and "life developed through evolution." Even though, as said, none of the commentators attempts to defend reductionism explicitly, reductionism as a background assumption clearly does shimmer through in several of the responses to our target article. **Hur et al.** state casually that "mental illness is undeniably based in brains and genes." **Redish, Kazinka, & Herman** (**Redish et al.**) argue that "[b]ecause all behavior arises from brain function, neurobiology is obviously critical for understanding psychiatric phenomena." **Pernu** declares that "[s]symptoms, however, are mere signs, marks of the underlying disorder or illness; and even in clinical practice, the ultimate aim is not just to remove the symptoms, but to cure the physiological condition they stem from." One does not get the impression that these commentators, in their own view, are launching spectacular scientific hypotheses. If they looked at their statements in this way, the commentators would probably have felt compelled to provide arguments for these theses. Rather, they make the impression that they are stating the obvious.

Explanatory reductionism, however, is not obvious. It is not a fact but a hypothesis that mental disorders originate in the brain; it is not a fact but a hypothesis that there are genes "for" mental disorders; and it is not a fact but a hypothesis that finding out "what goes wrong in the brain" is a necessary condition for progress in the science of mental disorders. It is a realistic possibility that increased understanding of the neuroscience involved in mental disorders will in fact establish that they are *not* brain disorders. The network theory opens up a general and scientifically defensible prospect that shows how this could be the case, as we argued in our target article.

R3. Networks = notworks?

For some commentators, the network approach does not work, for various reasons and in various ways (**Elbau, Binder, & Spoormaker** [**Elbau et al.**]; **Eronen & Bringmann; Hur et al.; Ward & Fischer**). One returning theme concerns the supposed "thinness" of network models – in other words that, supposedly, the network framework "fails to provide a deeper explanation … of where those patterns come from" (Hur et al.). Ward & Fischer acknowledge the usefulness of a network approach as a model at the level of phenomenology, that can be employed to "detect patterns among symptoms," but they argue that the approach does not represent "the mechanisms that constitute [symptoms]."

To analyze this situation, it is important to distinguish between a network as a *theoretical model* that aims to represent phenomena in the world (e.g., see Borsboom 2017; Cramer et al. 2016) and as a *statistical model* that one estimates on empirical data (e.g., see Epskamp & Fried 2018; Van Borkulo et al. 2014). If the commentators, when noting the thinness of symptom networks, are referring to the fact that the statistical model does not churn out a set of empirical, mechanistic facts about the world – for example, the connection between depressed mood and suicidal thoughts is primarily driven by a lack of self-worth – then we fully agree. Statistical models contain variables connected by parameters. A model fitting exercise will return parameter estimates that describe relations between variables, but it will not generate mechanistic explanations for why these relations exist. Statistical models, in this respect, indeed should be seen as delivering evidence on the *presence* of causal processes rather than on the *nature* of these processes.

However, the conditional argument set up in our paper does not concern such statistical models; it instead concerns network theories that do address the nature of psychopathology – namely, that psychopathological conditions are alternative stable states of complex networks formed by symptoms and interactions between symptoms (Borsboom 2017; Cramer et al. 2016). For instance, Cramer et al. (2016) showed that, for major depression, we can fairly precisely explain how an episode of major depression can come about, namely by the dynamics in a strongly connected network in which symptoms cause and maintain one another. The connections between symptoms in this theoretical model *do* lend themselves well for coupling with real-world mechanisms, including biological ones (see also **Ross**). For example, the theoretical connection between insomnia and fatigue is supported by a real-world mechanism, namely the homeostatic processes instantiated by the biological clock; the connection between hallucinations and anxiety likely involves fear mechanisms deploying the amygdala (**Hur et al.**); the connection between gambling and money shortage undoubtedly deploys the mechanical specifications of the Roulette table. The network theory of psychopathology holds that, in psychopathology, these real-world mechanisms reinforce each other to such a degree that a symptom network becomes self-sustaining (Borsboom 2017). Importantly, however, this theory is not necessarily tied to statistical network models; this can be seen from the fact that one can specify the theoretical idea independently of statistical modeling schemes whatsoever: it could, for instance, also be specified using deterministic models (e.g., dynamical systems models). Current statistical network modeling techniques should therefore not be seen as instantiations of this theory but as statistical devices that chart the structure of symptom-symptom connectivity (e.g., see Boschloo et al. 2015); the hope in using such models is not that these analyses will magically return a substantive theory all by themselves, but that they will uncover patterns of network connectivity between symptoms that can subsequently be used to inform theoretical models. However, statistical models are not themselves substantive theories, and the network approach is not defined by or limited to one particular type of statistical model.

We also can see this clearly from the fact that the relation between network theory and network model is not one-to-one: Network theories *may* map into statistical network models, but they do not *necessarily* do so. Recently, for instance, it was shown that a Curie-Weiss model – a special case of the Ising model in which all connections have the same strength – is statistically equivalent to the Rasch model (an important latent variable model; Epskamp et al. 2018; Marsman et al. 2017). This means that one can use latent variable models to specify statistical consequences of network theories, and one can use network

models to specify statistical consequences of latent variable theories. As one example, it has been recently suggested that network models are particularly well suited to identify the dimensionality of the latent factor space (and even outperform factor analysis in this respect; Golino & Epskamp 2017; Golino & Demetriou 2017). These equivalence theorems also explain why, as **Jayawickreme, Rasmussen, Karasz, Verkuilen, & Jayawickreme** [**Jayawickreme et al.**] note, network models and factor models can lead to very similar results; when network models are applied to data of questionnaires that have been developed using psychometric models (e.g., item response theory of factor analysis), they will mirror the structure of these models (i.e., every factor will produce a clique in the network). In view of this, Jayawickreme et al. make the important methodological point that one should be careful in selecting which variables to include in network models if the findings are to inform network theories; for instance, if one uses sets of items that are selected to conform to a unidimensional factor model (in practice, this typically results in sets of items that all correlate with about the same strength), one's network analysis will be predictably homogeneous (i.e., all nodes will feature roughly equal interconnections so that all nodes are equally central; Marsman et al. 2017). However, this may not always be informative with respect to the issue of how disorders are structured.

Therefore, we should take care in distinguishing network theories from network models. We think that complaints on the thinness of networks, or the idea that they are "merely phenomenological," are relevant to statistical network models, which indeed only represent conditional associations between symptoms. As **Van Loo & Romeijn** argue, such models can be interpreted in an instrumentalist fashion. And if one engages in such an interpretation, it is indeed somewhat hard to see exactly what is meant in a conditional argument that proceeds from the subjunctive conditional "if the network theory of psychopathology were true," because truth is a property of semantically interpreted theories, not of statistical models per se. However, such an interpretation is not suited to network theories, interpreted as formalized theories that represent actual problems people have (i.e., the problems encoded as symptoms in diagnostic manuals such as DSM-IV) and causal processes by means of which these problems maintain, promote, or inhibit each other. Interpreted in this way, network models are not susceptible to arguments that they are too thin in the sense of being merely about phenomenology. Symptoms, and relations between them, allow any depth of theoretical explanation one might envisage, including those ventured in learning theories, cognitive schema accounts, psychodynamic approaches, neuroscientific theories, and, ultimately, evolutionary theories (**O & Brüne**). For this reason, network theories do not rule out hierarchical explanatory accounts, as **Oller** suggests. However, what *is* very clear when one considers plausible candidate accounts that connect symptoms, is that it is spectacularly unlikely that *all* of these relations will be amenable to a neuroscientific account, as they will require reference to the world external to the body as well as the (partly culturally loaded) content of mental states, of which several commentators provided important examples (e.g., see **Crafa & Nagel; Field et al.; Jayawickreme et al.**). Network theories therefore almost by necessity lead to pluralism, as different symptoms (and relations between symptoms) require analyses at different levels.

R4. Intentionality, multiple realizability, and extended systems

Many commentators agree with the claim that symptoms often bear intentional content and that this constitutes a challenge for reductionists. As **Ross** rightly remarks, intentionality is a *general* obstacle to reductionist explanation, and not one that is specific for the explanation of mental disorder. Nevertheless, intentionality pops up as a specific challenge for reductionist explanation of mental disorders when it is claimed (as we do) that mental disorders should be understood *in terms of causally interrelated symptoms*. In other words, it is precisely the network approach to mental disorders that brings intentionality to the fore as a challenge for reductionism. **Eronen & Bringmann** are similarly correct to claim that "defending anti-reductionism does not require taking a network perspective," but this has no consequences for our argument because our claim rests on the sufficiency, but not the necessity, of the premise that network theory correctly describes psychopathology: We argue that the network approach constitutes a specific obstacle to reductionist explanation, because it locates the causal nexus at the level of symptoms, which are often inherently intentional. This has little to do with whether other obstacles to reductionism also exist (which is undeniably the case).

To avoid this difficulty, **Eronen & Bringmann** and **Van Loo & Romeijn** both bring up the idea that one could adopt an instrumentalist understanding of network models. Doing so would discharge us of the obligation to argue for the causal relevance of intentional symptoms: We could just show that modeling mental disorders as causally related symptoms with intentional content is a useful and predictive explanatory tool. However, we believe instrumentalism is not a serious option here. Adopting instrumentalism about models implies that the models one proposes cannot be said to be either true or false – and this means that what we attempt in the target article would not even be possible: One cannot develop a conditional argument on the truth of a model that cannot be said to be either true or false. More importantly, network models get their "bite" precisely from their ambition to provide an account of what mental disorders *are*; therefore, retreating to an instrumentalist position would render the approach toothless.

The network approach suggests that the intentional content itself figures in causal explanation. For example, **Field et al.** mention that addicts who believe that they suffer from a chronic disease are less likely to recover than those who believe they suffer from an unhealthy habit. Their implicit suggestion is that the specific content of their beliefs about their addiction causally affects their chance of recovery. But how should we understand this? According to **Eronen & Bringmann,** the hard-core reductionist could "accept the importance of intentional contents and their meaningful relationships, but nevertheless argue that the real causal work is done by brain states." To show that this is a dead end, Eronen & Bringmann argue that we might have to resort to an interventionist understanding of causation (a suggestion we already hinted at in the original article), and subsequently propose to start testing the hypothesis that interventions on the level of intentional content indeed have causal effects, a point also made by **Müller**.

While the emphasis on interventions is certainly worthwhile, **Eronen & Bringmann** seem to overlook the fact that we already have a substantial amount of empirical evidence for the thesis that manipulating the content of mental states (fears, beliefs) has content-specific causal effects in psychopathology. In particular, the research program surrounding cognitive behavioral therapy (CBT) offers a rich trove of findings that evaluate exactly this thesis. CBT, which is probably the most extensively researched and empirically supported psychological treatment in existence

(Hofmann et al. 2012), is based on the premise that psychopathology is at least in part caused and maintained by maladaptive cognitions (general schemas about the self and the world): intentional states *par excellence*. CBT explicitly targets the content of these cognitive representations through techniques such as cognitive restructuring (Beck 1970; Ellis 1962).

Panic Disorder (PD) is an important example of a case where interventions on the content of cognitive states are explicitly used. In PD, intentional states play a central role. Primary theories on the etiology of PD hold that panic attacks arise from a feedback loop between physiological arousal (e.g., increased heart rate) and a cognitive schema that specifies a particular intentional content, that is, *about* the arousal (e.g., "My heart is racing, so I must be having a heart attack"). The interpretation of the bodily signal as heralding an impending catastrophe then reinforces the arousal itself, which in turn reinforces the cognitive representation, resulting in a runaway feedback process that culminates in a panic attack (Clark 1986). CBT interventions in use involve a variety of techniques aimed at modifying these cognitive representations, for instance through controlled exposure, which teaches the person that the impending catastrophe in fact need not happen, and by training the person to replace the cognitive representation with another one. (For example, CBT may teach the person to consider at least one other hypothesis, apart from "I am having a heart attack," that could explain the physiological arousal when it arises.) Such interventions have been shown to be effective (e.g., Hofmann et al. 2012; Mitte 2005). Similar techniques are used in a variety of other cases ranging from eating disorders to depression and from somatoform disorders to psychosis (see Hoffmann et al. [2012] for an overview).

Thus, the evidence for the causal relevance of intentional states and the effectiveness of intervening on these states is overwhelming. However, according to **Eronen & Bringmann**, this would still leave open a reductionist rebuttal: The causal relevance of symptoms-with-content could still be brought about by neural or biological causes. Here we can point to the argument developed in **Ross**'s comment, which aims to show that the fact that neural processes are causally involved does not guarantee a successful reduction. What is needed for a reduction to succeed is to show that these neural processes can explain why a person has a belief with a certain content and not with another.

One of the arguments we bring forward for the claim that intentionality blocks reductionism, is that it is widely held that mental states with intentional content are multiply realizable and therefore not type identical with brain states. Some of the commentators argue that this is not convincing. **Pernu**, for example, claims that many apparent cases of multiple realizability can be dissolved by "kind splitting." This solution is based on the idea that one should analyze mental states at a level that is just as "fine-grained" as the level at which one analyzes physical states (Bechtel & Mundale 1999) – and that if one insists on a fine-grained analysis of physical states, one might have to conclude that two people who believe "my neighbor is a secret agent for the CIA" on the basis of different underlying physical processes, do not actually have the same belief at all (Polger 2009).

However, this response makes sense only if one already assumes that mental states such as beliefs and fears are, in fact, completely constituted by physical states in the brain. As said, it might be that in the future this will turn out to be true, but at present this is still very much an open issue. In fact, there is a growing chorus of voices in contemporary philosophy of mind arguing that we should not think about intentional states in this way. According to these approaches, what makes John's belief *about* "his neighbor being a secret agent for the CIA" is not that he is in a certain brain state, but that some coherent set of counterfactuals is true of him (Baker 1995). His belief is characterized through a set of counterfactual conditionals such as: If the neighbor were to start a friendly talk, John would respond nervously; if John passed the CIA headquarters, he would expect his neighbor to be there; and so on. What makes a state a state with a certain specific content, is thus thought to depend on the occurrence of a set of meaningfully related phenomena, embedded in a certain context.

Building on the by now almost mainstream assumption (explicitly or implicitly endorsed by several commentators: **Crafa & Nagel; Field et al.; Jayawickreme et al.**) that cognition is embedded and extended (see Clark & Chalmers 1998; Menary 2010), the content of one's mental state is thought to be at least partly determined by one's environment. This raises another barrier to purely biological explanations of mental disorders, for two reasons: First, the simple but profoundly important and easily forgotten fact that the environment is not in the brain. Second, the more complicated fact that the relevant features of the environment are unlikely to admit an effective characterization in terms of their purely physical features, more or less for the same reason that behaviorist analyses hardly ever managed to characterize classes of stimuli in purely physical terms, without indirect reference to mental states. The physical constitution of the Roulette table certainly matters to its function, but that physical constitution realizes a gambling apparatus only from the point of view of the human being, not from the point of view of physics; it is, in other words, unlikely that the Roulette table will ever become a "kind" of physics (or of neuroscience; Fodor 1974). Perhaps for this reason, the relevant environmental features in psychopathology are themselves almost always multiply realizable. One need only consider the fact that debts and poverty are crucial factors in the maintenance, and probably also in the genesis, of several disorders; these properties come down to a lack of money, and money is the quintessential example of a multiply realizable phenomenon.

Pernu, however, holds that multiple realizability does not offer sufficient grounds to reject reductionism, and that a more successful argument for the failure of reductionist explanation would be if something like *inverse* multiple realizability would be true. So, if it were true that one and the same brain state could realize different symptoms in different circumstances, this would show that even an ideally complete neural description of a symptom would not allow us to identify which symptom it is. Pernu argues that the truth of inverse multiple realizability would undermine even the basic supervenience thesis that he ascribes to us, turning such an argument into what he calls a "Pyrrhic victory." However, in our reference to supervenience in the target article, we merely suggested that "the strongest viable position that is still available would be non-reductive materialism along the lines of the supervenience thesis in the philosophy of mind," that is the thesis defended by Kim (1982; 1984). At the same time, as becomes clear in our argument, any form of non-reductive materialism that our position allows for would clearly need to take a broader spectrum of physical states into account than mere brain states (including, for example, states of the environment). In fact, we consider it highly plausible that something like inverse multiple realizability for brain states is correct, and thus that supervenience about neural states is wrong. As discussed above, **Ross** (who suggests that our tolerance for supervenience is an unwarranted

concession – about which he might be right) convincingly argues that the context is often what makes a disorder a disorder.

This means that it might not always be possible, even in science-fiction scenarios in which one would have access to the complete physical description of brain states, to determine purely on the basis of these brain states whether a person has a certain symptom or not. Indeed, such a conclusion would seem correct for many of the symptoms listed in diagnostic manuals such as DSM-IV. To start with, symptoms that involve truth conditions in the external world would seem to defy supervenience of psychological states on neural states more or less by definition. Elizabeth and Bob may both believe that they are persecuted by the CIA, and this belief may be instantiated in the exact same way in their brains. Depending on the external circumstances, however, this belief may count as a symptom or not – for instance, when the belief is veridical for Elizabeth (who is actually a Russian spy) but finds no grounding in reality for Bob.

A second important class of symptoms that is likely to violate supervenience with respect to neural states involves the many symptoms that code behaviors as "excessive" or "out of proportion" with respect to the circumstances: For example, persistent handwashing is not a symptom of Obsessive Compulsive Disorder (OCD) in a situation where there is a nontrivial risk of infection.

A third class of symptoms that will likely defy supervenience with respect to the brain refers to social norms. Two people may both exhibit the same level of systematic violence, with all of the same neural states that come with it, but if one is a professional boxer while the other is a choir boy, we may justifiably consider the behavior as a sign of psychopathology in the latter but not the former.

Fourth, symptoms that explicitly involve relations with the environment (e.g., having debts, having been in contact with law enforcement, etc.) will not satisfy supervenience with respect to brain states because in these cases what counts is what happened to the person rather than what brain states the person has.

Fifth, a large class of symptoms is defined in terms of a specific trigger in the environment; for instance, the fear response of a phobic who sees a spider may be physiologically indistinguishable of the panic attack in a patient with panic disorder. What makes the difference between these conditions could lie in how they are triggered by the environment.

It should be noted that the above examples putatively refute supervenience of symptoms with respect to brain states, but not with respect to the physical world at large. However, if supervenience of the mental on the physical is meant as supervenience on the physical state of the world *in its totality*, this would engender a form of non-reductive materialism with even fewer teeth than the thesis that mental states supervene on brain states.

