New Zealand Sociology Special Issue: Theorising Mental Health
2023, Issue 38, No. 1, 41–55

The Network Approach to Major Depressive Disorder: A Critical Realist Perspective

Lennox Johnson*

Abstract
Lack of progress in treating major depressive disorder (MDD) along with critiques of the reductionist biomedical model have motivated psychiatric researchers to develop alternative models of mental disorders. Over the last decade, network models have been developed that conceptualise mental disorders as the product of dynamic causal interactions between psychiatric symptoms, especially when these symptoms form self-reinforcing feedback loops. This article discusses how problems with the biomedical model have motivated research into network models of MDD. The network approach is then compared with critical realist (CR) theorising on MDD within the sociology of mental health. This article argues that critical realism and network approaches overlap significantly in their ontological and epistemological assumptions about mental disorders. Therefore, CR and network models are potentially complementary approaches for sociological theorising about MDD. The final section of this article discusses tensions between these two approaches and limitations of current network models of MDD from a CR perspective.

Keywords: critical realism; depression; mental disorders; network model

Introduction

According to the World Health Organization, major depressive disorder (MDD) is the leading cause of disability worldwide, affecting approximately five per cent of the global adult population (World Health Organization, 2021). Despite decades of research, progress in understanding the aetiology of MDD or in developing successful treatments has been slow. The global disease burden of MDD has increased over the past three decades (Jorm et al., 2017). Researchers have been unable to identify biomarkers that are useful in clinical settings, and the effectiveness of both pharmacological and psychological interventions remain limited (Scull, 2021). Critical psychiatrists and sociologists have argued that the lack of progress in mental health research is because the dominant model for understanding MDD, the biomedical model, is deeply flawed. Increasingly, these critiques are also coming from within academic psychiatry itself. This lack of progress has motivated research into alternative theories and models of mental disorders, such as network models of mental disorders (Borsboom, 2008).

Instead of viewing MDD as the result of a single underlying biological cause, the network approach focuses on individual symptoms associated with depression and the causal relationships between symptoms (Borsboom, 2017b). This article will analyse the network approach to depression from a critical realist (CR) perspective (Pilgrim & Bentall, 1999). It will start by describing the biomedical model of depression and the social context in which the modern diagnostic criteria for depression were established. Then it will discuss criticisms of the biomedical model and explain the CR perspective on depression. Next, it will introduce the network model of depression, compare the philosophical assumptions underlying both the network approach and CR, and argue that the network model of depression is, for the most part, compatible with a CR perspective. However, the network approach is limited by its overemphasis on biological factors and lack of reflexivity. A CR perspective can provide a useful theoretical framework for addressing the

* **Lennox Johnson** is a doctoral student in Sociology at Waipapa Taumata Rau, University of Auckland. His research interests include the sociology of mental health and the sociology of disasters, particularly when viewed through the lens of critical realism and complexity theory.
Email: ljoh674@aucklanduni.ac.nz

limitations of the network approach, with a combination of the network approach and a CR perspective providing an ontologically sophisticated and empirically grounded theory of depression.

The biomedical model of depression

According to the American Psychiatric Association (APA), depression "is a common and serious medical illness that negatively affects how you feel, the way you think and how you act" (American Psychiatric Association, n.d., para. 1). The fifth edition of the APA's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) is the principal tool for diagnosing depression within psychiatry (see American Psychiatric Association, 2013). An individual meets the diagnostic criteria for depression if they experience any five of the following symptoms for a period of at least two weeks, and one of the symptoms includes either: (1) depressed mood, or (2) loss of interest or pleasure (American Psychiatric Association, 2013, pp.160–161):

1. Depressed mood most of the day, nearly every day.
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day.
3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.
4. Insomnia or hypersomnia nearly every day.
5. Psychomotor agitation or retardation nearly every day.
6. Fatigue or loss of energy nearly every day.
7. Feelings of worthlessness or excessive or inappropriate guilt.
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day.
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

These symptoms must cause clinically significant distress or impairment, must not result from substance abuse or another medical condition, and must not be an expected or socially acceptable response to unexpected loss (that is, grief at the loss of a loved one).

The identification of depression as a medical illness illustrates how psychiatry primarily uses a biomedical model to understand the aetiology of depression. The *biomedical model* views the aetiology of mental illness as ultimately attributable to harmful variations in physiological or biochemical functioning within the brain (Strickland & Patrick, 2014). These harmful variations include genetic vulnerabilities to mental disorders, deficiencies in specific brain circuits, or other neurobiological mechanisms. Social factors may interact with these harmful biochemical variations, but the ultimate cause of most mental disorders is understood in biological terms. The dominance of the biomedical model of mental disorders within psychiatry has had a profound impact on the conceptualisation, diagnosis and treatment of medical disorders (Fried, 2015). According to the biomedical model, psychiatric research should seek to identify the structural variations within the brain that cause mental illness. Once a harmful variation is identified, treatment should focus on correcting the latent biological cause of the disease. When this latent cause is corrected, the patient will return to normal health.

According to Borsboom et al. (2019), explanatory reductionism in the context of mental health research is the view that mental disorders can ultimately be explained in terms of biology. Borsboom et al. (2019) argue that explanatory reductionism is endorsed by some of the most authoritative psychiatry researchers. For example, Insel and Cuthbert (2015, p. 500), former leaders of the National Institute of Mental Health, declared that "as new diagnostics will likely be redefining 'mental disorders' as 'brain circuit

disorders,' new therapeutics will likely focus on tuning these circuits." Borsboom et al. (2019) cite several other instances in prestigious journals that explicitly endorse explanatory reductionism in the context of mental health. They conclude that:

> 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. (Borsboom et al., 2019, p. 3)

The biomedical model can therefore be understood as a form of explanatory reductionism. It aims to explain mental disorders in terms of variations in brain structures or biochemical processes and to use standard scientific methods used in medicine to establish causal relationships between brain structures, biochemistry and mental illness. While the DSM and the biomedical model have been dominant within psychiatry since the publication of the DSM-III (American Psychiatric Association, 1980), both have been heavily criticised as an inadequate foundation for understanding and treating depression.

The social construction of depression

One approach that has challenged the dominance of the DSM and the biological model of mental illness is social constructionism. *Social constructionism* is a theoretical tradition within mental health research that emphasises that concepts such as depression are the product of a specific socio-historical context rather than an objective and neutral description of reality (White, 2017). Allan Horwitz (2011) has examined the social context that changed depression from a serious and rare condition to the most commonly diagnosed mental illness. This occurred shortly after the publication of the third edition of the DSM (DSM-III) (American Psychiatric Association, 1980). Depression now represents approximately 40 per cent of all psychiatric diagnoses (Horwitz, 2011).

Horwitz notes that psychiatry was facing a crisis of legitimacy in the decades leading up to 1980. Psychiatric diagnoses based on the criteria in the second edition of the DSM (DSM-II) (American Psychiatric Association, 1968) were criticised as unreliable and pseudo-scientific, and psychiatry faced intense competition from other professions offering treatment for mental distress, such as clinical psychologists, counsellors and social workers. Robert Spitzer, the APA-appointed task force head responsible for revising the DSM-II, sought to improve the inter-rater reliability of diagnoses in the DSM-II by creating a classification system based purely on the symptoms produced by each disorder and by purging underlying psychodynamic explanations from diagnostic criteria.