So how should we deal with accounting for the intentional content of symptoms? **Slors, Francken, & Strijbos** [**Slors et al.**] suggest that making sense of intentional content requires more than "just" assumptions of rationality. Instead, they argue that what is needed is "a range of further, non-rationalizing interpretive strategies, including simulation, empathizing, and using clinical knowledge and experience." They rightly note that, for example, delusions are precisely delusions insofar as they do *not* manifest the kind of rational relations to other states that we observe in regular beliefs. For example, whereas beliefs are at least to *some* extent susceptible to correcting evidence, delusions are generally not. In the target article, we acknowledge that pathological intentional states differ from non-pathological ones in terms of rationality. We tried to account for the inherent irrationality of psychopathology by pointing out that psychopathology stands out as, so to say, an irrational figure against a "minimally rational" background. Slors et al. turn the picture around, and speak of the search for "islands of reason in a sea of confusion." Which picture best reflects the "rational status" of someone suffering from psychopathological symptoms will clearly depend on the individual case; however, we believe our notion of rationality is much more minimal and bounded than Slors et al. have taken it to be. Our point in speaking of a background of rationality was to indicate that almost all persons suffering from symptoms still take part in the social practice of exchanging reasons: They take part in conversations about treatment options or what they want for dinner, and often even participate in complex practices such as "talking therapies" like cognitive behavioural therapy. It is in this sense that we claim that almost all persons suffering from psychopathology, still manifest a substantial amount of background rationality. So, whereas we fully agree that "making sense" requires a wider interpretative spectrum than just naive assumptions of full rationality, we do not think our arguments offer such a restricted view on interpretation.

R5. Beam me up, Scotty!

As **Hur et al.** state, "Clinical and translational neuroscience has historically been oversold and under-delivered." These are true words, except that overselling neuroscience is not a thing of the past but of the present as well. Even if the prospects for a neuroscientific analysis of mental disorders seem gloomy upon a sober evaluation of the evidence, many researchers keep pointing to the future, presumably under the assumption that, sooner or later, the biological ship must sail in and their utopic Reductionia will materialize.

For example, **Elbau et al.** keep their hopes up and state that "there is not yet reason to abandon the effort of biological reduction that has been most fruitful in all other fields of medicine." **Oller** thinks that "recognition of the abstract nature of mental/linguistic events does not diminish the importance of the neural impulses that, at another level, form the infrastructure *for* those events" (emphasis Oller's). And **Eronen & Bringmann** suggest that the reductionist could maintain that "it is pragmatically useful to describe and predict human behavior in terms of beliefs and desires, but this is consistent with the idea that the real causes of behavior are biological or neural." Perhaps the most adventurous commentators are **Perlovsky**, who envisages a "physics of mind," and **Pernu**, who hopes for an entirely new categorization scheme which, in contrast to the current set of disorders, would map neatly onto as yet unknown neurobiological explanations: "Therefore, the idea is that we should aim to abandon the superficial DSM classifications and replace them with more valid classifications based on physiological aetiologies." As a result, **Pernu** argues, "our current understanding of mental disorders – 'folk psychiatry' – will be fundamentally transformed, and the symptomatically defined notions of mental disorders (M) will give way to new notions, aligned with their neural-level realisers (M1 and M2)."

Of course, one can hope for such transformations to materialize, and one is free to pursue research that might deliver such beautifully delineated categories. We do insist, however, that we all keep our eyes on the ball. In the current scheme of things, explanatory reductionism is a remote possibility, not a realistic research target. We do not have biomarkers that are sufficiently

reliable and predictive for diagnostic use. We have not identified genes that are specific to disorders and explain an appreciable amount of variance. We have not obtained insight into pathogenetic pathways in the brain that are sufficiently secure to inform treatment. If anything, we should wonder why the massive investments in research, that should have uncovered these factors, have not pushed back the prevalence of common mental disorders by a single percentage point. Of course, everybody wants a penicillin for psychopathology, but we do not do anyone a favor by supposing that we are almost there, or even on the way. Therefore, despite the strong scientific image that is built up by brain scanners and genome sequencing machines, the situation as sketched by Insel and Cuthbert (2015), for example, and echoed by **Pernu** should not be understood as science but as science fiction.

If the role of biology is not to teleport us into explanatory reductionism, what exactly is it? **Ward & Fisher** aim in a different and, in our view, more plausible direction, stating that "the primary causes of mental disorders could be biological in nature without being disease entities." Their idea is that, while network models may identify the relevant connections between symptoms, "mechanistic models shed light on the way different symptoms exert their effects and influence each other" — and some of these mechanistic models may well be neurobiological in nature. This is certainly true and can be easily seen from some of the prosaic examples we discussed: Disruptions in the sleep-wake cycle cause concentration problems and fatigue; lack of appetite causes weight loss and lack of energy; prolonged use of drugs often causes tolerance and dependence. There is no doubt whatsoever that these mechanisms are grounded in neurobiology, and that the explanation of why one network structure obtains rather than another will have to rely on such neurobiology as well.

An important question, however, is how far one wants to move down this explanatory ladder and whether it would make equal sense to do so for different symptom-symptom connections. The cases of insomnia → fatigue, lack of appetite → weight loss, and substance use → tolerance are obvious enough, but it is no coincidence that these particular connections do not typically invoke the content of mental states. For connections that *do* rest on the content of mental states, it is much more difficult to see how they are to be explained through neurobiological mechanisms. For instance, John and Jane are both depressed and convinced they are a burden to their environment; more specifically, John thinks he makes his family unhappy, while Jane feels that her friends despise her behind her back. John and Jane both have feelings of worthlessness, but the content of their ideations differs, as will their brain states. How could information on those states possibly help us understand why John and Jane have the feelings and thoughts they have, or in what sense they both instantiate feelings of worthlessness despite the different contents and neural realizations? Similarly, what about the distinct neural mechanisms that must underlie compulsive handwashing (obviously engaging the motor cortex) versus checking locks (probably engaging the visual cortex)? What about gambling addicts who are preoccupied by Roulette tables versus those who cannot help thinking about slot machines? Does anybody seriously expect to be able to identify these phenomena *as being* feelings of worthlessness, compulsive behaviors, or preoccupations with gambling, and to successfully differentiate them from each other and group them in the right categories, merely by looking at the brain?

One can hope for advances in mind-reading neuroscience that would allow one to detect suicidal ideations without having to ask the person what they think. But even if such science-fiction devices were to obtain, it would seem that primacy for identifying the content of mental states will remain with the intentional level of description rather than with the neurological one. One is hard-pressed to accept a situation in which the mind-reading machine detects the presence of persistent suicidal ideation, whereas one subjectively experiences only persistent thinking about how to fix a leaking tap. The success of any mind-reading machine will have to be measured against the ordinary folk psychological level of intentional description (after all, it's supposed to be *reading the mind*). In that sense, the intentional level will always have epistemic priority – at best, a mind-reading machine might become exceptionally well calibrated to that intentional level. As a result, in cases that involve the content of mental states, whatever we expect to find at the level of neurobiology must be *epistemically slaved* to the mental states we already identify at the intentional level. If this is correct, we can never expect the identification of the relevant mental states and behavioral patterns to be successfully executed at the level of neuroscience itself, and this means that whatever neuroscientific explanation we will get is going to be parasitic on "folk psychology" rather than an alternative to it.

R6. A little less observation, a little more experiment

Müller contests the testability of network models, stating that: "Empirical network approaches simply purport testability by bioinformatic large data analysis." Although network models can imply testable constraints on conditional independence relations in the data, we agree that in most applications of statistical network models the results should be taken as exploratory, and that it is important to develop ways of critically testing complexity approaches. The question of how exactly network theories should be empirically interrogated is, however, not trivial; the same holds for the question of how to decide whether a given disorder is amenable to a network theory. As **Tabb** suggests, the assumption that a network model is broadly correct may not hold for all mental disorders (see also Fried & Cramer 2017), and as such, it is important to develop ways of deciding which approach is most likely to be fruitful; for instance, whether a common cause model or a network model would be more appropriate. Simply pitting the statistical models against one another is difficult because that different explanatory models can have very similar consequences at the level of correlations in an observational dataset (**Jayawickreme et al.**). How should we then tackle the challenge of picking the most suitable candidate model for a given disorder?

One potentially fruitful strategy is increased reliance on experimental instead of observational data. In clinical psychology and psychiatry, network analyses have frequently relied on observational data in which no single variable is manipulated, for example, data from major population surveys in which a large number of people are asked once about the presence/absence or severity of psychopathological symptoms (for an overview, see Fried et al. 2017). While useful in many respects (e.g., as a means to obtain prevalence and comorbidity estimates, or to develop hypotheses on network structure), such designs are limited when it comes to understanding the nature of mental disorders. We can fit ever more complex statistical models to these data, but given the data-driven, non-confirmatory nature of current network models, and their possible statistical equivalence with latent variable models, this may not be the most productive way of deciding

between different generative theories of psychopathology. We emphasize that this is so regardless of whether or not one includes physiological and neuroscientific data into the network model, except in the rare cases where one has a very strong argument for the etiological primacy of factors measured at such a level. A biological correlate in principle remains just that: a correlate.

Experiments might, however, offer fruitful avenues for distinguishing between different frameworks. Suppose, for example, that a researcher is able to show in the lab that people, somehow manipulated to have depressed mood, subsequently suffer from more thoughts of death than people whose mood was not manipulated. That is, the researcher has experimentally revealed a connection between two symptoms of major depression. This finding would be consistent with a full network perspective (as advocated in the target article) and with a hybrid perspective in which both a (local) common cause and network dynamics play a role (Fried & Cramer 2017). However, importantly, such findings would be inconsistent with a common cause conceptualization of major depression (and *a fortiori* with the reductionist interpretation of that common cause). In a common cause model, symptoms are merely effects of a latent condition, and manipulation of effects does not change their causes. Just as we cannot manipulate the behavior of one thermometer to induce changes in another – only the common cause *temperature* can change readings on both thermometers – so it does not make sense that symptoms can cause one another if their covariance is truly due to an underlying abnormality – whether it be biological or psychological (**Nephew, Febo, & Santos** [**Nephew et al.**]) – that causes the overt symptomatology of psychopathology. Experimental data in which symptoms are locally manipulated thus should be able to shed light on which disorders are amenable to network theory and which are not. This may also illuminate whether current hierarchical factor models of the covariance between disorders (Kotov et al. 2017), which typically include putative common causes in the form of latent variables, reflect causal order or are merely descriptive, as **Bornstein** suggests.

However, current experimental work – of which the randomized controlled trial (RCT) is its most prominent example – can (and perhaps should) be modified in the coming years to better incorporate a non-reductionist, multifactorial perspective on mental disorders. We agree with **Ioannidis** that, for example, RCTs could (and perhaps should) be redesigned so that they better capture non-biological outcomes over a longer period of time. We add to these helpful suggestions the idea that RCTs may test the efficacy of an intervention that targets either a symptom (e.g., targeting the most central node *insomnia* with a sleep hygiene protocol) or a network connection (e.g., cognitive restructuring to weaken the connection between the symptoms *depressed mood* and *suicidal thoughts*); as a result of which one may be able to deduce whether network approaches, as a possible intervention strategy, are fruitful or not.

Including time series data on dynamics as dependent variables in the experimental design (Snippe et al. 2017; Wichers et al. 2016) and/or modeling experimental manipulations as treatment factors in a statistical network models (Bekhuis et al. 2018; Blanken et al. 2019) could be two important starting points of such analyses. Another possible starting point would be to develop dynamical network models that could be used to model treatment effects through surgical (or not-so-surgical, i.e. "fat hand") interventions in the model structure, perhaps analogous to how this is done in modern approaches in causal inference (Pearl 2000). This would involve extensions of the current causal modeling apparatus (most importantly the inclusion of nonlinearity and feedback), but there seems to be no principled reason why suitably adapted interventionist models of causation (Woodward 2003) could not be employed in, say, a dynamical systems framework. The often-used Ising model (van Borkulo et al. 2014) is a simple toy example that would seem to support models for causal interventions (Marsman et al. 2017); for example, if the Ising model is true, then intervening on one of its nodes may be modeled using the conditioning operation, which induces a new Ising model (Epskamp et al. 2018, Eq. 30.6). However, many other approaches are conceivable, too. If we were able to simulate dynamical networks well enough to deliver reasonable predictions on what should happen under various interventions to the model, that would be a significant advance. Clearly, current models and theories cannot handle even such comparably easy questions, as **Müller** correctly suggests, and there is accordingly a huge space of opportunity for the development of network theories and analysis techniques that optimally model experimental interventions, in addition to the extensions of RCTs suggested by **Ioannidis**.

R7. A system of … what, actually?

Many commentators **(Crafa & Nagel**; **Hens et al.**; **Jayawickreme et al.**; **Redish et al.)** noted, correctly, that a systems perspective on mental disorders necessitates having a clear sense of what the elements of such a system are. In this respect, it is obvious that network researchers have not yet developed systematic approaches to assess this issue. In our target article we have focused on (DSM) symptoms as the key nodes of a psychopathological system (see also Borsboom 2017), but it should be an important part of future research endeavors to investigate the extent to which this is the most optimal characterization. This matters, as Redish et al. point out, because one's conclusion based on network analysis (or any statistical analysis, for that matter) – for example, the most central node is X – heavily relies on the variables that feature as nodes in the network structure. When using DSM depression symptoms, the most central node could be *depressed mood* while, when using RDoC factors, the most central node might be *perception and understanding of the self*.

The notion of a "symptom" itself is also problematic. To many, it suggests the idea that a symptom is an indication of something else – a disease/disorder – which exists independently of the symptom. This may not be the right way to think about symptoms in the context of network models, however. In network theory, the phenomena that we call symptoms can also be guides to diagnosis, but they do not signify the presence of a distinct entity, as is often the case in medicine; rather, psychopathology symptoms should be seen as signalling the disturbance of the network as a whole, just like abnormal amounts of algae signal the alternative stable state of a turbid lake (Scheffer et al. 2001). One could wonder, however, whether symptoms (traditionally evaluated for their epistemic quality as indicators of an underlying abnormality) are the only and/or most crucial elements of a psychopathological system from a causal point of view – that is, whether they are the crucial drivers of network dynamics that result in a disordered state (Fried & Cramer 2017; Jones et al. 2017). Even without a network approach in mind, there is reason to be critical regarding the capacity of symptoms to capture the essential features of a disorder. For example, **Hens et al.** convincingly argue that ASD symptoms do not capture "what it actually means to have such a disorder." In addition, as **Hyland** argues, the covariation of somatic symptoms in the case of functional disorders "cannot

be explained in terms of symptom-to-symptom causality." So, what other variables may be relevant for psychopathology networks? Many candidate factors could (and should) be considered, primarily within the context of explicit theoretical models (i.e., formalized dynamical systems); **Crafa & Nagel** even advocate including social flexibility in psychopathology networks, as it may underpin the degree to which a person is susceptible to intentionality itself – that is, the degree to which "exposure to new social information will reinforce or undermine existing processes." Jones et al. (2017) similarly provide a motivation for a number of cases where interplay between symptoms may not suffice to characterize disorders.

Potentially moving away from symptoms means relying less on existing data and collecting more new data with the specific purpose of capturing networks and their dynamics. This is something **Jayawickreme et al.** also explicitly advocate, but for another reason: Many of the items that feature in existing datasets were specifically developed with the intent of accurately measuring one and the same latent variable. This usually means that such items tap into similar aspects of the construct targeted. Take, for example, the items "I like a desk without clutter" and "I sort my socks by color." Such items are, by necessity and design, highly correlated; and in a network analysis, this will show up as a relatively strong connection. Importantly, such a strong connection may not stem from an actual direct relation – for example, liking a desk without clutter causing someone to sort their socks by color – but, rather, from the fact that these two items really *do* measure the same thing, namely liking order (or not). This will prove an additional challenge in future research: telling network fact (actual relations) from fiction (spurious correlations). Clearly, if many variables that overlap too strongly in their semantics are included in a network structure, this may yield inadequate solutions (Costantini 2014).

Successful modeling of complex systems therefore requires a judicious choice of the key variables in the system. This has proven true for dynamical systems modeled in fields ranging from ecosystems to meteorology, and is undoubtedly true for psychopathology as well. As more theoretically informed network theories are developed, we hope it will become clear which variables are essential and how the dynamic interaction between them unfolds in time. Current symptomatology is likely to include some key variables already, but is unlikely to contain a definitive list; hence, we expect considerable progress to unfold in this respect in both substantive and methodological directions.

R8. Mereological matryoshkas? Integrating multiple levels of analysis

As we have stressed throughout our target paper, blocking explanatory reductionism by means of a network perspective is not analogous to blocking biology-oriented research altogether. That is, although searching for a biological common cause (e.g., etiological brain pathways as mentioned by **Ioannidis**) is not fruitful if a mental disorder is the outcome of symptom-symptom network dynamics, searching for biological processes that are implicated in symptoms (e.g., sclerotic plaque as the process that gives rise to the symptom *chest pains*; **Elbau et al.**) and connections between them (e.g., homeostatic mechanisms that give rise to interindividual differences in connectivity between insomnia and fatigue) is by no means a pointless exercise. Symptoms in particular that either have what **McNally** refers to as "formal" features (e.g., auditory hallucinations), or that "appear to lack intentional content altogether" (e.g., exaggerated startle and emotional numbing), are suitable candidates for research programs that aim at biological elucidation. That is, there are interesting biological/psychological correlates to be found that may partially account for the non-intentional component of certain symptoms and connections between them. The main difference with existing research programs from a network perspective is that these correlates are not the holy grail, as in **Tabb's** "triumphalist reductionism," but instead would be modeled and analyzed as making up or informing the structure and dynamics of the complex system that drives psychopathology.

A challenge for the coming years is to develop sensible ways in which to integrate various levels of network analysis – for example, combining biological with more psychological variables (as suggested by **Baran, Hyland, Oller, Pasqualotto,** and **Pessoa**). Addressing this challenge properly will be anything but simple. The most straightforward solution – that is, just estimating a network structure for all of these biological and psychological variables simultaneously – is suboptimal for various reasons. For instance, it is well known that correlations between data coming from different sources will typically be low. As a result, using current state-of-the-art network estimation methods that use regularization in order to avoid false positives will result in very small or even absent conditional dependence relations – while, in fact, such a relation between a biological and psychological variable may be important. One way to deal with this problem is to find statistical solutions (e.g., by locally relaxing penalty parameters of the statistical regularization procedure), but most of these solutions naturally come at the expense of increasing the probability of finding spurious relations. A similar way of including biological variables in a symptom network is by treating the biological variables as moderator variables that determine the strengths of connections between symptoms.

It is certainly important to develop and try out such techniques, but it should be realized that the above strategies analyze distinct variables (e.g., psychological and biological ones) as if they are functionally distinct entities, which may not be appropriate. Another approach would be to assume that the "biological" and the "psychological" represent different network structures, as **Hyland** argues for example, particularly in the case of functional disorders such as fibromyalgia. Such a solution, while highly interesting, is not trivial. Two major questions are how these network structures relate to one another at a theoretical level, and which statistical model, if any, best captures this relation. Under the assumption that psychological variables are not simply higher-level realizations of lower-level biological processes, as we have argued in our target paper, relating "biological" to "psychological" network structures may operate in at least two ways. First, they may be related through a *mereological* structure, where the biological parts do not cause a particular psychological variable but rather form that variable, just as parts of a tangerine form a tangerine yet do not cause it. In this case, subsets of the biological variables could be modeled psychometrically as formative indicators of psychological variables (Kievit et al. 2011). Second, psychological and biological networks may be related through a *Russian doll* structure, where biological network structures are "nested" in a psychological network without one causing the other, just as one smaller matryoshka does not cause a larger matryoshka. One should, in this case, find a sensible way of relating the state and architecture of elements in the psychological network structure to those of the biological network structures. This perhaps could be done by making the value of each psychological

variable a function of the state of the embedded biological network, while the state of the embedded biological network is a function of the relations in the symptom network; however, as far as we know, there is currently no modeling framework to encode this idea.

Interestingly, as soon as one moves from verbal descriptions of relations between "levels of analysis" to formalized models of these relations, it becomes unclear exactly how the formalization should be done (apart from the most simple models; see also Kievit et al. 2011). Perhaps it would be useful to construct a "simplest non-trivial case" in the form of a simulation model that explicitly codes relations between psychological and biological networks. Such a model could also be used to focus the debate. In any case, it is clear that the question of how to integrate biological and psychological levels of analysis is wide open in psychopathology research, and that there are considerable opportunities for progress in addressing this issue.

R9. Conclusion

We believe our target article, together with the commentaries, establishes three clear conclusions: (1) Mental disorders are not brain disorders in the everyday understanding of these terms; (2) Explanatory reductionism is both unlikely to be correct and insufficiently popular to engender considerable support among commentators when challenged; and (3) If network theory is broadly correct, reductionism is in an awfully structural sort of trouble. However, it also became clear that simply throwing a bunch of symptoms into a statistical analysis will not by itself answer the question of how psychopathology arises and what it is, especially in relation to the complex configuration of biological, psychological, and social levels of description that will enter into such networks. If a complex network perspective is more broadly adopted in the field, it will need to address the questions of what the constituent components of networks are, how they play out dynamically in time, and what the role of biology is in such a network. If current network models define the minimum level of complexity needed to properly characterize psychopathology, as we indeed believe they do, that blocks the road to reductionism; however, at the same time, the methodological and substantive challenges to a successful analysis of disorders are sufficiently intimidating to motivate scientific modesty of network theorists as well. Network approaches offer tantalizing possibilities for integrating different levels of analysis into a comprehensive system, but much more work is needed before such prospects can be realized.

Notes

1. https://qbi.uq.edu.au/brain/brain-diseases/depression.
2. https://www.volksgezondheidenzorg.info/onderwerp/hersenaandoeningen/overzicht.

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Abed R. T. & Abbas M. J. (2011) A reformulation of the social brain theory for schizophrenia: The case for out-group intolerance. *Perspectives in Biology and Medicine* **54**:132–51. [OJ]

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**:1361–74. [aDB]

Adam D. (2013) Mental health: On the spectrum. [Editorial]. *Nature* **496**:416–18. [aDB, LPer]

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

Alves P., Sales C. & Ashworth M. (2017) Does outcome measurement of treatment for substance use disorder reflect the personal concerns of patients? A scoping review of measures recommended in Europe. *Drug and Alcohol Dependence* **179**:299–308. [JPAI]

Ambadyp N. & Bharucha J. (2009) Culture and the brain. *Current Directions in Psychological Science* **18**:342–45. [DC]

American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders*, 5th edition (DSM-5). American Psychiatric Association. [aDB, RJM, ADR, LT]

Anderson M. L. (2010) Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences* **33**:245–66. [AP, TKP]

Anderson M. L., Kinnison J. & Pessoa L. (2013) Describing functional diversity of brain regions and brain networks. *Neuroimage* **73**:50–58. [LPes]

Andreasen N. C. (1984) *The broken brain: The biological revolution in psychiatry*. Harper and Row. [aDB]

Andrews K. (2012) *Do apes read minds? Towards a new folk-psychology*. MIT Press. [MS]

Anonymous. (2013) No dishonour in depression. [Editorial]. *Nature* 498: article 137. (Posted on 12 June, 2013). Available at: http://www.nature.com/news/no-dishonour-in-depression-1.13170. [aDB]

Archpru Akaka M. & Chandler J. D. (2011) Roles as resources: A social roles perspective of change in value networks. *Marketing Theory* **11**(3):243–60. [DC]

Badiani A. (2013) Substance-specific environmental influences on drug use and drug preference in animals and humans. *Current Opinion in Neurobiology* **23**:588–96. [CPM]

Bak P. (1997) *How nature works*. Oxford University Press. [LT]

Baker L. R. (1995) *Explaining attitudes: A practical approach to the mind*. Cambridge University Press. [arDB]

Bakermans-Kranenburg M. J. & Van Ijzendoorn M. H. (2006) Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Developmental Psychobiology* **48**:406–409. [OJ]

Barabási A.-L. (2012) The network takeover. *Nature Physics* **8**:14–16. doi:10.1038/nphys2188. [aDB, RJM]

Barabási A.-L., Gulbahce N. & Loscalzo J. (2011) Network medicine: A network-based approach to human disease. *Nature Reviews Genetics* **12**:56–68. [RJM]

Barabási A.-L. & Oltvai Z. N. (2004) Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics* **5**:101–14. [RJM]

Baran N. M. (2017) Sensitive periods, vasotocin-family peptides, and the evolution and development of social behavior. *Frontiers in Endocrinology* 8: article no. 00189. (Online journal). Available at: https://doi.org/10.3389/fendo.2017.00189. [NMB]

Barbara J.-G. (2008) L'adaptation biologique et les neurosciences [Biological adaptation and neuroscience]. Paper presented at the REHSEIS Conference on Adaptations, UPMC, Université Diderot, REHSEIS, CNRS UMR7596, Paris, October 21, 2008 (journée «Adaptations», journée de rentrée du REHSEIS, organisée par J. G. Barbara, C. Lefève, G. Gachelin, 21 octobre 2008, Université Diderot, CNRS UMR7596). Available at: http://www.rehseis.cnrs.fr/spip.php?article313; and at : http://www.biusante.parisdescartes.fr/chn/docpdf/adaptation.pdf. [MD]

Bassett D. S. & Sporns O. (2017) Network neuroscience. *Nature Neuroscience* **20**:353–64. [RJM]

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* **46**:3359–69. [aDB]

Bechtel W. & Mundale J. (1999) Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science* **66**:175–207. [arDB, TKP]

Beck A. T. (1970) Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy* **1**:184–200. [rDB]

Beck A. T. (2008) The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry* **165**:969–77. [aDB]

Beck A. T. & Bredemeier K. (2016) A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical Psychological Science* **4**(4):596–619. Available at: https://doi.org/10.1177/2167702616628523. [TW]

Beer A. L., Plank T. & Greenlee M. W. (2011) Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research* **213**:299–308. [AP]

Bekhuis E., Schoevers R., de Boer M., Peen J., Dekker J., Van H. & Boschloo L. (2018) Symptom-specific effects of psychotherapy versus combined therapy in the treatment of mild to moderate depression: A network approach. *Psychotherapy and Psychosomatics* **87**:121–23. [rDB]

Belsky J. (2016) The differential susceptibility hypothesis: Sensitivity to the environment for better and for worse. *JAMA Pediatrics* **170**(4):321–22. [AT]

Bem S. & De Jong H. L. (2013) *Theoretical issues in psychology: An introduction*, 3rd edition. Sage. [MEH]

Bentall R. (2003) *Madness explained: Psychosis and human nature*. Penguin. [aDB]
Berkman E. T., Hutcherson C. A., Livingston J. L., Kahn L. E. & Inzlicht M. (2017) Self-control as value-based choice. *Current Directions in Psychological Science* **26** (5):422–28. [MF]
Bewernick B. H., Kayser S., Sturm V. & Schlaepfer T. E. (2012) Long-term effects of nucleus accumbens deep brain stimulation in treatment-resistant depression: Evidence for sustained efficacy. *Neuropsychopharmacology* **37**:1975–85. [JH]
Bickle J. (1998) *Psychoneural reduction: The new wave*. MIT Press. [aDB, TKP]
Bickle J. (2003) *Philosophy and neuroscience: A ruthlessly reductive account*. Kluwer. [TKP]
Bickle J. (2016) Multiple realizability. In: *The Stanford Encyclopedia of Philosophy* (Spring 2016 edition), ed. E. N. Zalta. Stanford University. Available at: https://plato.stanford.edu/archives/spr2016/entries/multiple-realizability/. [MIE]
Birn R. M., Shackman A. J., Oler J. A., Williams L. E., McFarlin D. R., Rogers G. M., Shelton S. E., Alexander A. L., Pine D. S., Slattery M. J., Davidson R. J., Fox A. S. & Kalin N. H. (2014) Evolutionarily conserved dysfunction of prefrontal-amygdalar connectivity in early-life anxiety. *Molecular Psychiatry* **19**:915–22. [JH]
Birnbaum R. & Weinberger D. R. (2017) Genetic insights into the neurodevelopmental origins of schizophrenia. *Nature Reviews Neuroscience* **18**:727–40. [JH]
Blanken T. F., Van Der Zweerde T., Van Straten A., Van Someren E. J. W., Borsboom D. & Lancee J. (2019) Introducing Network Intervention Analysis to investigate sequential, symptom-specific treatment effects: A demonstration in co-occurring insomnia and depression. *Psychotherapy and Psychosomatics*, e-pub ahead of print. doi: 10.1159/00049504. [rDB]
Bliksted V., Fagerlund B., Weed E., Frith C. & Videbech P. (2014) Social cognition and neurocognitive deficits in first-episode schizophrenia. *Schizophrenia Research* **153**: 9–17. [DC]
Bornstein R. F. (2011) An interactionist perspective on interpersonal dependency. *Current Directions in Psychological Science* **20**:124–28. [RFB]
Bornstein R. F. & Becker-Matero N. (2011) Reconnecting psychoanalysis to mainstream psychology: Metaphor as glue. *Psychoanalytic Inquiry* **31**:172–84. [RFB]
Borsboom D. (2008) Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology* **9**:1089–1108. [aDB, DC]
Borsboom D. (2017) A network theory of mental disorders. *World Psychiatry* **16**(1):5–13. doi:10.1002/wps.20375. [arDB, RJM, BCN]
Borsboom D. & Cramer A. O. J. (2013) Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology* **9**:91–121. [aDB, NJ, RJM]
Borsboom D., Cramer A. O., Schmittmann V. D., Epskamp S. & Waldorp L. J. (2011) The small world of psychopathology. *PLoS One* **6**(11):e27407. Available at: https://doi.org/10.1371/journal.pone.0027407. [ADR]
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**:610–14. [aDB]
Borsboom D., Fried E. I., Epskamp S., Waldorp L. J., van Borkulo C. D., van der Maas H. L. J. & Cramer A. O. J. (2017) False alarm? A comprehensive reanalysis of "Evidence that psychopathology symptom networks have limited replicability" by Forbes, Wright, Markon & Krueger (2017). *Journal of Abnormal Psychology* **126**:989–99. [RJM]
Bortolotti L. (2010) *Delusions and other irrational beliefs*. Oxford University Press. [aDB]
Boschloo L., van Borkulo C. D., Rhemtulla M., Keyes K. M., Borsboom D. & Schoevers R. A. (2015) The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PLoS ONE* **10**:e0137621. doi:10.1371/journal.-pone.0137621. [arDB]
Bowie C. R., Leung W. W., Reichenberg A., McClure M. M., Patterson T. L., Heaton R. K. & Harvey P. D. (2008) Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biological Psychiatry* **63**:505–11. [DC]
Boyd R. (1991) Realism, antifoundationalism and the enthusiasm for natural kinds. *Philosophical Studies* **61**:127–48. [aDB]
Boyd R. (1999) Homeostasis, species, and higher taxa. In: *Species: New interdisciplinary essays*, ed. R. A. Wilson, pp. 141–85. MIT Press. [aDB]
Braid F., Baiardini I., Molinengo G., Garuti S., Ferrari M., Mantero M., Blasi F. & Canonica G. W. (2016) Choose your outcomes: From the mean to the personalized assessment of outcomes in COPD. An exploratory pragmatic survey. *European Journal of Internal Medicine* **34**:85–88. [JPAI]
Brentano F. (1874) *Psychology from an empirical standpoint*. Routledge & Kegan Paul. [aDB]
Bressler S. L. & Menon V. (2010) Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences* **14**:277–90. [AP]
Bringmann L. F. & Eronen M. I. (2018) Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review* **125**(4):606–15. doi: 10.1037/rev0000108. [MIE, RJM, HMvL]
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. [aDB]
Brüne M. (2014) On aims and methods of psychiatry: A reminiscence of 50 years of Tinbergen's famous questions about the biology of behavior. *BMC Psychiatry* 14: article 1695. (Online article). doi: 10.1186/s12888-014-0364-y. Available at: https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-014-0364-y. [OJ]
Brüne M. (2016) *Textbook of evolutionary psychiatry and psychosomatic medicine: The origins of psychopathology*, 2nd edition. Oxford University Press. [OJ]
Brüne M. (2016) Borderline personality disorder: Why "fast and furious"?. *Evolution Medicine and Public Health* **1**:52–66. [AT]
Bulteel K., Tuerlinckx F., Brose A. & Ceulemans E. (2016) Using raw VAR regression coefficients to build networks can be misleading. *Multivariate Behavioral Research* **51**(2–3):330–44. [MIE]
Cacioppo J. T. & Tassinary L. G. (1990) Inferring psychological significance from physiological signals. *American Psychologist* **45**(1):16–28. [LPes]
Campbell J. (2009) What does rationality have to do with psychological causation? Propositional attitudes as mechanisms and as control variables. In: *Psychiatry as cognitive neuroscience: Philosophical perspectives*, ed. M. Broome & L. Bortoletti, pp. 137–49. Oxford University Press. [aDB]
Carroll S. B. (2005) *Endless forms most beautiful: The new science of evo devo and the making of the animal kingdom*. W. W. Norton. [DKO]
Case A. & Deaton A. (2015) Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences USA* **112**(49):15078–83. [MF]
Caspi A., Hariri A. R., Holmes A., Uher R. & Moffitt T. E. (2010) Genetic sensitivity to the environment: The case of the serotonin transporter gene and its implications for studying complex diseases and traits. *FOCUS* **8**(3):398–416. Available at: https://doi.org/10.1176/foc.8.3.foc398. [NMB]
Caspi A., Houts R. M., Belsky D. W., Goldman-Mellor S. J., Harrington H., Israel S., Meier M. H., Ramrakha S., Shalev I., Poulton R. & Moffitt T. E. (2014) The p factor: One general psychopathology factor in the structure in the structure of psychiatric disorders? *Clinical Psychological Science* **2**:119–37. [RJM]
Caspi A. & Moffitt T. E. (2006) Gene-environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience* **7**:583–90. [CPM]
Caspi A., Sugden K., Moffitt T. E., Taylor A., Craig I. W., Harrington H., McClay J., Mill J., Martin J., Braithwaite A. & Poulton R. (2003) Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science* **301**:386–89. [RJM]
Chamberlain S. R. & Menzies L. (2009) Endophenotypes of obsessive–compulsive disorder: Rationale, evidence and future potential. *Expert Review of Neurotherapeutics* **9**:1133–46. [DC]
Chandler D. J. (2013) Something's got to give: Psychiatric disease on the rise and novel drug development on the decline. *Drug Discovery Today* **18**:202–206. [JPAI]
Chhim S. (2012) Baksbat (broken courage): The development and validation of the inventory to measure baksbat, a Cambodian trauma-based cultural syndrome of distress. *Culture, Medicine and Psychiatry* **36**:640–59. [NJ]
Chiao J. Y. (2009) Cultural neuroscience: A once and future discipline. *Progress in Brain Research* **178**:287–304. [AP]
Chua K. J., Lukaszewski A. W., Grant D. M. & Sng O. (2017) Human life history strategies. *Evolutionary Psychology* **15**(1):1474704916677342. doi: 10.1177/1474704916677342 (Online article). Available at: http://journals.sagepub.com/doi/10.1177/1474704916677342. [AT]
Cipriani A., Furukawa T. A., Salanti G., Chaimani A., Atkinson L. Z., Ogawa Y., Leucht S., Ruhe H. G., Turner E. H., Higgins J. P. T., Egger M., Takeshima N., Hayasaka Y., Imai H., Shinohara K., Tajika A., Ioannidis J. P. A. & Geddes J. R. (2018) Comparative efficacy and acceptability of first- and second-generation antidepressants in the acute treatment of major depressive disorder: A network meta-analysis. *The Lancet* **391**:1357–66. [JPAI]
Clark A. & Chalmers D. (1998) The extended mind. *Analysis* **58**:7–19. [arDB]
Clark D. M. (1986) A cognitive approach to panic. *Behavior Research and Therapy* **24**:461–70. [rDB]
Clementz B. A., Sweeney J. A., Hamm J. P., Ivleva E. I., Ethridge L. E., Pearlson G. D., Keshavan M. S. & Tamminga C. A. (2016) Identification of distinct psychosis biotypes using brain–based biomarkers. *American Journal of Psychiatry* **173**(4):373–84. doi:10.1176/appi.ajp.2015.14091200. [DKO]
Collignon O., Dormal G., Albouy G., Vandewalle G., Voss P., Phillips C. & Lepore F. (2013) Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain* **136**(Part 9):2769–83. [MD]
Convento S., Vallar G., Galantini C. & Bolognini N. (2013) Neuromodulation of early multisensory interactions in the visual cortex. *Journal of Cognitive Neuroscience* **25**:685–96. [AP]
Conway C. C., Forbes M. K., Forbush K. T., Fried E. I., Hallquist M. N., Kotov R., Mullins-Sweatt S. N., Shackman A. J., Skodol A. E., South S. C., Sunderland M., Waszczuk M. A., Zald D. H., Afzali M. H., Bornovalova M. A., Carragher N., Docherty A. R., Jonas K. G., Krueger R. F., Patalay P., Pincus A. L., Tackett J. L., Reininghaus U., Waldman I. D., Wright A. G. C., Zimmerman J., Bach B., Bagby R. M., Chmielewski M., Cicero D. C., Clark L. A., Dalgleish T., DeYoung C. G., Hopwood C. J., Ivanova M. Y., Latzman R. D., Patrick C. J., Ruggero C. J.,

Samuel D. B., Watson D. & Eaton N. R. (2018) A hierarchical taxonomy of psychopathology can reform mental health research. *Perspectives on Psychological Science*. [Also in: *PsyArXiv Preprints*, pp. 1–43. Available at: https://psyarxiv.com/wsygp/]. [JH]

Costantini G. (2014) Network analysis: A new perspective on personality psychology. Unpublished doctoral dissertation, Department of Psychology, University of Milano-Bicocca. [rDB]

Crafa D. & Nagel S. (2013) Accounting for heterogeneity: The culture-brain-behavior interaction model. Poster presented at the Meeting of the International Cultural Neuroscience Consortium, Evanston, Illinois, June 2013. [DC]