In the case of depression, the DSM-III largely adopted the Feignher criteria, with minor modifications. The *Feighner criteria for depression* required three conditions to be met for diagnosis. First, the individual must have a dysphoric mood as indicated by feelings of sadness or hopelessness. Second, the individual must also have at least five of the following symptoms: poor appetite or weight loss, sleep difficulties, loss of energy, agitation or retardation, loss of interest in usual activities, feelings of guilt, difficulty concentrating, and suicidal thoughts. And third, these symptoms must have lasted for at least one month and must not have been caused by other psychiatric or medical conditions (Feighner et al., 1972). Horwitz notes that the Feignher criteria was only one of many possible diagnostic schemes for depression. The adoption of the Feignher criteria was largely the result of close professional relationships between committee members on the DSM-III task force, who represented only one wing of the psychiatric profession and who were committed to legitimising psychiatry as a medical discipline. Horwitz (2011, p. 42) argues that the DSM-III definition of depression:

> …was not the result of the development of a well-established body of research but instead was grounded in attempts by one wing of the psychiatric profession to gain professional dominance and scientific legitimacy and to distinguish itself from professional competitors.

Diagnostic criteria for depression in earlier editions of the DSM treated it as a rare and serious condition. The choice of Feighner criteria in the DSM-III transformed depression into an extremely common condition. Changes to the diagnostic criteria had dramatic implications for mental health practice, and pharmaceutical companies, in particular, capitalised on the expanded definition to sell a new generation of antidepressant drugs: selective serotonin reuptake inhibitors (SSRIs). Horwitz (2011) concludes that it was ultimately rather esoteric changes in diagnostic criteria resulting from intra-professional competition that ushered in a dramatic increase in rates of depression over the last four decades.

Philosophical and empirical critiques of the biomedical model

The biomedical model and the DSM definition of depression have also been subject to philosophical and empirical critiques. Increasingly, these critiques are coming from within psychiatric research itself. Eiko Fried (2015) argues that two common assumptions about depression are not supported by the empirical evidence: first, the assumption that depression is a discrete condition, and second, that symptoms of depression are interchangeable and reflect an underlying latent disorder. Fried (2015) notes that research into depression has progressed remarkably slowly over recent decades, and this slow progress is, to a considerable extent, due to research being based on these two false assumptions.

Fried (2015) argues that depression is not a distinct disease category, like tuberculosis or pneumonia. One strong indication of this is that current research practice combines an extraordinarily diverse set of psychiatric symptoms into one undifferentiated category. A diagnosis of depression can result from a large variety of symptom combinations, and it is possible for two people to be diagnosed with depression while having almost no symptoms in common. In an earlier study, Fried and Nesse (2015) found that in a sample of 3703 individuals diagnosed with MDD, there were 1030 unique symptom combinations, averaging just 3.6 patients per unique combination. The most common combination of symptoms occurred in just 1.8 per cent of patients (Fried and Nesse, 2015). Furthermore, the symptoms associated with depression, such as insomnia, sad mood and appetite loss, are themselves heterogeneous and differ from each other in their underlying aetiology.

Fried (2015) argues that the use of sum-scores and threshold values to assign individuals to the category of MDD has reinforced the notion of MDD as a discrete category rather than a diverse, dimensional and heterogeneous set of symptoms. Fried (2015, p. 3) concludes that:

> The quest for biomarkers and more efficacious treatment has been disappointing at best. This lack of progress is partly because the definition of MD [major depressive disorder] as disease entity has encouraged lack of attention to specific symptoms and their dynamic interactions.

Fried (2015) also argues that MDD is not a latent condition that independently causes each of its symptoms. The latent variable model of depression is shown in Figure 1. (See also Fried (2020) for an up-to-date discussion of latent models of mental disorders.) D represents depression as a latent condition that is the common cause of symptoms s1–s14. This common cause framework views each symptom as roughly equally connected to the underlying cause. On this view, symptoms are all equally important and interchangeable. The use of sum-scores for depression diagnoses, for the most part, makes this assumption, albeit two symptoms are given hierarchical priority. This assumption is intuitive and useful when describing

certain kinds of medical condition, such as a bacterial infection; however, the model is deeply flawed when applied to mental disorders.