Crafa D. & Nagel S. K. (in press) Traces of culture: The feedback loop between brain, behavior, and disorder. *Transcultural Psychiatry*. [DC]

Cramer A. O. J., Borsboom D., Aggen S. H. & Kendler K. S. (2012) The pathoplasticity of dysphoric episodes: Differential impact of stressful life events on the pattern of depressive symptom intercorrelations. *Psychological Medicine* **42**:957–65. [RJM]

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**:270–71. [aDB]

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. doi:10.1371/journal.pone.0167490. [arDB]

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**:137–50; discussion 150–93. [aDB, RJM]

Craver C. & Kaplan D. M. (in press) Are more details better? On the norms of completeness for mechanistic explanations. *Philosophy of Science*. Available at: https://doi.org/10.1093/bjps/axy015. [TW]

Cromer K. D. & Villodas M. T. (2017) The role of posttraumatic stress as a pathway to psychopathology among youth at high risk for victimization. *Psychology of Violence* **7**:12–21. [RFB]

Cuthbert B. N. (2014) The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* **13**:28–35. [TKP]

Cuthbert B. N. & Kozak M. J. (2013) Constructing constructs for psychopathology: The NIMH Research Domain Criteria. *Journal of Abnormal Psychology* **122**:928–37. [TKP]

Dalmau J., Tüzün E., Wu H., Masjuan J., Rossi J. E., Voloschin A., Baehring J. M., Shimazaki H., Koide R., King D., Mason W., Sansing L. H., Dichter M. A., Rosenfeld M. R. & Lynch D. R. (2007) Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with ovarian teratoma. *Annals of Neurology* **61**(1):25–36. Available at: https://doi.org/10.1002/ana.21050. [HMvL]

Davidson D. (1984) *Inquiries into truth and interpretation*. Clarendon Press. [aDB, LT]

Davies J. (2018) Addiction is not a brain disease. (Editorial). *Addiction Research and Theory* **26**(1):1–2. [MF]

Deary I. J. (2012) Intelligence. *Annual Review of Psychology* **63**:453–82. [RJM]

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

Dehaene S. (2007) *Les neurones de la lecture*. Odile Jacob. [MD]

De Jaegher H. (2013) Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience* 7: article 15. Available at: https://doi.org/10.3389/fnint.2013.00015, and at: https://www.frontiersin.org/articles/10.3389/fnint.2013.00015/full. [KH]

Del-Ben C. M., Ferreira C. A., Sanchez T. A., Alves-Neto W. C., Guapo V. G., de Araujo D. B. & Graeff F. G. (2012) Effects of diazepam on BOLD activation during the processing of aversive faces. *Journal of Psychopharmacology* **26**:443–51. [JH]

Del Giudice M. (2014) An evolutionary life history framework for psychopathology. *Psychological Inquiry* **25**:261–300. [OJ]

Del Giudice M., Ellis B. J. & Shirtcliff E. A. (2011) The adaptive calibration model of stress responsivity. *Neuroscience and Biobehavioral Review* **35**(7):1562–92. [AT]

De Meo R., Murray M. M., Clarke S. & Matusz P. J. (2015) Top-down control and early multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience* 9: article 17. doi: 10.3389/fnint.2015.00017. Available at: https://www.frontiersin.org/articles/10.3389/fnint.2015.00017/full. [AP]

Dennett D. C. (1987) *The intentional stance*. MIT Press. [aDB, MIE, MS, LT]

Dennett D. C. (1991) Real patterns. *The Journal of Philosophy* **88**:27–51. [aDB, DKO]

Dennett D. C. (1995) *Darwin's dangerous idea: Evolution and the meanings of life*. Penguin. [MF]

Dennett D. C. (1997) *Kinds of minds. Towards an understanding of consciousness*. Basic Books. [aDB]

de Ricqlès A. (2015) Adaptation: Adaptation biologique. In: *Dans Encyclopedia Universalis*. (Online publication). Available at: http://www.universalis.fr/encyclopedie/adaptation-adaptation-biologique/. [MD]

De Schryver M., Vindevogel S., Rasmussen A. E. & Cramer A. O. J. (2015) Unpacking constructs: A network approach for studying war exposure, daily stressors and post-traumatic stress disorder. *Frontiers in Psychology* 6: article 1896. (Online publication). doi:10.3389/fpsyg.2015.01896. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01896/full. [NJ]

Deserno M. K., Borsboom D., Begeer S. & Geurts H. M. (2017) Multicausal systems ask for multicausal approaches: A network perspective on subjective well-being in individuals with autism spectrum disorder. *Autism* **21**:960–71. [rDB]

Deserno M. K., Borsboom D., Begeer S. & Geurts H. M. (2018) Relating ASD symptoms to well-being: Moving across different construct levels. *Psychological Medicine* **48**:1179–89. [rDB]

Desseilles M. & Phillips C. (2016) How cognition affects perception: Brain activity modelling to unravel top-down dynamics. *Behavioral and Brain Sciences* **39**:e238. doi: 10.1017/S0140525X15002757. [MD]

Di Chiara G. & Imperato A. (1988) Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences USA* **85**:5274–78. [CPM]

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**:309–18. [aDB]

Dretske F. I. (1997) *Naturalizing the mind*. MIT Press. [aDB]

Drysdale A. T., Grosenick L., Downar J., Dunlop K., Mansouri F., Meng Y., Fetcho R. N., Zebley B., Oathes D. J., Etkin A., Schatzberg A. F., Sudheimer K., Keller J., Mayberg H. S., Gunning F. M., Alexopoulos G. S., Fox M. D., Pascual-Leone A., Voss H. U., Casey B. J., Dubin M. J. & Liston C. (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine* **23**(1):28–38. doi:10.1038/nm.4246. Available at: https://www.nature.com/articles/nm.4246. [JH, BCN]

Durisko Z., Mulsant B. H., McKenzie K. & Andrews P. W. (2016) Using evolutionary theory to guide mental health research. *The Canadian Journal of Psychiatry* **61**:159–65. [OJ]

Easton A. C., Lucchesi W., Lourdusamy A., Lenz B., Solati J., Golub Y., Lewczuk P., Fernandes C., Desrivieres S., Dawirs R. R., Moll G. H., Kornhuber J., Frank J., Hoffmann P., Soyka M., Kiefer F., Schumann G., Giese K. P. & Müller C. P. (2013) AlphaCaMKII autophosphorylation controls the establishment of alcohol drinking behavior. *Neuropsychopharmacology* **38**:1636–47. [CPM]

Edelman G. M. & Gally J. A. (2001) Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences USA* **98**(24):13763–68. [LPes]

Eiser J. R. & Van der Pligt J. (1986) "Sick" or "hooked": Smokers' perceptions of their addiction. *Addictive Behaviors* **11**(1):11–15. [MF]

Eiser J. R., van der Pligt J., Raw M. & Sutton S. R. (1985) Trying to stop smoking: Effects of perceived addiction, attributions for failure, and expectancy of success. *Journal of Behavioral Medicine* **8**(4):321–41. [MF]

Eliasmith C. (2013) *How to build a brain: A neural architecture for biological cognition*. Oxford University Press. [TW]

Ellis A. (1962) *Reason and emotion in psychotherapy*. Lyle Stuart. [rDB]

Elton M. (2003) *Daniel Dennett. Reconciling science and our self-conception*. Polity Press. [aDB]

Endicott R. P. (1993) Species-specific properties and more narrow reductive strategies. *Erkenntnis* **38**(3):303–21. [aDB]

Epskamp S. & Fried E. I. (2018) A tutorial on regularized partial correlation networks. *Psychological Methods*. doi:10.1037/met0000167. (Advance Online publication. Available at: http://dx.doi.org/10.1037/met0000167). [rDB]

Epskamp S., Maris G., Waldorp L. J. & Borsboom D. (2018) Network psychometrics. In: *Handbook of Psychometrics*, ed. P. Irwing, D. Hughes & T. Booth, pp. 953–87. Wiley. [arDB]

Eronen M. I. (2013) No levels, no problems: Downward causation in neuroscience. *Philosophy of Science* **80**:1042–52. [aDB]

Eronen M. I. (2017) Interventionism for the intentional stance: True believers and their brains. *Topoi*. (Published Online, 2nd December, 2017). Available at: https://doi.org/10.1007/s11245-017-9513-5. [aDB, MIE, MS]

Evangelou E., Warren H. R., Mosen-Ansorena D., Mifsud B., Pazoki R., Gao H., … Caulfield M. J. & the Million Veteran Program (2018) Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature Genetics* **50**:1412–25. doi:10.1038/s41588-018-0205-x. [JH]

Fair D. A., Cohen A. L., Power J. D., Dosenbach N. U. F., Church J. A., Miezin F. M., Schlaggar B. L. & Petersen S. E. (2009) Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology* **5**:e1000381. (Online publication). Available at: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000381. [AP]

Feinstein J. S., Adolphs R., Damasio A. & Tranel D. (2011) The human amygdala and the induction and experience of fear. *Current Biology* **21**:1–5. [JH]

Felleman D. J. & Van Essen D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* **1**:1–47. [AP]

First M. B. (2012) The National Institute of Mental Health Research Domain Criteria (RDoC) project: Moving towards a neuroscience-based diagnostic classification in psychiatry. In: *Philosophical issues in psychiatry II: Nosology*, ed. K. S. Kendler & J. Parnas. Oxford University Press. [TKP]

Flagel S. B., Pine D. S., Ahmari S. E., First M. B., Friston K. J., Mathys C., Redish A. D., Schmack K., Smoller J. W. & Thapar A. (2016) A novel framework for improving

psychiatric diagnostic nosology. In: *Computational psychiatry: New perspectives on mental illness*, ed. A. D. Redish & J. A. Gordon, Ch. 10, pp. 167–99. MIT Press. [ADR]
Fodor J. A. (1974) Special sciences (or: The disunity of science as a working hypothesis). *Synthese* **28**:97–115. [arDB, TKP]
Fodor J. A. (1985) The modularity of mind. *Behavioral and Brain Sciences* **8**(1):1–5. [AP]
Folke C., Carpenter S., Walker B., Scheffer M., Chapin T. & Rockström J. (2010) Resilience thinking: Integrating resilience, adaptability and transformability. *Ecology and Society* **15**(4): article 20. Online publication. Available at: http://www.ecologyandsociety.org/vol15/iss4/art20/. [DC]
Forbes M. K., Wright A. G. C., Markon K. E. & Krueger R. F. (2017a) Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology* **126**:969–88. [RJM]
Forbes M. K., Wright A. G. C., Markon K. E. & Krueger R. F. (2017b) Further evidence that psychopathology networks have limited replicability and utility: Response to Borsboom et al. (2017) and Steinley et al. (2017). *Journal of Abnormal Psychology* **126**:1011–16. [RJM]
Fornito A., Zalesky A. & Breakspear M. (2015) The connectomics of brain disorders. *Nature Reviews Neuroscience* **16**:159–72. [AP]
Forst C. V. (2002) Network genomics: A novel approach for the analysis of biological systems in the post-genomic era. *Molecular Biology Reports* **29**:265–80. [RJM]
Fox A. S., Oler J. A., Birn R. M., Shackman A. J., Alexander A. L. & Kalin N. H. (2018) Functional connectivity within the primate extended amygdala is heritable and predicts early-life anxious temperament. *Journal of Neuroscience* **38**:7611–7621. Available at: https://doi.org/10.1523/JNEUROSCI.0102-18.2018. [JH]
Fox A. S., Oler J. A., Shackman A. J., Shelton S. E., Raveendran M., McKay D. R., Converse A. K., Alexander A. L., Davidson R. J., Blangero J., Rogers J. & Kalin N. H. (2015) Intergenerational neural mediators of early-life anxious temperament. *Proceedings of the National Academy of Sciences USA* **112**:9118–22. [JH]
Fox A. S., Oler J. A., Shelton S. E., Nanda S. A., Davidson R. J., Roseboom P. H. & Kalin N. H. (2012) Central amygdala nucleus (Ce) gene expression linked to increased trait-like Ce metabolism and anxious temperament in young primates. *Proceedings of the National Academy of Sciences USA* **109**:18108–13. [JH]
Fox A. S. & Shackman A. J. (in press) The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. *Neuroscience Letters*. [JH]
Foxe J. J. & Schroeder C. E. (2005) The case for feedforward multisensory convergence during early cortical processing. *NeuroReport* **16**:419–23. [AP]
Frances A. (2014) Resuscitating the biopsychosocial model. *The Lancet Psychiatry* **1**:496–97. [ADR]
Francken J. C. & Slors M. (2014) From commonsense to science, and back: The use of cognitive concepts in neuroscience. *Consciousness and Cognition* **29**:248–58. Available at: https://doi.org/10.1016/j.concog.2014.08.019. [aDB, MS]
Francken J. C. & Slors M. (2018) Neuroscience and everyday life: Facing the translation problem. *Brain and Cognition* **120**:67–74. Available at: https://doi.org/10.1016/j.bandc.2017.09.004. [MS]
Franić S., Dolan C. V., Borsboom D. & Boomsma D. I. (2012) Structural equation modeling in genetics. In: *Handbook of structural equation modeling*, ed. R. H. Hoyle, pp. 617–35. Guilford Press. [aDB]
Fried E. I. (2015) Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology* 6: article 309. (Online publication). doi: 10.3389/fpsyg.2015.00309. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00309/full. [JH]
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**:256–65. [aDB]
Fried E. I. & Cramer A. O. J. (2017) Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science* **12**(6):999–1020. Available at: https://doi.org/10.1177/1745691617705892 and at: http://journals.sagepub.com/doi/abs/10.1177/1745691617705892. [arDB, NJ]
Fried E. I. & Nesse R. M. (2015) Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders* **172**:96–102. [JH, BCN]
Fried E. I., van Borkulo C. D., Cramer A. O. J., Boschloo L. B., 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. [arDB, RJM]
Friston K. (2010) The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* **11**(2):127–38. doi:10.1038/nrn2787. [MD, IGE]
Friston K. J., Redish A. D. & Gordon J. A. (2017) Computational nosology and precision psychiatry. *Computational Psychiatry* **1**(1):2–23. [IGE]
Furlan M., Wann J. P. & Smith A. T. (2013) A representation of changing heading direction in human cortical areas pVIP and CSv. *Cerebral Cortex* **24**:2848–58. [AP]
Galatzer-Levy I. R. & Bryant R. A. (2013) 636,120 Ways to have posttraumatic stress disorder. *Perspectives on Psychological Science* **8**(6):651–62. Available at: https://doi.org/10.1177/1745691613504115. [JH]
Galilei G. (1615/2012) Observations on the Copernican theory. In: *Selected writings of Galileo Galilei*, trans. W. R. Shea & M. Davie, pp. 96–114. Oxford University Press. [RJM]
Gallese V. (2008) Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience* **3**:317–33. [AP]
Gandal M. J., Leppa V., Won H., Parikshak N. N. & Geschwind D. H. (2016) The road to precision psychiatry: Translating genetics into disease mechanisms. *Nature Neuroscience* **19**(11):1397–407. doi: 10.1038/nn.4409. [JH]
Gargiulo G. J. (1998) Meaning and metaphor in psychoanalytic education. *Psychoanalytic Review* **85**:413–22. [RFB]
Gerrans P. (2013) Delusional attitudes and default thinking. *Mind and Language* **28**(1):83–102. Available at: https://doi.org/10.1111/mila.12010. [MS]
Geschwind D. H. & Flint J. (2015) Genetics and genomics of psychiatric disease. *Science* **349**(6255):1489–94. doi: 10.1126/science.aaa8954. [JH]
Geurts H. M., Corbett B. & Solomon M. (2009) The paradox of cognitive flexibility in autism. *Trends in Cognitive Science* **13**:74–82. [DC]
Ghazanfar A. A. & Schroeder C. E. (2006) Is neocortex essentially multisensory? *Trends in Cognitive Science* **10**:278–85. [AP]
Gillett C. (2003) The metaphysics of realization, multiple realization and the special sciences. *Journal of Philosophy* **100**:591–603. [aDB]
Global Burden of Disease Collaborators (2016) Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* **388**:1545–602. [JH]
Golino H. F. & Demetriou A. (2017) Estimating the dimensionality of intelligence like data using exploratory graph analysis. *Intelligence* **62**:54–70. [rDB]
Golino H. F. & Epskamp S. (2017) Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE* **12**(6): e0174035. [rDB]
Gone J. P. & Kirmayer L. J. (2010) On the wisdom of considering culture and context in psychopathology. In: *Contemporary directions in psychopathology: Scientific foundations of the DSM-V and ICD-11*, ed. T. Millon, R. F. Krueger & E. Simonsen, pp. 72–96. Guilford Press. [NJ]
Gong Q. & He Y. (2015) Depression, neuroimaging and connectomics: A selective overview. *Biological Psychiatry* **77**(3):223–35. doi:10.1016/j.biopsych.2014.08.009. [BCN]
Gordon J. A. & Redish A. D. (2016) On the cusp. Current challenges and promises in psychiatry. In: *Computational psychiatry: New perspectives on mental illness*, ed. A. D. Redish & J. A. Gordon, pp. 3–14. MIT Press. [JH]
Gordon R. M. (1996) "Radical" simulationism. In: *Theories of theories of mind*, ed. P. Carruthers & P. K. Smith, pp. 11–21. Cambridge University Press. [MS]
Gottlieb G. (2002) Developmental-behavioral initiation of evolutionary change. *Psychological Review* **109**:211–18. [DKO]
Gotzsche P. C. (2013) *Deadly medicines and organized crime*. CRC Press. [JPAI]
Granger H. J. (1998) Cardiovascular physiology in the twentieth century: Great strides and missed opportunities. *The American Journal of Physiology* **275**(6, Pt 2): H1925–36. [IGE]
Grant S. G. (2015) The molecular evolution of the vertebrate behavioural repertoire. *Philosophical Transactions of the Royal Society, B: Biological Sciences* **371**:1–9. Available at: http://dx.doi.org/10.1098/rstb.2015.0051. [DKO]
Greenberg G. (2013) *The book of woe: The DSM and the unmaking of psychiatry*. Penguin. [aDB]
Greer S. M., Trujillo A. J., Glover G. H. & Knutson B. (2014) Control of nucleus accumbens activity with neurofeedback. *Neuroimage* **96**:237–44. [JH]
Grunspan D. Z., Nesse R. M., Barnes M. E. & Brownell S. E. (2018) Core principles of evolutionary medicine. *Evolution, Medicine, and Public Health* **2018**:13–23. [OJ]
Gulbins E., Palmada M., Reichel M., Luth A., Bohmer C., Amato D., Müller C. P., Tischbirek C. H., Groemer T. W., Tabatabai G., Becker K. A., Tripal P., Staedtler S., Ackermann T. F., van Brederode J., Alzheimer C., Weller M., Lang U. E., Kleuser B., Grassme H. & Kornhuber J. (2013) Acid sphingomyelinase-ceramide system mediates effects of antidepressant drugs. *Nature Medicine* **19**:934–38. [CPM]
Guloksuz S., Pries L. K. & van Os J. (2017) Application of network methods for understanding mental disorders: Pitfalls and promise. *Psychological Medicine* **47**(16):2743–52. doi: 10.1017/S0033291717001350. [aDB]
Guze S. B. (1989) Biological psychiatry: Is there any other kind? *Psychological Medicine* **19**:315–23. [aDB]
Hacking I. (1999) *The social construction of what?* Harvard University Press. [aDB, AP]
Hallquist M. N. & Pilkonis P. A. (2012) Refining the phenotype of borderline personality disorder: Diagnostic criteria and beyond. *Personality Disorders: Theory, Research, and Treatment* **3**(3):228–46. [RFB]
Han S., Northoff G., Vogeley K., Wexler B. E., Kitayama S. & Varnum M. E. (2013) A cultural neuroscience approach to the biosocial nature of the human brain. *Annual Review of Psychology* **64**:335–59. [AP]