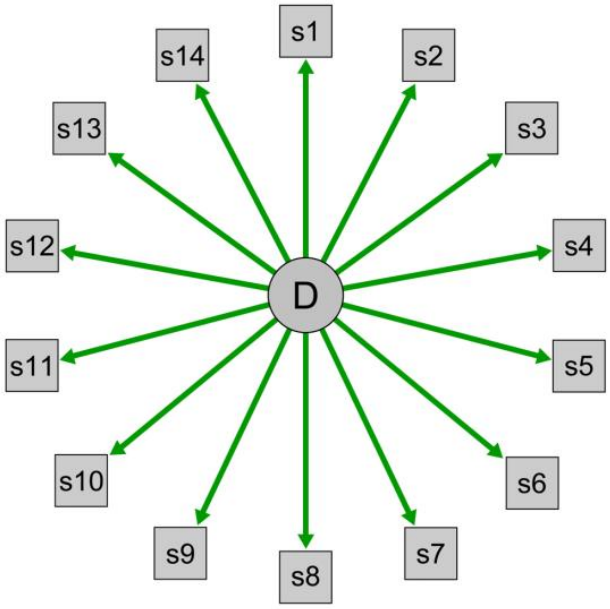

Figure 1. The latent variable model of depression.

Note: An underlying cause, D, independently causes symptoms s1–s14. Symptoms do not causally interact with one another.

The common cause model assumes that symptoms are locally independent; that is, symptoms such as insomnia and concentration difficulties do not causally influence one another. This assumption does not match common-sense intuitions about how symptoms are related to one another. Insomnia is causally related to fatigue, which can trigger concentration problems. Viewing these as locally independent symptoms caused by some mysterious underlying condition is both unnecessary and unintuitive.

Critical realism and depression

David Pilgrim and Richard Bentall (1999) are critical of both the naïve positivism of standard biomedical approaches to depression as well as strong social constructionist approaches that problematise all empirical claims about the reality and causality of mental distress. Pilgrim and Bentall (1999) use a CR framework for analysing mental distress associated with depression. CR combines three fundamental assumptions (Pilgrim, 2017): (1) ontological realism, (2) epistemological relativism, and (3) judgemental rationality.

Like positivists, critical realists accept *ontological realism*, which is the view that there is an objective reality that exists independently of the way we perceive or describe it. In tension with this view is a commitment to *epistemological relativism*, which recognises that our understanding of the world is context-dependent and can shift across time and place as well as between individuals. For example, as shown by Horwitz (2011), our understanding of depression was shaped by intra-professional pressures within psychiatry during the creation of the DSM-III. However, CR is also committed to *judgemental rationality*, which is the capacity for humans to weigh up evidence for a scientific theory within a particular context. This is contrasted with *judgemental relativism*, which takes the view that truths about the world cannot be established and that there are only equally valid differences in perspective.

Critical realism is partly aligned with the *positivist approach* to mental distress in that it recognises that mental distress is real and that empirical claims about the causes of this distress can accurately describe reality. However, it aims to avoid naïve positivism by recognising that *our theories of reality* and the

methodological tools that are chosen to investigate reality are shaped by social context. In this sense, CR combines aspects of both positivism and social constructionism. Pilgrim and Bentall (1999, p. 262) summarise their position by stating that "deconstruction has a part to play in this exercise, but human science should not be reduced methodologically to this position alone. We can, and should, make attempts at investigating reality in itself, but do so cautiously and critically."

Pilgrim (2017) states that CR is a philosophical framework rather than a theory or model of mental distress. However, it is clear that CR is not just committed to ontological realism in any form, but to a specific form of ontological realism that views reality as complex and multi-layered. CR views the social world as a complex interconnected whole in a state of constant flux. Dynamic interactions within society create unpredictable emergent properties, and reductionist attempts to establish fixed, permanent and context-independent categories are unlikely to be successful. Instead, attempts to understand causality in complex systems will require epistemic humility and multifactorial explanations. This is a core difference between naïve and critical realism.