Hare E., Glahn D. C., Dassori A., Raventos H., Nicolini H., Ontiveros A., Medina R., Mendoza R., Jerez A., Munoz R., Almasy L. & Escamilla M. A. (2010) Heritability of age of onset of psychosis in schizophrenia. *American Journal of Medical Genetics: Part B, Neuropsychiatric Genetics* **153B**(1):298–302. Available at: https://doi.org/10.1002/ajmg.b.30959. [DKO]

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. [aDB]

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**:151–62. [aDB]

Hart C. (2013) *High price: Drugs, neuroscience and discovering myself*. Penguin. [MF]

Hasin D. S., Shmulewitz D., Stohl M., Greenstein E., Aivadyan C., Morita K., Saha T., Aharonovich E., Jung J., Zhang H., Nunes E. V. & Grant B. F. (2015) Procedural validity of the AUDADIS-5 depression, anxiety and post-traumatic stress disorder modules: Substance abusers and others in the general population. *Drug and Alcohol Dependence* **152**:246–56. [JH]

Haslam N., Holland E. & Kuppens P. (2012) Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine* **42**:903–20. [RFB]

Heather N. (2018) Rethinking addiction. *The Psychologist* **31**(1):24–28. [MF]

Heather N., Best D., Kawalek A., Field M., Lewis M., Rotgers F., Reinout W. & Heim D. (2018) Challenging the brain disease model of addiction: European launch of the addiction theory network. (Editorial). *Addiction Research and Theory* **26**(4):249–55. doi:10.1080/16066359.2017.1399659. [MF]

Heilig M. & Koob G. F. (2007) A key role for corticotropin-releasing factor in alcohol dependence. *Trends in Neurosciences* **30**:399–406. [CPM]

Herpertz S. C., Huprich S. K., Bohus M., Chanen A., Goodman M., Mehlum L., Moran P., Newton-Howes G., Scott L. & Sharp C. (2017) The challenge of transforming the diagnostic system of personality disorders. *Journal of Personality Disorders* **31**:577–89. [RFB]

Heyman G. M. (1996) Resolving the contradictions of addiction. *Behavioral and Brain Sciences* **19**(4):561–610. [MF]

Heyman G. M. (2013) Quitting drugs: Quantitative and qualitative features. *Annual Review of Clinical Psychology* **9**:29–59. [MF]

Hidaka B. H. (2012) Depression as a disease of modernity: Explanations for increasing prevalence. *Journal of Affective Disorders* **140**:205–14. [OJ]

Hills T. T. (2006) Animal foraging and the evolution of goal-directed cognition. *Cognitive Science* **30**:3–41. [DKO]

Hinton D. E. & Good B. J. (2016) The culturally sensitive assessment of trauma: Eleven analytic perspectives, a typology of errors, and the multiplex models of distress generation. In: *Culture and PTSD: Trauma in global and historical perspective*, ed. B. J. Good & D. E. Hinton, pp. 50–113. University of Pennsylvania Press. [NJ]

Hochstein E. (2016) One mechanism, many models: A distributed theory of mechanistic explanation. *Synthese* 193:1387–407. Available at: https://doi.org/10.1007/s11229-015-08. [TW]

Hoffmann S. G., Asnaani A., Vonk I. J. J., Sawyer A. T. & Fang A. (2012) The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research* **36**:427–40. [rDB]

Hollingsworth A., Ruhm C. & Simon K. (2017) Macroeconomic conditions and opioid abuse. NBER Working Paper No. 23192. NBER Working Paper Series. National Bureau of Economic Research. Available at: http://www.nber.org/papers/w23192. [MF]

Hoogman M., Bralten J., Hibar D. P., Mennes M., Zwiers M. P., Schweren L. S. J., van Hulzen K. J. E., Medland S. E., Shumskaya E., Jahanshad N., Zeeuw P., Szekely E., Sudre G., Wolfers T., Onnink A. M. H., Dammers J. T., Mostert J. C., Vives-Gilabert Y., Kohls G., Oberwelland E., Seitz J., Schulte-Rüther M., Ambrosino S., Doyle A. E., Høvik M. F., Dramsdahl M., Tamm L., van Erp T. G. M., Dale A., Schork A., Conzelmann A., Zierhut K., Baur R., McCarthy H., Yoncheva Y. N., Cubillo A., Chantiluke K., Mehta M. A., Paloyelis Y., Hohmann S., Baumeister S., Bramati I., Mattos P., Tovar-Moll F., Douglas P., Banaschewski T., Brandeis D., Kuntsi J., Asherson P., Rubia K., Kelly C., Martino A. D., Milham M. P., Castellanos F. X., Frodl T., Zentis M., Lesch K. P., Reif A., Pauli P., Jernigan T. L., Haavik J., Plessen K. J., Lundervold A. J., Hugdahl K., Seidman L. J., Biederman J., Rommelse N., Heslenfeld D. J., Hartman C. A., Hoekstra P. J., Oosterlaan J., Polier G. V., Konrad K., Vilarroya O., Ramos-Quiroga J. A., Soliva J. C., Durston S., Buitelaar J. K., Faraone S. V., Shaw P., Thompson P. M. & 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**:310–19. [aDB]

Hopwood C. J. & Bornstein R. F., eds. (2014) *Multimethod clinical assessment*. Guilford Press. [RFB]

Horgan T. (1993) Nonreductive materialism and the explanatory autonomy of psychology. In: *Naturalism: A critical appraisal*, ed. S. Wagner & R. Warner, pp. 295–320. University of Notre Dame Press. [aDB]

Horner R. H. (2010) Positive behavior supports focus on autism and other developmental disabilities. 15(2):97–105. doi:10.1177/108835760001500205. [LT]

Horst S. (2007) *Beyond reduction: Philosophy of mind and post-reductionist philosophy of science*. Oxford University Press. [DR]

Horvath S. & Morf C. C. (2009) Narcissistic defensiveness: Hypervigilance and avoidance of worthlessness. *Journal of Experimental Social Psychology* **25**:1252–58. [RFB]

Huhn M., Tardy M., Spineli L.-M., Kissling W., Förstl H., Pitschel-Walz G., Leucht C., Samara M., Dold M., Davis J. M. & Leucht S. (2014) Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: A systematic overview of meta-analyses. *JAMA Psychiatry* **71**:706–15. [JPAI]

Hurn J. D. (1998) *The history of general paralysis of the insane in Britain 1830 to 1950*. Doctoral dissertation. University College London. Available at: http://discovery.ucl.ac.uk/1349281/1/339949.pdf. [aDB]

Hutchinson L. & Romero D. (2016) Precision or imprecision medicine? *Clinical Oncology* **13**:712–13. [RFB]

Huys Q. J., Maia T. V. & Frank M. J. (2016) Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience* **19**(3):404–13. doi: 10.1038/nn.4238. [ADR]

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

Hyland M. E. (2017) A new paradigm to explain functional disorders and the adaptive network theory of chronic fatigue syndrome and fibromyalgia syndrome. In: *Resistance and renewal in theoretical psychology*, ed. G. B. Sullivan, J. Cresswell, B. Ellis, M. Morgan & E. Schraube, pp. 21–31. Captus University Publications. [MEH]

Hyland M. E., Hinton C., Hill C., Whalley B., Jones R. C. & Davies A. F. (2016) Explaining unexplained pain to fibromyalgia patients: Finding a narrative that is acceptable to patients and provides a rationale for evidence-based interventions. *British Journal of Pain* **10**(3):156–61. [MEH]

Hyman S. E., Malenka R. C. & Nestler E. J. (2006) Neural mechanisms of addiction: The role of reward-related learning and memory. *Annual Reviews in Neurosciences* **29**:565–98. [CPM]

Immordino-Yang M. H. & Damasio A. (2007) We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education* **1**:3–10. [AP]

Inman C. S., Bijanki K. R., Bass D. I., Gross R. E., Hamann S. & Willie J. T. (in press) Human amygdala stimulation effects on emotion physiology and emotional experience. *Neuropsychologia*. [JH]

Insel T. R. (2013) Transforming diagnosis. *My blog; Tom Insel, M.D., NIMH Director*, April. 29, 2013. Available at: http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml?utm_source=rss_readers&utm_medium=rss&utm_campaign=rss-full. [RJM]

Insel T. R., Cuthbert B., Garvey M., Heinssen R., Pine D. S., Quinn K., Sanislow C. & Wang P. (2010) Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry* **167**(7):748–51. doi:10.1176/appi.ajp.2010.09091379. [IGE, RJM, TKP]

Insel T. R. (2014) The NIMH Research Domain Criteria (RDoC) project: Precision medicine for psychiatry. *American Journal of Psychiatry* **171**(4):395–97. [ADR]

Insel T. R. & Cuthbert B. N. (2015) Brain disorders? Precisely. *Science* **348**:499–500. [arDB]

Ioannidis J. P. (2008) Effectiveness of antidepressants: An evidence myth constructed of a thousand clinical trials? *Philosophy, Ethics, and Humanities in Medicine* **3**:14. doi: 10.1186/1747-5341-3-14. Available at: https://peh-med.biomedcentral.com/articles/10.1186/1747-5341-3-14. [JPAI]

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**:187–96. [aDB]

Jaspers K. (1923/1963) *General psychopathology*. Manchester University Press. [aDB]

Jayawickreme N., Mootoo C., Fountain C., Rasmussen A., Jayawickreme E. & Bertuccio R. (2017) Post-conflict struggles as networks of problems: A network analysis of trauma, daily stressors and psychological distress among Sri Lankan war survivors. *Social Science and Medicine* **190**:119–32. [NJ]

Johnstone L. & Dallos R. (2013) *Formulation in psychology and psychotherapy: Making sense of people's problems*. Routledge. [AP]

Jones P. J., Heeren A. & McNally R. J. (2017) Commentary: A network theory of mental disorders. *Frontiers of Psychology* 8: article 1305. doi:10.3389/fpsyg.2017.01305. (Online publication). Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01305/full. [rDB]

Jordans M. J. D., Semrau M., Thornicroft G. & van Ommeren M. (2012) Role of current perceived needs in explaining the association between past trauma exposure and distress in humanitarian settings in Jordan and Nepal. *British Journal of Psychiatry* **201**:276–81. [NJ]

Juignet P. (2015) Edgar Morin et la complexité [Edgar Morin and complexity]. *Philosophie, Science et Société*. (Online publication). Available at: https://philosciences.com/philosophie-generale/complexite-systeme-organisation-emergence/17-edgar-morin-complexite. [MD]

Kaiser M. I. (2015) *Reductive explanation in the biological sciences*. Springer. [KT]

Kalin N. H., Fox A. S., Kovner R., Riedel M. K., Fekete E. M., Roseboom P. H., Tromp D. P., Grabow B. P., Olsen M. E., Brodsky E. K., McFarlin D. R., Alexander A. L.,

Emborg M. E., Block W. F., Fudge J. L. & Oler J. A. (2016) Overexpressing corticotropin-releasing hormone in the primate amygdala increases anxious temperament and alters its neural circuit. *Biological Psychiatry* **80**:345–55. [JH]

Kalis A. (2011) *Failures of agency: Irrational behavior and self-understanding*. Lexington Books. [aDB]

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

Kambeitz J., Kambeitz-Ilankovic L., Cabral C., Dwyer D. B., Calhoun V. D., van den Heuvel M. P., Falkai P., Koutsouleris N. & Malchow B. (2016) Aberrant functional whole-brain network architecture in patients with schizophrenia: A meta-analysis. *Schizophrenia Bulletin* **42**(Suppl. 1):S13–S21. doi:10.1093/schbul/sbv174. [BCN]

Kanwisher N., McDermott J. & Chun M. M. (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* **17**:4302–11. [AP]

Kapur S., Phillips A. G. & Insel T. R. (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it?" *Molecular Psychiatry* **17**(12):1174–79. [KT]

Katahira K. & Yamashita Y. (2017) A theoretical framework for evaluating psychiatric research strategies. *Computational Psychiatry* **1**(1):184–207. [IGE]

Kauffman S. A. (1971) Articulation of parts explanations in biology and the rational search for them. *Boston Studies in the Philosophy of Science* **8**:257–72. [KT]

Kauffman T., Hamilton R., Keenan J. P., Warde A. & Pascual-Leone A. (2000) The role of visual cortex in tactile Braille reading: The early blind, the sighted, and the blindfolded. *Annals of Neurology* **48**:418–19. [AP]

Kawamichi H., Kitada R., Yoshihara K., Takahashi H. K. & Sadato N. (2015) Interpersonal touch suppresses visual processing of aversive stimuli. *Frontiers in Human Neuroscience* 9: article 164. (Online article). doi: 10.3389/fnhum.2015.00164. Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2015.00164/full. [AP]

Kazdoba T. M., Leach P. T., Yang M., Silverman J. L., Solomon M. & Crawley J. N. (2016) Translational mouse models of autism: Advancing toward pharmacological therapeutics. In: *Translational Neuropsychopharmacology*, ed. T. W. Robbins & B. J. Sahakian, pp. 1–52. [Series: *Current Topics in Behavioral Neurosciences, vol. 28*]. Springer. doi/10.1007/7854_2015_5003. [NMB]

Keiflin R. & Janak P. H. (2015) Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron* **88**(2):247–63. doi:10.1016/j.neuron.2015.08.037. [IGE]

Keller M. C., Neale M. C. & Kendler K. S. (2007) Association of different adverse life events with distinct patterns of depressive symptoms. *American Journal of Psychiatry* **164**(10):1521–29. [AT]

Keller M. C. & Nesse R. M. (2006) The evolutionary significance of depressive symptoms: Different adverse situations lead to different depressive symptom patterns. *Journal of Personality and Social Psychology* **91**:316–30. [AT]

Kemmelmeier M. & Kühnen U. (2012) Culture as process: The dynamics of cultural stability and change. *Social Psychology* **43**:171–73. [DC]

Kendler K. S. (2005) Toward a philosophical structure for psychiatry. *American Journal of Psychiatry* **162**:433–40. [aDB]

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

Kendler K. S. (2012a) Levels of explanation in psychiatric and substance use disorders: Implications for the development of an etiologically based nosology. *Molecular Psychiatry* **17**:11–21. [aDB]

Kendler K. S. (2012b) The dappled nature of causes of psychiatric illness: Replacing the organic-functional/hardware-software dichotomy with empirically based pluralism. *Molecular Psychiatry* **17**:377–88. [JH]

Kendler K. S. (2013) What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn. *Molecular Psychiatry* **18**(10):1058–66. [KT]

Kendler K. S. (2014) The structure of psychiatric science. *The American Journal of Psychiatry* **171**(9):931–38. Available at: https://doi.org/10.1176/appi.ajp.2014.13111539. [HMvL]

Kendler K. S., Zachar P. & Craver C. (2011) What kinds of things are psychiatric disorders? *Psychological Medicine* **41**:1143–50. [aDB]

Kendrick T., El-Gohary M., Stuart B., Gilbody S., Churchill R., Aiken L., Bhattacharya A., Gimson A., Brütt A. L., de Jong K. & Moore M. (2016) Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews* 7: article CD011119. Available at: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011119.pub2/media/CDSR/CD011119/CD011119_standard.pdf. [JPAI]

Kessler R. C., Amminger G. P., Aguilar-Gaxiola S., Alonso J., Lee S. & Ustun T. B. (2007) Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry* **20**:359–64. Available at: https://doi.org/10.1097/YCO. [DKO]

Kessler R. C., Chiu W. T., Demler O. & Walters E. E. (2005) Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**:617–27. [JH]

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

Kim H. U. (2012) Autism across cultures: Rethinking autism. *Disability and Society* **27**(4):535–45. Available at: https://doi.org/10.1080/09687599.2012.659463. [KH]

Kim J. (1982) Psychophysical supervenience. *Philosophical Studies* **41**:51–70. [arDB]

Kim J. (1984) Concepts of supervenience. *Philosophy and Phenomenological Research* **45**:153–76. [arDB]

Kim J. (1992) Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research* **52**(1):1–26. [MIE]

Kincaid H., Daniels E., Dellis A., Hofmeyr A., Rousseau J., Sharp C. & Ross D. (2013) A taxometric analysis of the performance of the problem gambling severity index in a South African national urban prevalence sample. *International Journal of Gambling Studies* **29**:377–92. [DR]

Kirmayer L. J. & Ryder A. G. (2016) Culture and psychopathology. *Current Opinion in Psychology* **8**:143–48. [NJ]

Kirsch I. & Hyland M. E. (1987) How thoughts affect the body: A metatheoretical framework. *The Journal of Mind and Behavior* **8**:417–34. [MEH]

Kitada R., Johnsrude I. S., Kochiyama T. & Lederman S. J. (2010) Brain networks involved in haptic and visual identification of facial expressions of emotion: An fMRI study. *Neuroimage* **49**:1677–89. [AP]

Kitayama S. & Park J. (2010) Cultural neuroscience of the self: Understanding the social grounding of the brain. *Social Cognitive and Affective Neuroscience* **5**:111–29. [AP]

Kitayama S. & Uskul A. K. (2011) Culture, mind, and the brain: Current evidence and future directions. *Annual Review of Psychology* **62**:419–49. [DC]

Kleinman A. (1988) *Rethinking psychiatry: From cultural category to personal experience*. The Free Press. [NJ]

Knudsen E. I. (2011) Control from below: The role of a midbrain network in spatial attention. *European Journal of Neuroscience* **33**(11):1961–72. Available at: https://doi.org/10.1111/j.1460-9568.2011.07696.x. [NMB]

Koob G. F. (1992) Drugs of abuse: Anatomy, pharmacology and function of reward pathways. *Trends in Pharmacological Sciences* **13**:177–84. [CPM]

Koob G. F. & Volkow N. D. (2016) Neurobiology of addiction: A neurocircuitry analysis. *Lancet Psychiatry* **3**:760–73. [CPM]

Kotchen T. A. (2011) Historical trends and milestones in hypertension research: A model of the process of translational research. *Hypertension* **58**(4):522–38. doi:10.1161/HYPERTENSIONAHA.111.177766. [IGE]