Pilgrim (2017) notes that any theory, model or methodology that blends ontological realism, epistemological relativism and judgemental rationality is compatible with a CR approach. For example, Pilgrim (2015) notes that the biopsychosocial model is compatible with CR. In the following sections, I will describe the network approach to depression, show how it avoids many of the limitations of the naïve biomedical model, and argue that this model is compatible with a CR approach to mental health, although, like the biopsychosocial model, it also has limitations from a CR perspective.

Basic principles of network models

The network approach to mental disorders was first proposed by Borsboom in 2008 (Borsbook, 2008), and this approach has grown substantially within psychiatric research over the last decade. (For an overview, see Borsboom, 2017b; Borsboom et al., 2019; Robinaugh et al., 2020). Proponents of the network model note that psychiatry has failed to find underlying biological mechanisms for mental disorders such as depression and argue that the reason for this is because these mechanisms simply do not exist (Borsboom et al, 2019). The network approach, therefore, rejects both the disease model of mental disorders and the assumption that mental disorders are discrete categories with demarcated boundaries. In an article summarising the network approach, Borsboom (2017b, p. 5) explains that "given the heterogeneity of the problems that psychiatry and clinical psychology deal with, it would perhaps be best to categorize them broadly as 'problems of living'." Coincidentally, or perhaps not, this is the same phrase Thomas Szasz (1974) used in his critique of psychiatry when he argued that mental illness was a myth.

Instead of viewing mental disorders as discrete diseases or as a socially constructed myth, the network approach views mental disorders as the product of causal interactions between symptoms. It starts from the simple observation that symptoms associated with mental disorders directly influence one another. For example, insomnia and tiredness are both symptoms of MDD. However, these symptoms do not covary because an underlying neurological mechanism independently causes both insomnia and tiredness; instead, insomnia directly causes tiredness. Similarly, tiredness causes concentration difficulties and a lack of motivation. Concentration difficulties and decreased motivation may lead to poor job performance, which may lead to feelings of worthlessness, which leads to a depressed mood and worse sleep quality, resulting in a self-reinforcing feedback loop. According to Borsboom et al. (2019, p. 4), mental disorders grow out of these networks of symptom-symptom relations, in particular, when they "are sufficiently forceful to lead the network to sustain its own activation" so that feedback loops between symptoms persist even after the precipitating cause has waned. Borsboom (2017b, pp. 7–9) proposes five principles that form the foundation of the network approach:

1. *Complexity:* psychopathology is the product of complex interactions between different components (biological, psychological, or social) within a network.
2. *Symptom-component correspondence:* properties of the psychopathology networks correspond to problems that have been codified as symptoms throughout the history of psychiatry.
3. *Direct causal connections between symptoms:* symptoms within a network causally affect one another.
4. *Mental disorders follow a network structure:* symptoms are connected in non-trivial ways. Symptoms may be closely or weakly connected to one another. Some groups of symptoms tend to cluster together more than others. When certain groups of symptoms cluster together, this creates a recognisable phenomenological manifestation of certain mental disorders. For example, symptoms associated with the construct of 'depression' tend to cluster together, while symptoms associated with the construct 'schizophrenia' are different from depression. They also cluster together in different ways, and the causal relationships between symptoms in patients diagnosed with schizophrenia are different from depression.
5. *Hysteresis:* In strongly connected networks, feedback loops between symptoms can be self-reinforcing. An adverse life event may trigger a subset of symptoms, such as depressed mood and sleep difficulties. If these symptoms are strongly causally connected, they may trigger other symptoms, such as loss of appetite or difficulty concentrating. If all of these connections are strong enough, the network can become self-sustaining and remain active even after the adverse life event is over. This property is called hysteresis and is common within complex systems. An example of this effect may be post-traumatic stress disorder, in which symptoms persist long after the triggering event has disappeared. In contrast, weakly connected symptom networks are not self-reinforcing, and once the adverse event is over, the network structure returns to its original equilibrium state. The network approach, therefore, provides a plausible explanatory mechanism for why adverse life events may trigger persistent psychopathological symptoms, but only in some people.