Kotov R., Krueger R. F., Watson D., Achenbach T. M., Althoff R. R., Bagby R. M., Brown T. A., Carpenter W. T., Caspi A., Clark L. A., Eaton N. R., Forbes M. K., Forbush K. T., Goldberg D., Hasin D., Hyman S. E., Ivanova M. Y., Lynam D. R., Markon K., Miller J. D., Moffitt T. E., Morey L. C., Mullins-Sweatt S. N., Ormel J., Patrick C. J., Regier D. A., Rescorla L., Ruggero C. J., Samuel D. B., Sellbom M., Simms L. J., Skodol A. E., Slade T., South S. C., Tackett J. L., Waldman I. D., Waszczuk M. A., Widiger T. A., Wright A. G. C. & Zimmerman M. (2017) The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology* **126**:454–77. [RFB, JH, RJM, rDB]

Koukouli F., Rooy M., Tziotis D., Sailor K. A., O'Neill H. C., Levenga J., Witte M., Nilges M., Changeux J. P., Hoeffer C. A., Stitzel J. A., Gutkin B. S., DiGregorio D. A. & Maskos U. (2017) Nicotine reverses hypofrontality in animal models of addiction and schizophrenia. *Nature Medicine* **23**:347–54. [CPM]

Koutsouleris N., Kambeitz-Ilankovic L., Ruhrmann S., Rosen M., Ruef A., Dwyer D. B., Paolini M., Chisholm K., Kambeitz J., Haidl T., Schmidt A., Gillam J., Schultze-Lutter F., Falkai P., Reiser M., Riecher-Rössler A., Upthegrove R., Hietala J., Salokangas R. K. R., Pantelis C., Meisenzahl E., Wood S. J., Beque D., Brambilla P., Borgwardt S. & PRONIA Consortium. (2018) Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: A multimodal, multisite machine learning analysis. *JAMA Psychiatry* **75**(11):1156–72. doi:10.1001/jamapsychiatry.2018.2165. [JH]

Kozak M. J. & Cuthbert B. N. (2016) The NIMH research domain criteria initiative: Background, issues, and pragmatics. *Psychophysiology* **53**:286–97. [JH]

Kraepelin E. & Lange J. (1927) *Psychiatrie, vol. I*, 9th edition. J. A. Barth. [aDB]

Krakauer J. W., Ghazanfar A. A., Gomez-Marin A., MacIver M. A. & Poeppel D. (2017) Neuroscience needs behavior: Correcting a reductionist bias. *Neuron* 93(3):480–90. Available at: https://doi.org/10.1016/j.neuron.2016.12.041. [NMB]

Kreek M. J., Nielsen D. A., Butelman E. R. & Laforge K. S. (2005) Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nature Neuroscience* **8**:1450–57. [CPM]

Krueger R. F. & Markon K. E. (2014) The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology* **10**:477–501. [RFB]

Lacasse J. R. & Leo J. (2005) Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine* **2**:e392. Available at: https://doi.org/10.1371/journal.pmed.0020392. [aDB]

Ladyman J. & Ross D. (2007) *Every thing must go: Metaphysics naturalized*. Oxford University Press. [DR]

Lahey B. B., Van Hulle C. A., Singh A. L., Waldman I. D. & Rathouz P. J. (2011) Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry* **68**:181–89. [RJM]

Le Bihan D. (2003) Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience* **4**:469–80. [AP]

Lederman S. J., Klatzky R. L., Abramowicz A., Salsman K., Kitada R. & Hamilton C. (2007) Haptic recognition of static and dynamic expressions of emotion in the live face. *Psychological Science* **18**:158–64. [AP]

Ledford H. (2014) Medical research: If depression were cancer. [Editorial]. *Nature* **515**:182–84. [aDB]

Lekkas S. & Mikhailov L. (2010) Evolving fuzzy medical diagnosis of Pima Indians diabetes and dermatological conditions. *Artificial Intelligence in Medicine* **50**:117–26. [RFB]

Lenfant C. (2010) Chest pain of cardiac and noncardiac origin. *Metabolism* **59**(Suppl. 1): S41–S46. doi:10.1016/j.metabol.2010.07.014. [IGE]

Leshner A. (1997) Addiction is a brain disease, and it matters. *Science* **278**(5335): 45–47. [MF]

Leucht S., Leucht C., Huhn M., Chaimani A., Mavridis D., Helfer B., Samara M., Rabaioli M., Bächer S., Cipriani A., Geddes J. R., Salanti G. & Davis J. M. (2017) Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, Bayesian meta-analysis, and meta-regression of efficacy predictors. *American Journal of Psychiatry* **174**:927–42. [JPAI]

Levy N. (2013) Addiction is not a brain disease (and it matters). *Frontiers in Psychiatry* 4: article 24. (Online publication). doi: 10.3389/fpsyt.2013.00024 Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2013.00024/full. [MF]

Lewis D. K. (1966) An argument for the identity theory. *The Journal of Philosophy* **63**: 17–25. [aDB]

Lewis M. (2017) Addiction and the brain: Development, not disease. *Neuroethics* **10**: 7–18. [MF]

Lilienfeld S. O. (2014) The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy* **62**:129–39. [JH]

Lindquist K. A. & Barrett L. F. (2012) A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in Cognitive Sciences* **16** (11):533–40. [LPes]

Littlewood R (2002) *Pathologies of the West: An anthropology of mental illness in Europe and America*. Cornell University Press. [NJ]

Lloyd D. (2000) Virtual lesions and the not so-modular brain. *Journal of the International Neuropsychological Society* **6**:627–35. [AP]

Lo C.-Y. Z., Su T.-W., Huang C.-C., Hung C.-C., Chen W.-L., Lan T.-H., Lin C. P. & Bullmore E. T. (2015) Randomization and resilience of brain functional networks as systems-level endophenotypes of schizophrenia. *Proceedings of the National Academy of Sciences USA* **112**(29):9123. doi:10.1073/pnas.1502052112. [BCN]

Loomis J. M., Klatzky R. L. & Lederman S. J. (1991) Similarity of tactual and visual picture recognition with limited field of view. *Perception* **20**:167–77. [AP]

Loriedo C. (2005) Resilienza e fattori di protezione nella psicoterapia familiare sistemica [Resilience and protection factors in the Systemic family psychotherapy]. *Rivista di Psicoterapia Relazionale* **21**:1000–24. [AP]

Luhrmann T. M., Padmavati R., Tharoor H. & Osei A. (2014) Differences in voice-hearing experiences of people with psychosis in the USA, India and Ghana: Interview-based study. *British Journal of Psychiatry* **206**:41–44. [DC]

Luria A. R. (2012) *Higher cortical functions in man*. Springer Science and Business Media. [AP]

Macaluso E. & Driver J. (2005) Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences* **28**:264–71. [AP]

MacDonald III A. W., Zick J. L., Netoff T. I. & Chafee M. V. (2016) The computation of collapse: Can reliability engineering shed light on mental illness? In: *Computational psychiatry: New perspectives on mental illness*, ed. A. D. Redish & J. A. Gordon, Ch. 9, pp. 153–66. MIT Press. [ADR]

Macefield R. C., Jacobs M., Korfage I. J., Nicklin J., Whistance R. N., Brookes S. T., Sprangers M. A. & Blazeby J. M. (2014) Developing core outcomes sets: Methods for identifying and including patient-reported outcomes (PROs). *Trials* **15**:49. doi: 10.1186/1745-6215-15-49. [JPAI]

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: article 1379. (Online publication). doi: 10.3389/fpsyg.2014.01379. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2014. 01379/full. [aDB]

Mandy W. & Lai M.-C. (2017) Towards sex- and gender-informed autism research. (Editorial to the Special Issue on women and girls on the autism spectrum). *Autism* **21**(6):643–45. Available at: https://doi.org/10.1177/1362361317706904. [KH]

Marcks B. A., Weisberg R. B., Edelen M. O. & Keller M. B. (2010) The relationship between sleep disturbance and the course of anxiety disorders in primary care patients. *Psychiatry Research* **178**(3):487–92. [ADR]

Markou A., Chiamulera C., Geyer M. A., Tricklebank M. & Steckler T. (2009) Removing obstacles in neuroscience drug discovery: The future path for animal models. *Neuropsychopharmacology* **34**(1):74–89. Available at: https://doi.org/10.1038/npp.2008.173. [NMB]

Marmanidis H., Holme G. & Hafner R. J. (1994) Depression and somatic symptoms: A cross cultural study. *Australian and New Zealand Journal of Psychiatry* **28**:274–78. [aDB]

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* **7**:1–21. doi:10.1080/00273171.2017.1379379. [arDB]

Martinez-Lavin M., Infante O. & Lerma C. (2008) Hypothesis: The chaos and complexity theory may help our understanding of fibromyalgia and similar maladies. *Seminars in Arthritis and Rheumatism* **37**(4):260–64. [MEH]

Maruska K. P. & Fernald R. D. (2018) Astatotilapia burtoni: A model system for analyzing the neurobiology of behavior. *ACS Chemical Neuroscience* **9**(8):1951–62. doi:10.1021/acschemneuro.7b00496. [NMB]

Mattick R. P., Breen C., Kimber J. & Davoli M. (2009) Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. *Cochrane Database of Systematic Reviews* **3**:CD002209. doi: 10.1002/14651858.CD002209.pub2 [MF]

Mayr E. (1982) *The growth of biological thought*. Harvard University Press. [AT]

McBride W. J., Murphy J. M. & Ikemoto S. (1999) Localization of brain reinforcement mechanisms: Intracranial self-administration and intracranial place-conditioning studies. *Behavioural Brain Research* **101**:129–52. [CPM]

McClelland J. L., McNaughton B. L. & O'Reilly R. C. (1995) Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* **102**(3):419–57. [ADR]

McCreary A. C., Müller C. P. & Filip M. (2015) Psychostimulants: Basic and clinical pharmacology. *International Review Neurobiology* **120**:41–83. [CPM]

McDowell J. (1996) *Mind and world*. Harvard University Press. [aDB]

McGinn C. (1982) *The character of mind*. Oxford University Press. [aDB]

McGraw L. A. & Young L. J. (2010) The prairie vole: An emerging model organism for understanding the social brain. *Trends in Neurosciences* **33**(2):103–109. doi:10.1016/j.tins.2009.11.006. [NMB]

McNally R. J. (1990) Psychological approaches to panic disorder: A review. *Psychological Bulletin* **108**:403–19. [aDB]

McNally R. J. (2002) Anxiety sensitivity and panic disorder. *Biological Psychiatry* **52**: 938–46. [aDB]

McNally R. J. (2011) *What is mental illness?* The Belknap Press of Harvard University Press. [RJM]

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

McNally R. J., Heeren A. & Robinaugh D. J. (2017a) A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *European Journal of Psychotraumatology* **8** (Suppl. 3):e1341276. Available at: http://doi.org/10.1080/20008198.2017.1341276. [aDB]

McNally R. J., Mair P., Mugno B. & Riemann B. (2017b) Co-morbid obsessive–compulsive disorder and depression: A Bayesian network approach. *Psychological Medicine* **47**:1204–14. doi:10.1017/S0033291716003287. [aDB]

McNally R. J., Robinaugh D. J., Wu G. W. Y., Wang L., Deserno M. & Borsboom D. (2015) Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science* **3**:836–49. [RJM]

Meaney M. J. (2017) Epigenetics and the biology of gene × environment interactions. In: *Gene-Environment Transactions in Developmental Psychopathology*, ed. P. H. Tolan & B. L. Leventhal, pp. 59–94. Springer. doi:10.1007/978-3-319-49227-8_4. [NMB]

Melidis C., Denham S. L. & Hyland M. E. (2018) A test of the adaptive network explanation of functional disorders using a machine learning analysis of symptoms. *Biosystems* **185**:22–30. [MEH]

Menary R., ed. (2010) *The extended mind*. MIT Press. [arDB]

Menon V. & Uddin L. Q. (2010) Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function* **214**:655–67. [DC]

Mesibov G. B., Shea V. & Schopler E. (2004) *The TEACCH approach to autism spectrum disorders*. Plenum Press. [LT]

Michaelson J. J., Shi Y., Gujral M., Zheng H., Malhotra D., Jin X., Jian M., Liu G., Greer D., Bhandari A., Wu W., Corominas R., Peoples A., Koren A., Gore A., Kang S., Lin G. N., Estabillo J., Gadomski T., Singh B., Zhang K., Akshoomoff N., Corsello C., McCarroll S., Iakoucheva L. M., Li Y., Wang J. & Sebat J. (2012) Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* **151**(7):1431–42. [CPM]

Mielenz D., Reichel M., Jia T., Quinlan E. B., Stöckl T., Mettang M., Zilske D., Kirmizi-Alsan E., Schönberger P., Praetner M., Huber S. E., Amato D., Schwarz M., Purohit P., Brachs S., Spranger J., Hess A., Büttner C., Ekici A. B., Perez-Branguli F., Winner B., Rauschenberger V., Banaschewski T., Bokde A. L., Büchel C., Conrod P. J., Desrivieres S., Flor H., Frouin V., Gallinat J., Garavan H., Gowland P., Heinz A., Martinot J. L., Lemaitre H., Nees F., Paus T., Smolka M. N., Schambony A., Bäuerle T., Eulenburg V., Alzheimer C., Lourdusamy A., Schumann G. & Müller C. P. (2018) EFhd2/Swiprosin-1 is a common genetic determinator for sensation-seeking/low anxiety and alcohol addiction. *Molecular Psychiatry* **23**(5):1303–19. doi:10.1038/mp.2017.63. [CPM]

Miller G. A. (2010) Mistreating psychology in the decades of the brain. *Perspectives on Psychological Science* **5**(6):716–43. [MIE, JH]
Miller W. R., Westerberg V. S., Harris R. J. & Tonigan J. S. (1996) What predicts relapse? Prospective testing of antecedent models. *Addiction* **91**(Suppl.):S155–72. [MF]
Millikan R. G. (1984) *Language, thought, and other biological categories*. MIT Press. [DR]
Mitte K. (2005) A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders* **88**:27–45. [rDB]
Mölder B. (2010) *Mind ascribed: An elaboration and defence of interpretivism*. John Benjamins. [aDB]
Morgan M. S. & Morrison M., eds. (1999) *Models as mediators*. Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511660108. [HMvL]
Morin E. (2005) *Introduction à la pensée complexe*. Editions du Seuil. [MD]
Morris M. W., Chiu C. & Liu Z. (2015) Polycultural psychology. *Annual Review of Psychology* **66**:631–59. [NJ]
Mucha P. J., Richardson T., Macon K., Porter M. A. & Onnela J. P. (2010) Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328** (5980):876–78. [LPes]
Müller C. P. (2018) Animal models of psychoactive drug use and addiction: Present problems and future needs for translational approaches. *Behavioural Brain Research* **352**:109–15. doi:10.1016/j.bbr.2017.06.028. [CPM]
Müller C. P. & Homberg J. (2015) The role of serotonin in drug use and addiction. *Behavioural Brain Research* **277C**:146–92. [CPM]
Müller C. P. & Huston J. P. (2006) Determining the region-specific contributions of 5-HT receptors to the psychostimulant effects of cocaine. *Trends in Pharmacological Sciences* **27**:105–12. [CPM]
Müller C. P., Kalinichenko L. S., Tiesel J., Witt M., Stöckl T., Sprenger E., Fuchser J., Beckmann J., Praetner M., Huber S. E., Amato D., Mühle C., Büttner C., Ekici A. B., Smaga I., Pomierny-Chamiolo L., Pomierny B., Filip M., Eulenburg V., Gulbins E., Lourdusamy A., Reichel M. & Kornhuber J. (2017) Paradoxical antidepressant effects of alcohol are related to acid sphingomyelinase and its control of sphingolipid homeostasis. *Acta Neuropathologica* **133**(3):463–83. [CPM]
Müller C. P. & Kornhuber J. (2017) Biological evidence for paradoxical improvement of psychiatric disorder symptoms by addictive drugs. *Trends in Pharmacological Sciences* **38**(6):501–502. [CPM]
Müller C. P. & Schumann G. (2011) Drugs as an instrument: A new framework for non-addictive psychoactive drug use. *Behavioral and Brain Sciences* **34**(6):293–347. [CPM]
Müller G. B. & Newman S. A. (2003) *Origination of organismal form: Beyond the gene in developmental and evolutionary biology*. MIT Press. [DKO]
Murphy D. (2005) *Psychiatry in the scientific image*. MIT Press. [aDB, AP]
Murray M. M., Molholm S., Michel C. M., Heslenfeld D. J., Ritter W., Javitt D. C., Schroeder C. E. & Foxe J. J. (2005) Grabbing your ear: Rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex* **15**:963–74. [AP]
Nagel E. (1961) *The structure of science: Problems in the logic of scientific explanation*. Routledge & Kegan Paul. [aDB]
National Institute of Mental Health. (2018) Research Domain Criteria (RDoC). Available at: https://www.nimh.nih.gov/research-priorities/rdoc. [ADR]
Nephew B. C., Carini L. M., Sallah S., Cotino C., Alyamani R. A. S., Pittet F., Bradburn S. & Murgatroyd C. (2017) Intergenerational accumulation of impairments in maternal behavior following postnatal social stress. *Psychoneuroendocrinology* **82**:98–106. doi:10.1016/j.psyneuen.2017.05.011. [BCN]
Nesse R. M. (2013) Tinbergen's four questions, organized: A response to Bateson and Laland. *Trends in Ecology and Evolution* **28**:681–82. [OJ]
Nestler E. J. & Hyman S. E. (2010) Animal models of neuropsychiatric disorders. *Nature Neuroscience* **13**(10):1161–69. Available at: https://doi.org/10.1038/nn.2647. [NMB]
Newman S. A. & Müller G. B. (2000) Epigenetic mechanisms of character origination. *Journal of Experimental Zoology, B: Molecular and Developmental Evolution* **288**:304–17. [DKO]
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. [aDB]
Nordenfeld L. (2007) *Rationality and compulsion*. Oxford University Press. [aDB]
Norman K. A., Polyn S. M., Detre G. J. & Haxby J. V. (2006) Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences* **10**:424–30. [AP]
Nugent A. C., Diazgranados N., Carlson P. J., Ibrahim L., Luckenbaugh D. A., Brutsche N., Herscovitch P., Drevets W. C. & Zarate Jr. C. A. (2014) Neural correlates of rapid antidepressant response to ketamine in bipolar disorder. *Bipolar Disorders* **16**:119–28. [JH]
Nutt D. J., Lingford-Hughes A., Erritzoe D. & Stokes P. R. (2015) The dopamine theory of addiction: Forty years of highs and lows. *Nature Reviews Neuroscience* **16**:305–12. [CPM]
O J. (2018a) Learned helplessness from an evolutionary mismatch perspective. In: *Encyclopedia of evolutionary psychological science*, ed. T. K. Shackelford & V. A. Weekes-Shackelford, p. 36 ff.. Springer Nature. Available at: https://doi.org/10.1007/978-3-319-16999-6_1056-1. [OJ]
O J. (2018b) Self-efficacy, animal phobias and evolutionary mismatch. In: *Encyclopedia of evolutionary psychological science*, ed. T. K. Shackelford & V. A. Weekes-Shackelford, p. 58 ff. Springer Nature. Available at: https://doi.org/10.1007/978-3-319-16999-6_1057-1. [OJ]
Ohashi K., Anderson C. M., Bolger E. A., Khan A., McGreenery C. E. & Teicher M. H. (2017) Childhood maltreatment is associated with alteration in global network fibertract architecture independent of history of depression and anxiety. *NeuroImage* **150**:50–59. doi:10.1016/j.neuroimage.2017.02.037. [BCN]
Ohayon M. M. & Roth T. (2003) Place of chronic insomnia in the course of depressive and anxiety disorders. *Journal of Psychiatric Research* **37**(1):9–15. [ADR]
Olbert C. M., Gala G. J. & Tupler L. A. (2014) Quantifying heterogeneity attributable to polythetic diagnostic criteria: Theoretical framework and empirical application. *Journal of Abnormal Psychology* **123**(2):452–62. [JH, BCN]
Oler J. A., Fox A. S., Shackman A. J. & Kalin N. H. (2016) The central nucleus of the amygdala is a critical substrate for individual differences in anxiety. In: *Living without an amygdala*, ed. D. G. Amaral & R. Adolphs, pp. 218–51. Guilford Press. [JH]
Oller D. K., Griebel U. & Warlaumont A. S. (2016) Vocal development as a guide to modeling the evolution of language. *Topics in Cognitive Science* **8**(2):382–92. [Special Issue: *New frontiers in language evolution and development*, ed. D. K. Oller, R. Dale & U. Griebel.] . doi: 10.1111/tops.12198. [DKO]
Oller D. K., Niyogi P., Gray S., Richards J. A., Gilkerson J., Xu D., Yapanel U. & Warren S. F. (2010) Automated vocal analysis of naturalistic recordings from children with autism, language delay and typical development. *Proceedings of the National Academy of Sciences USA* **107**(30):13354–59. doi:10.1073/pnas.1003882107. [DKO]
Oppenheim P. & Putnam H. (1958) Unity of science as a working hypothesis. *Minnesota Studies in the Philosophy of Science* **2**:3–36. [aDB]
Pankevich D. E., Altevogt B. M., Dunlop J., Gage F. H. & Hyman S. E. (2014) Improving and accelerating drug development for nervous system disorders. *Neuron* **84**:546–53. [JH]
Partch C. L., Green C. B. & Takahashi J. S. (2013) Molecular architecture of the mammalian circadian clock. *Trends in Cell Biology* **24**:90–99. [aDB]
Pascual-Leone A. & Hamilton R. (2001) The metamodal organization of the brain. *Progress in Brain Research* **134**:427–45. [AP]
Pasqualotto A. (2016) Multisensory integration substantiates distributed and overlapping neural networks. *Behavioral and Brain Sciences* **39**:127–28. [AP]
Pasqualotto A., Dumitru M. L. & Myachykov A. (2016) Editorial: Multisensory integration: Brain, body, and world. *Frontiers in Psychology* 6: article 2046. (Online publication). doi:10.3389/fpsyg.2015.02046. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.02046/full. [AP]
Pasqualotto A., Finucane C. M. & Newell F. N. (2013) Ambient visual information confers a context-specific, long-term benefit on memory for haptic scenes. *Cognition* **128**:363–79. [AP]
Pasqualotto A. & Proulx M. J. (2015) Two-dimensional rubber-hand illusion: The Dorian Gray hand illusion. *Multisensory Research* **28**:101–10. [AP]
Passingham R. E., Stephan K. E. & Kötter R. (2002) The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience* **3**(8):606–16. [LPes]
Patten E., Belardi K., Baranek G. T., Watson L. R., Labban J. D. & Oller D. K. (2014) Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disabilities* **44**:2413–28. [DKO]
Payne J. D. & Nadel L. (2004) Sleep, dreams, and memory consolidation: The role of the stress hormone cortisol. *Learning and Memory* **11**(6):671–78. [ADR]
Pearl J. (2000) *Causality: Models, reasoning, and inference*. Cambridge University Press. [rDB]
Perlovsky L. I. (2006) Toward physics of the mind: Concepts, emotions, consciousness, and symbols. *Physics of Life Reviews* **3**(1):22–55. [LPer]
Perlovsky L. I. (2010a) Intersections of mathematical, cognitive, and aesthetic theories of mind. *Psychology of Aesthetics, Creativity, and the Arts* **4**(1):11–17. doi:10.1037/a0018147. [LPer]
Perlovsky L. I. (2010b) Musical emotions: Functions, origin, evolution. *Physics of Life Reviews* **7**(1):2–27. [LPer]
Perlovsky L. I. (2016) Physics of the mind. *Frontiers in Systems Neuroscience* 10: article 84. [Online publication: 15 November 2016]. Available at: https://doi.org/10.3389/fnsys.2016.00084; and at: https://www.frontiersin.org/articles/10.3389/fnsys.2016.00084/full. [LPer]
Perlovsky L. I. (2017a) *Music: Passions and cognitive functions*. Academic Press. [LPer]
Perlovsky L. I. (2017b) Physics of the mind, dynamic logic, and monotone Boolean functions. In: *Uncertainty modeling*, ed. V. Kreinovich, pp. 193–231. Springer. [LPer]
Pernu T. K. (2018) Mental causation via neuroprosthetics? A critical analysis. *Synthese* **195**:5159–74. Available at: https://link.springer.com/article/10.1007/s11229-018-1713-z. [TKP]
Pernu T. K. (2019) Causal explanation in psychiatry. In: *The Bloomsbury companion to philosophy of psychiatry*, ed. Ş. Tekin & R. Bluhm. Bloomsbury Academic. [TKP]
Persons J. B. (1986) The advantages of studying psychological phenomena rather than psychiatric diagnoses. *American Psychologist* **41**(11):1252–60. [MIE]

Pessoa L. (2013) *The cognitive-emotional brain: From interactions to integration*. MIT Press. [LPes]
Pessoa L. (2014) Understanding brain networks and brain organization. *Physics of Life Reviews* **11**(3):400–35. [LPes]
Pessoa L. (2017) A network model of the emotional brain. *Trends in Cognitive Sciences* **21** (5):357–71. [LPes]
Pessoa L. (2018) Understanding emotion with brain networks. *Current Opinion in Behavioral Sciences* **19**:19–25. [LPes]
Peters A., McEwen B. S. & Friston K. (2017) Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology* **156**:164–88. [IGE]
Petry N. M. (2011) *Contingency management for substance abuse treatment*. Routledge. [ADR]
Pfenning A. R., Hara E., Whitney O., Rivas M. V., Wang R., Roulhac P. L., Howard J. T., Wirthlin M., Lovell P. V., Ganapathy G., Mountcastle J., Moseley M. A., Thompson J. W., Soderblom E. J., Iriki A., Kato M., Gilbert M. T. P., Zhang G., Bakken T., Bongaarts A., Bernard A., Lein E., Mello C. V., Hartemink A. J. & Jarvis E. D. (2014) Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* **346**(6215):1256846. Available at: https://doi.org/10.1126/science.1256846. [NMB]
Pizzagalli D. A. (2014) Depression, stress, and anhedonia: Toward a synthesis and integrated model. *The Annual Review of Clinical Psychology* **10**:393–423. [JH]
Polanyi M. (1968) Life's irreducible structure. *Science* **160**:1308–12. [DKO]
Polger T. W. (2004) Neural machinery and realization. *Philosophy of Science* **71**(5): 997–1006. [rDB]
Polger T. W. (2009) Evaluating the evidence for multiple realization. *Synthese* **167**:457–72. [rDB]
Polger T. W. & Shapiro L. A. (2016) *The multiple realization book*. Oxford University Press. [MIE, TKP]
Putnam H. W. (1967) Psychological predicates. In: *Art, mind, and religion*, ed. W. H. Capitan & D. D. Merrill. University of Pittsburgh Press. [TKP, aDB]
Pylyshyn Z. (1984) *Computation and cognition*. MIT Press. [aDB]
Quinones S. (2016) *Dreamland: The true tale of America's opiate epidemic*. Bloomsbury Press. [MF]
Rachman S. (1998) A cognitive theory of obsessions: Elaborations. *Behaviour Research and Therapy* **36**:385–401. [aDB]
Radden J. (2011) *On delusion*. Routledge. [MS]
Radua J., Ramella-Cravaro V., Ioannidis J. P. A., Reichenberg A., Phiphopthatsanee N., Amir T., Yenn Thoo H., Oliver D., Davies C., Morgan C., McGuire P., Murray R. M. & Fusar-Poli P. (2018) What causes psychosis? An umbrella review of risk and protective factors. *World Psychiatry* **17**:49–66. [JPAI]
Ramos R. T., Sassi R. B. & Piqueira J. R. C. (2011) Self-organized criticality and the predictability of human behavior. *New Ideas in Psychology* **29**(1):38–48. doi:10.1016/j.newideapsych.2009.12.001. [LT]
Ramos-Sánchez L. & Atkinson D. R. (2009) The relationships between Mexican American acculturation, cultural values, gender, and help-seeking intentions. *Journal of Counseling and Development* **87**:62–71. [DC]
Rangel A., Camerer C. & Montague P. R. (2008) A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience* **9**(7):545–56. Available at: https://www.nature.com/articles/nrn2357. [ADR]
Rasch B. & Born J. (2013) About sleep's role in memory. *Physiological Reviews* **93**(2):681–766. [ADR]
Rasmussen A., Katoni B., Keller A. S. & Wilkinson J. (2011) Posttraumatic idioms of distress among Darfur refugees: Hozun and Majnun. *Transcultural Psychiatry* **48**:392–415. [NJ]
Rasmussen A., Nguyen L., Wilkinson J., Vundla S., Raghavan S., Miller K. & Keller A. (2010) Rates and impact of trauma and current stressors among Darfuri refugees in Eastern Chad. *American Journal of Orthopsychiatry* **80**:227–36. [NJ]
Ratcliffe M. (2015) *Experiences of depression: A study in phenomenology*. Oxford University Press. [MS]
Redei E. E., Andrus B. M., Kwansy M. J., Seok J., Ho J. & Mohr D. C. (2014) 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. Available at; http://www.nature.com/articles/tp201466. [aDB]
Redish A. D. (2013) *The mind within the brain: How we make decisions and how those decisions go wrong*. Oxford University Press. [ADR]
Redish A. D. & Gordon J. A. (2016) *Computational psychiatry: New perspectives on mental illness*. MIT Press. [ADR]
Redish A. D., Jensen S. & Johnson A. (2008) A unified framework for addiction: vulnerabilities in the decision process. *Behavioral and Brain Sciences* **31**:415–87. [ADR]
Regier D. A., Narrow W. E., Clarke D. E., Kraemer H. C., Kuramoto S. J., Kuhl E. A. & Kupfer D. J. (2013) DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry* **170**:59–70. [JH, RJM]
Regier P. S. & Redish A. D. (2015) Contingency management and deliberative decision-making processes. *Frontiers in Psychiatry* **6**: article 76. (Online publication). doi: 10.3389/fpsyt.2015.00076. Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2015.00076/full. [ADR]
Reise S. P. & Waller N. G. (2009) Item response theory and clinical measurement. *Annual Review of Clinical Psychology* **5**:27–48. [aDB]
Reiss S. (1991) Expectancy model of fear, anxiety, and panic. *Clinical Psychology Review* **11**:141–53. [aDB]
Robinaugh D. J., LeBlanc N. J., Vuletich H. J. & McNally R. J. (2014) Network analysis of persistent complex bereavement disorder in conjugally bereaved adults. *Journal of Abnormal Psychology* **123**:510–22. [aDB, RJM]
Robinson T. E. & Kolb B. (2004) Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology* **47**:33–46. [CPM]
Roseboom P. H., Nanda S. A., Fox A. S., Oler J. A., Shackman A. J., Shelton S. E., Davidson R. J. & Kalin N. H. (2014) Neuropeptide Y receptor gene expression in the primate amygdala predicts anxious temperament and brain metabolism. *Biological Psychiatry* **76**:850–57. [JH]
Ross D., Sharp C., Vuchinich R. & Spurrett D. (2008) *Midbrain mutiny: The behavioral economics and neuroeconomics of disordered gambling*. MIT Press. [DR]
Ryder A. G., Ban L. M. & Chentsova-Dutton Y. E. (2011) Towards a cultural-clinical psychology. *Social and Personality Psychology Compass* **5**/12:960–75. [NJ]
Salamone J. D. (1996) The behavioral neurochemistry of motivation: Methodological and conceptual issues in studies of the dynamic activity of nucleus accumbens dopamine. *Journal of Neuroscience Methods* **64**:137–49. [CPM]
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**:347–72. [aDB]
Samek D. R., Hicks B. M., Keyes M. A., Iacono W. G. & McGue M. (2017) Antisocial peer affiliation and externalizing disorders: Evidence for Gene × Environment × Development interaction. *Development and Psychopathology* **29**(1):155–72. doi:10.1017/s0954579416000109. [BCN]
Santos Jr. H., Fried E. I., Asafu-Adjei J. & Ruiz R. J. (2017) Network structure of perinatal depressive symptoms in Latinas: Relationship to stress and reproductive biomarkers. *Research in Nursing and Health* **40**(3):218–28. [BCN]
Sarkar S. (1992) Models of reduction and categories of reductionism. *Synthese*, **91**:167–94. [KT]
Satel S. & Lilienfeld S. (2014) Addiction and the brain-disease fallacy. *Frontiers in Psychiatry* **4**: article 141. (Online publication). doi: 10.3389/fpsyt.2013.00141. Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2013.00141/full. [MF]
Satel S. & Lilienfeld S. (2017) Calling it "brain disease" makes addiction harder to treat. *Boston Globe,* June 22, 2017. Available at: https://www.bostonglobe.com/ideas/2017/06/22/calling-brain-disease-makes-addiction-harder-treat/ehaJs5ZYIXpPottG89KOGK/story.html. [MF]
Schaffner K. (1974) Reductionism in biology: Prospects and problems. *Proceedings of the Biennial Meeting of the Philosophy of Science Association* **32**:613–32. [aDB]
Schaffner K. (2006) Reduction: The Cheshire cat problem and a return to roots. *Synthese* **151**:377–402. [aDB, KT]
Schaffner K. F. (2016) *Behaving: What's genetic, what's not, and why should we care?* Oxford University Press. [KT]
Schaffner K. F. & Tabb K. (2014) Varieties of social constructionism and the problem of progress in psychiatry. In: *Philosophical issues in psychiatry III: The nature and sources of historical change*. Oxford University Press. [aDB]
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. [aDB]
Scheffer M., Carpenter S. R., Foley J. A., Folke C. & Walker B. (2001) Catastrophic shifts in ecosystems. *Nature* **413**:591–96. [rDB]
Schlaepfer T. E., Cohen M. X., Frick C., Kosel M., Brodesser D., Axmacher N., Joe A. Y., Kreft M., Lenartz D. & Sturm V. (2008) Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. *Neuropsychopharmacology* **33**:368–77. [JH]
Schneider M., Levant B., Reichel M., Gulbins E., Kornhuber J. & Müller C. P. (2017) Lipids in psychiatric disorders and preventive medicine. *Neuroscience and Biobehavioral Reviews* **76**:336–62. [CPM]
Schoeller F., Perlovsky L. I. & Arseniev D. (2018) Physics of the mind: Experimental confirmations of theoretical predictions. *Physics of Life Reviews* **25**:45–68. [LPer]
Schuckit M. A., Tipp J. E., Bergman M., Reich W., Hesselbrock V. M. & Smith T. L. (1997) Comparison of induced and independent major depressive disorders in 2,945 alcoholics. *American Journal of Psychiatry* **154**:948–57. [CPM]
Schüll N. D. (2012) *Addiction by design*. Princeton University Press. [DR]
Schumann G., Binder E. B., Holte A., de Kloet E. R., Oedegaard K. J., Robbins T. W., Walker-Tilley T. R., Bitter I., Brown V. J., Buitelaar J., Ciccocioppo R., Cools R., Escera C., Fleischhacker W., Flor H., Frith C. D., Heinz A., Johnsen E., Kirschbaum C., Klingberg T., Lesch K. P., Lewis S., Maier W., Mann K., Martinot J. L., Meyer-Lindenberg A., Müller C. P., Müller W. E., Nutt D. J., Persico A., Perugi G., Pessiglione M., Preuss U. W., Roiser J. P., Rossini P. M., Rybakowski J. K., Sandi C.,

Stephan K. E., Undurraga J., Vieta E., van der Wee N., Wykes T., Haro J. M. & Wittchen H. U. (2014) Stratified medicine for mental disorders. *European Neuropsychopharmacology* **24**(1):5–50. doi:10.1016/j.euroneuro.2013.09.010. [IGE, CPM]

Schumann G., Liu C., O'Reilly P., Gao H., Song P., Xu B., Ruggeri B., Amin N., Jia T., Preis S., Segura L. M., Akira S., Barbieri C., Baumeister S., Cauchi S., Clarke T. K., Enroth S., Fischer K., Hallfors J., Harris S. E., Hieber S., Hofer E., Hottenga J. J., Johansson A., Joshi P. K., Kaartinen N., Laitinen J., Lemaitre R., Loukola A., Luan J., Lyytikainen L. P., Mangino M., Manichaikul A., Mbarek H., Milaneschi Y., Moayyeri A., Mukamal K., Nelson C., Nettleton J., Partinen E., Rawal R., Robino A., Rose L., Sala C., Satoh T., Schmidt R., Schraut K., Scott R., Smith A. V., Starr J. M., Teumer A., Trompet S., Uitterlinden A. G., Venturini C., Vergnaud A. C., Verweij N., Vitart V., Vuckovic D., Wedenoja J., Yengo L., Yu B., Zhang W., Zhao J. H., Boomsma D. I., Chambers J., Chasman D. I., Daniela T., de G. E., Deary I., Eriksson J. G., Esko T., Eulenburg V., Franco O. H., Froguel P., Gieger C., Grabe H. J., Gudnason V., Gyllensten U., Harris T. B., Hartikainen A. L., Heath A. C., Hocking L., Hofman A., Huth C., Jarvelin M. R., Jukema J. W., Kaprio J., Kooner J. S., Kutalik Z., Lahti J., Langenberg C., Lehtimaki T., Liu Y., Madden P. A., Martin N., Morrison A., Penninx B., Pirastu N., Psaty B., Raitakari O., Ridker P., Rose R., Rotter J. I., Samani N. J., Schmidt H., Spector T. D., Stott D., Strachan D., Tzoulaki I., van der Harst P., van Duijn C. M., Marques-Vidal P., Vollenweider P., Wareham N. J., Whitfield J. B., Wilson J., Wolffenbuttel B., Bakalkin G., Evangelou E., Liu Y., Rice K. M., Desrivieres S., Kliewer S. A., Mangelsdorf D. J., Müller C. P., Levy D. & Elliott P. (2016) KLB is associated with alcohol drinking, and its gene product beta-Klotho is necessary for FGF21 regulation of alcohol preference. *Proceedings of the National Academy of Sciences USA* **113**:14372–77. [CPM]

Searle J. R. (1983) *Intentionality*. Cambridge University Press. [aDB]

Searle J. R. (1998) *Mind language and society: Philosophy in the real world*. Basic Books. [aDB]

Seibold B. (2015) A mathematical introduction to traffic flow theory. *IPAM Tutorials*. Available at: http://helper.ipam.ucla.edu/publications/tratut/tratut_12985.pdf. [ADR]

Shackman A. J. & Fox A. S. (2018) Getting serious about variation: Lessons for clinical neuroscience. *Trends in Cognitive Sciences* **22**:368–69. [JH]

Shackman A. J., Fox A. S., Oler J. A., Shelton S. E., Davidson R. J. & Kalin N. H. (2013) Neural mechanisms underlying heterogeneity in the presentation of anxious temperament. *Proceedings of the National Academy of Sciences USA* **110**:6145–50. [JH]

Shackman A. J., Kaplan C. M., Stockbridge M. D., Tillman R. M., Tromp D. P. M., Fox A. S. & Gamer M. (2016a) The neurobiology of anxiety and attentional biases to threat: Implications for understanding anxiety disorders in adults and youth. *Journal of Experimental Psychopathology* **7**:311–42. [JH]

Shackman A. J., Tromp D. P. M., Stockbridge M. D., Kaplan C. M., Tillman R. M. & Fox A. S. (2016b) Dispositional negativity: An integrative psychological and neurobiological perspective. *Psychological Bulletin* **142**:1275–314. [JH]

Shapiro L. A. (2000) Multiple realizations. *Journal of Philosophy* **97**:635–54. [TKP]

Shay J. (1994) *Achilles in Vietnam: Combat trauma and the undoing of character*. Simon and Schuster. [ADR]

Skene N. G., Roy M. & Grant S. G. (2017) A genomic lifespan program that reorganizes the young adult brain is targeted in schizophrenia. *eLife* 1–30. (Online research article). eLife 2017;**6**:e17915 doi: 10.7554/eLife.17915. Available at: https://elifesciences.org/articles/17915. [DKO]

Skrondal A. & Rabe-Hesketh S. (2004) *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Chapman & Hall/CRC Press. [NJ]

Slors M. V. P. (2007) Intentional systems theory, mental causation and empathic resonance. *Erkenntnis* **67**:321–36. Available at: https://doi:org/10.1007/s10670-007-9074-x. [MS]

Snippe E., Viechtbauer W., Geschwind N., Klippel A., de Jonge P. & Wichers M. (2017) The impact of treatments for depression on the dynamic network structure of mental states: Two randomized controlled trials. *Scientific Reports* **7**: article 46523. (Online publication). Available at: https://www.nature.com/articles/srep46523. [rDB]

Solomon A. (2014) *The noonday demon*. Scribner. [aDB]

Spanagel R. (2009) Alcoholism: A systems approach from molecular physiology to addictive behavior. *Physiological Reviews* **89**:649–705. [CPM]

Spanagel R. & Kiefer F. (2008) Drugs for relapse prevention of alcoholism: Ten years of progress. *Trends in Pharmacological Sciences* **29**:109–15. [CPM]

Spencer H. (1864–1867) Adaptation. In: *The principles of biology, vol. 1*, Part II, Ch. V, p. 184-200. Williams and Norgate. [Book available in pdf at: https://gallica.bnf.fr/ark:/12148/bpt6k779237/f2.image.r=The%20principles%20of%20biology] [MD]

Stark K. H., Barnes J. C., Young N. D. & Gabriels R. L. (2015) Brief Report: Understanding crisis behaviors in hospitalized psychiatric patients with autism spectrum disorder – Iceberg assessment interview. *Journal of Autism and Developmental Disorders* **45**(11):3468–74. doi:10.1007/s10803-015-2552-0. [LT]

Stead L. F., Perera R., Bullen C., Mant D., Hartmann-Boyce J., Cahill K. & Lancaster T. (2012) Nicotine replacement therapy for smoking cessation. *Cochrane Database of Systematic Reviews* **12**:CD000146. doi: 10.1002/14651858.CD000146.pub4. [MF]

Stearns S. C. (2000) Life history evolution: Successes, limitations, and prospects. *Naturwissenschaften* **87**:476–86. [OJ]

Stein B. E. (2012) *The new handbook of multisensory processing*. MIT Press. [AP]

Steinberg L. (2008) A social neuroscience perspective on adolescent risk-taking. *Developmental Review* **28**:78–106. [AP]

Stephan K. E., Manjaly Z. M., Mathys C. D., Weber L. A. E., Paliwal S., Gard T., Tittgemeyer M., Fleming S. M., Haker H., Seth A. K. & Petzschner F. H. (2016) Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience* **10**. doi:10.3389/fnhum.2016.00550. Online article, available at: https://www.frontiersin.org/articles/10.3389/fnhum.2016.00550/full. [IGE]

Stone J., Wojcik W., Durrance D., Carson A., Lewis S., MacKenzie L., Warlow C. P. & Sharpe M. (2002) What should we say to patients with symptoms unexplained by disease? The "number needed to offend". *British Medical Journal* **325**(7378): 1449–50. [MEH]

Striedter G. F., Belgard T. G., Chen C.-C., Davis F. P., Finlay B. L., Güntürkün O., Hale M. E., Harris J. A., Hecht E. E., Hof P. R., Hofmann H. A. K., Holland L. Z., Iwaniuk A. N., Jarvis E. D., Karten H. J., Katz P. S., Kristan W. B., Macagno E. R., Mitra P. P., Moroz L. L., Preuss T. M., Ragsdale C. W., Sherwood C. C., Stevens C. F., Stüttgen M. C., Tsumoto T. & Wilczynski W. (2014) NSF workshop report: Discovering general principles of nervous system organization by comparing brain maps across species. *Brain, Behavior and Evolution* **83**(1):1–8. doi:10.1159/000360152. [NMB]

Stringaris A., Vidal-Ribas Belil P., Artiges E., Lemaitre H., Gollier-Briant F., Wolke S., Vulser H., Miranda R., Penttila J., Struve M., Fadai T., Kappel V., Grimmer Y., Goodman R., Poustka L., Conrod P., Cattrell A., Banaschewski T., Bokde A. L., Bromberg U., Buchel C., Flor H., Frouin V., Gallinat J., Garavan H., Gowland P., Heinz A., Ittermann B., Nees F., Papadopoulos D., Paus T., Smolka M. N., Walter H., Whelan R., Martinot J. L., Schumann G., Paillere-Martinot M. L. & the IMAGEN Consortium (2015) The brain's response to reward anticipation and depression in adolescence: Dimensionality, specificity, and longitudinal predictions in a community-based sample. *American Journal of Psychiatry* **172**:1215–23. [JH]

Sturmberg J. P., Bennett J. M., Martin C. M. & Picard M. (2017) "Multimorbidity" as the manifestation of network disturbances. *Journal of Evaluation in Clinical Practice* **23** (1):199–208. [MEH]

Suo X., Lei D., Li K., Chen F., Li F., Li L., Huang X., Lui S., Li L., Kemp G. J. & Gong Q. (2015) Disrupted brain network topology in pediatric posttraumatic stress disorder: A resting-state fMRI study. *Human Brain Mapping* **36**(9):3677–86. doi:10.1002/hbm.22871. [BCN]

Swedo S. E., Rapoport J. L., Leonard H., Lenane M. & Cheslow D. (1989) Obsessive–compulsive disorder in children and adolescents: Clinical phenomenology of 70 consecutive cases. *Archives of General Psychiatry* **46**:335–41. [ADR]

Tabb K. (2015) Psychiatric progress and the assumption of diagnostic discrimination. *Philosophy of Science* **82**(5):1047–58. [KT]

Tabb K. & Schaffner K. F. (2017) Causal pathways, random walks and tortuous paths: Moving from the descriptive to the etiological in psychiatry. In: *Philosophical issues in psychiatry IV: Nosology*, ed. K. S. Kendler & J. Parnas, pp. 342–60. Oxford University Press. [aDB]

Talarico J. M. & Rubin D. C. (2003) Confidence, not consistency, characterizes flashbulb memories. *Psychological Science* **14**(5):455–61. [ADR]

Taylor M. A. (1981) *The neuropsychiatric mental status examination*. Spectrum. [RJM]

Thase M. E. (2014) Using biomarkers to predict treatment response in major depressive disorder: Evidence from past and present studies. *Dialogues in Clinical Neuroscience* **16**:539–44. [JPAI]

Thompson E. (2007) *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press. [LPes]

Thornton T. (2010) Psychiatric explanation and understanding. *European Journal of Analytic Philosophy* **6**:95–111. [aDB]

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. [aDB]

Tomm K. (1984) One perspective on the Milan systemic approach: Part I. Overview of development, theory and practice. *Journal of Marital and Family Therapy* **10**: 113–25. [AP]

Tonello L., Giacobbi L., Pettenon A., Scuotto A., Cocchi M., Gabrielli F. & Cappello G. (2018) Crisis behavior in autism spectrum disorders: A self-organized criticality approach. *Complexity* **2018**: Article ID 5128157. doi:10.1155/2018/5128157. (Published Online January 31, 2018). Available at: https://www.hindawi.com/journals/complexity/2018/5128157/. [LT]

Troisi A. (2017) *The painted mind. Behavioral science reflected in great paintings*. Oxford University Press. [AT]

Turkheimer E. (1998) Heritability and biological explanation. *Psychological Review* **105**:782–91. [JH]

U.S. Burden of Disease Collaborators (2018) The state of US health, 1990–2016. Burden of diseases, injuries, and risk factors among US states. *Journal of the American Medical Association (JAMA)* **319**:1444–72. [JH]

Uesaki M., Takemura H. & Ashida H. (2018) Computational neuroanatomy of human stratum proprium of interparietal sulcus. *Brain Structure and Function* **223**: 489–507. [AP]

Ungless M. A., Argilli E. & Bonci A. (2010) Effects of stress and aversion on dopamine neurons: Implications for addiction. *Neuroscience and Biobehavioral Reviews* **35**: 151–56. [CPM]

Van Bork R., Epskamp S., Rhemtulla M., Borsboom D. & van der Maas H. L. J. (2017) What is the *p*-factor of psychopathology? Some risks of general factor modelling. *Theory and Psychology* **27**:759–73. [RJM]

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**: article 5918. Available at: https://www.nature.com/articles/srep05918. [arDB]

Van Borkulo C. D., Boschloo L., Borsboom D., Pennix B. W. J. H., Waldorp L. J. & Schoevers R. A. (2015) Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry* **72**:1219–26. [RJM]

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 USA* **111**:87–92. doi:10.1073/pnas.1312114110. [aDB]

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**:842–61. [aDB]

Van Loo H. M., De Jonge P., Romeijn J.-W., Kessler R. C. & Schoevers R. A. (2012) Data-driven subtypes of major depressive disorder: A systematic review. *BMC Medicine* **10**(1):156. Available at: https://doi.org/10.1186/1741-7015-10-156. [HMvL]

Van Loo H. M. & Romeijn J.-W. (2015) Psychiatric comorbidity: Fact or artifact? *Theoretical Medicine and Bioethics* **36**(1):41–60. [HMvL]

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**:363–72. [aDB]

Varela F. J. & Shear J. (1999) First-person methodologies: What, why, how. *Journal of Consciousness Studies* **6**:1–14. [DC]

Varela F. J., Thompson E. & Rosch E. (1991) *The embodied mind: Cognitive science and human experience*. MIT Press. [LPes]

Varela F., Thompson E. & Rosch E. (1991) *The embodied mind: Cognitive science and human experience*. MIT Press. [MD]

Varghese M., Keshav N., Jacot-Descombes S., Warda T., Wicinski B., Dickstein D. L., Harony-Nicolas H., de Rubeis S., Drapeau E., Buxbaum J. D. & Hof P. R. (2017) Autism spectrum disorder: Neuropathology and animal models. *Acta Neuropathologica* **134**(4):537–66. doi: 10.1007/s00401-017-1736-4. [NMB]

Volkow N. D., Koob G. F. & McLellan A. T. (2016) Neurobiologic advances from the brain disease model of addiction. *New England Journal of Medicine* **374**(4): 363–71. [MF]

Wan X., Zhou X., Woods A. T. & Spence C. (2015) Influence of the glassware on the perception of alcoholic drinks. *Food Quality and Preference* **44**:101–10. [AP]

Warlaumont A. S., Richards J. A., Gilkerson J. & Oller D. K. (2014) A social feedback loop for speech development and its reduction in autism. *Psychological Science* **25**(7):1314–24. doi: 10.1177/0956797614531023. [DKO]

Waszczuk M. A., Zimmerman M., Ruggero C., Li K., MacNamara A., Weinberg A., Hajcak G., Watson D. & Kotov R. (2017) What do clinicians treat: Diagnoses or symptoms? The incremental validity of a symptom-based, dimensional characterization of emotional disorders in predicting medication prescription patterns. *Comprehensive Psychiatry* **79**:80–88. [JH]

Waterhouse L., London E. & Gillberg C. (2016) ASD validity. *Review Journal of Autism and Developmental Disorders* **3**(4):302–29. Available at: https://doi.org/10.1007/s40489-016-0085-x. [KH]

Watson D. & Stasik S. M. (2014) Examining the comorbidity between depression and the anxiety disorders from the perspective of the quadripartite model. In: *Oxford handbook of depression and comorbidity*, ed. C. S. Richards & M. W. O'Hara, pp. 46–65. Oxford University Press. [JH]

Wei H. (2017) Construction of a hierarchical gene regulatory network centered around a transcription factor. *Brief Bioinform* **27**:e4662947. doi: 10.1093/bib/bbx152. Available at: https://www.researchgate.net/publication/321318939_Construction_of_a_hierarchical_gene_regulatory_network_centered_around_a_transcription_factor. [BCN]

Wessely S., Nimnuan C. & Sharpe M. (1999) Functional somatic syndromes: One or many? *The Lancet* **354**(9182):936–39. [MEH]

Westerink B. H. (1995) Brain microdialysis and its application for the study of animal behaviour. *Behavioural Brain Research* **70**:103–24. [CPM]

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**:147–57. [aDB]

Whitfield P. (2009) Shakespeare, pedagogy and dyslexia. *Voice and Speech Review* **6**: 254–62. [AP]

Whitfield P. (2016) A facilitation of dyslexia through a remediation of Shakespeare's text. *Research in Drama Education: The Journal of Applied Theatre and Performance* **21**:385–400. [AP]

Whitfield P. (2017) The micro grasp and macro gestus strategy as a facilitation of dyslexia in actor-training: Reconstructing the written text when performing Shakespeare. *Theatre, Dance and Performance Training* **8**:329–47. [AP]

Wichers M. (2014) The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine* **44**(7): 1349–60. [aDB]

Wichers M., Groot P. C., Psychosystems, ESM Group, & EWS Group. (2016) Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics* **85**:114–16. [rDB]

Wichers M., Wigman J. T. W., Bringmann L. F. & de Jonge P. (2017) Mental disorders as networks: Some cautionary reflections on a promising approach. *Social Psychiatry and Psychiatric Epidemiology* **52**(2):143–45. Available at: https://doi.org/10.1007/s00127-016-1335-z. [HMvL]

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**:611–24. [aDB]

Williams M. J. & Adinoff B. (2008) The role of acetylcholine in cocaine addiction. *Neuropsychopharmacology* **33**:1779–97. [CPM]

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

Wise R. A. (2002) Brain reward circuitry: Insights from unsensed incentives. *Neuron* **36**:229–40. [CPM]

Wittenborn A. K., Rahmandad H., Rick J. & Hosseinichimeh N. (2016) Depression as a systemic syndrome: Mapping the feedback loops of major depressive disorder. *Psychological Medicine* **46**(3):551–62. doi:10.1017/s0033291715002044. [BCN]

Wolfram S. (2002) *A new kind of science*. Wolfram Media. [MEH]

Woo C. W., Chang L. J., Lindquist M. A. & Wager T. D. (2017) Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience* **20**:365–77. [JH]

Woodward J. (2003) *Making things happen*. Oxford University Press. [rDB]

Woodward N. D. & Cascio C. J. (2015) Resting-state functional connectivity in psychiatric disorders. *JAMA Psychiatry* **72**(8):743–44. doi:10.1001/jamapsychiatry.2015.0484. [BCN]

Woolgar A., Jackson J. & Duncan J. (2016) Coding of visual, auditory, rule and response information in the brain: Ten years of multivoxel pattern analysis. *Journal of Cognitive Neuroscience* **28**:1433–54. [AP]

World Health Organization (2010) *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization. Available at: www.who.int/classifications/icd/en/bluebook.pdf. [LT]

Wray N. R., Ripke S., Mattheisen M., Trzaskowski M., Byrne E. M., Abdellaoui A., *et al.* (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nature Genetics 50:668–81. [rDB]

Wright A. G. C. (2017) Factor analytic support for the five-factor model. In: *The Oxford handbook of the five factor model of personality*, ed. T. A. Widiger, pp. 217–42. Oxford University Press. [NJ]

Wright A. G. C., Hopwood C. J. & Simms L. J. (2015) Daily interpersonal and affective dynamics in personality disorders. *Journal of Personality Disorders* **29**:503–29. [RFB]

Yun R. J., Stern B. L., Lenzenweger M. F. & Tiersky L. A. (2013) Refining personality disorder subtypes and classification using finite mixture modeling. *Personality Disorders: Theory, Research, and Treatment* **4**:121–28. [RFB]

Zachar P. & Kendler K. S. (2017) The philosophy of nosology. *Annual Review of Clinical Psychology* **13**:49–71. doi:10.1080/09608789908571020. [IGE]

Zangaladze A., Epstein C. M., Grafton S. T. & Sathian K. (1999) Involvement of visual cortex in tactile discrimination of orientation. *Nature* **401**:587–90. [AP]

Zhang R., Geng X. & Lee T. M. C. (2017) Large-scale functional neural network correlates of response inhibition: An fMRI meta-analysis. *Brain Structure and Function* **222** (9):3973–90. doi:10.1007/s00429-017-1443-x. [BCN]

Zhu X., Need A. C., Petrovski S. & Goldstein D. B. (2014) One gene, many neuropsychiatric disorders: Lessons from Mendelian diseases. *Nature Neuroscience* **17**:773–81. [JH]

Zinberg N. E. (1984) *Drug, set, and setting: The basis for controlled intoxicant use*. Yale University Press. [CPM]

Zoorob M. & Salemi J. (2017) Bowling alone, dying together: The role of social capital in mitigating the drug overdose epidemic in the United States. *Drug and Alcohol Dependence* **173**:1–9. [MF]