Figure 2 presents a simplified example of a network model featuring four symptoms (S_1–S_4). Symptoms that causally interact with one another are connected by lines. In this example, S_1 directly interacts with S_2 and S_3 but only indirectly interacts with S_4 through S_3. External factors in the environment (E_1–E_3) can also interact with symptoms, represented in this diagram by the external field.

Robinaugh et al. (2020, p. 1), in a literature review of 363 articles using network models, note that interest in network models has grown substantially since it was originally proposed in 2008, and it is "now a full-fledged area of psychiatric research". Network models of symptoms associated with MDD (Wittenborn et al., 2016), post-traumatic stress disorder (McNally et al., 2015), and psychosis (Isvoranu et al., 2016) have been developed, and research in this area is fast-moving and gaining recognition within academic psychiatry (Fried et al., 2017). Network science has been flourishing within biology and medicine for the last 20 years, and now, according to McNally (2019, p. 26), "the 'network takeover' has reached psychopathology, inspired by Borsboom and his colleagues."

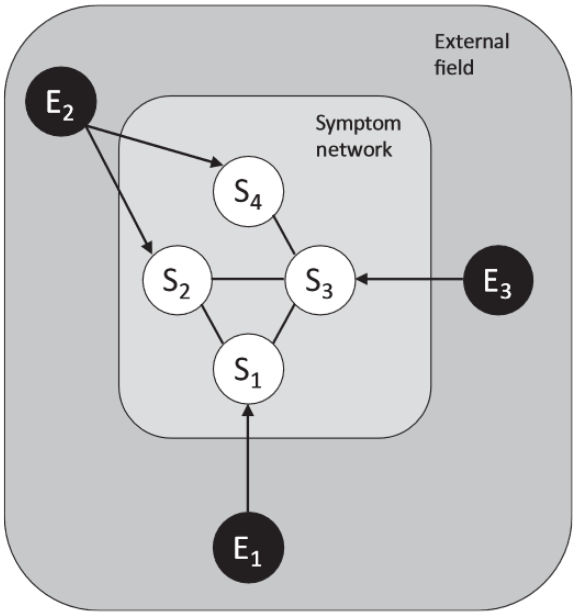

Figure 2. A symptom network featuring four symptoms (S_1-S_4).

Note: If two symptoms causally interact with one another, they are connected by a line. Symptoms that are not directly connected to one another (for example, S_1–S_4) may still synchronise if they share a neighbour. External factors that causally interact with the network are represented by the external field (E_1–E_3).

Source: D. Borsboom, (2017b, p. 6). Copyright 2017 by John Wiley & Sons. Reprinted with permission.

The network approach to depression in practice

Based on these philosophical principles, the network approach aims to develop mathematical models mapping the connections between symptoms associated with MDD. These models can then potentially be used to test hypotheses about the nature of psychiatric conditions. For example, one hypothesis within the network literature states that individuals are more likely to be vulnerable to MDD if their symptoms are highly connected to one another. This is because tightly connected symptom networks are more likely to produce self-reinforcing feedback loops within the network. A second hypothesis suggests that some symptoms may be more centrally connected than others within a particular psychiatric condition. Symptoms that are most central (that is, most interconnected) within the network may be particularly promising targets for intervention. For example, if lack of energy is highly connected to many other symptoms within the network, then interventions that help to address a patient's lack of energy may have a particularly strong effect on the rest of the symptom network. In contrast, interventions that target symptoms that are less centrally connected are less likely to affect the overall structure of the network. The network approach, therefore, aims to identify the most centrally connected symptoms in the hopes of developing more targeted interventions to treat MDD.

To address these open questions, Hakulinen et al. (2020) investigated the network structure of depressive symptoms among Finnish participants with MDD (DD+) and those without (DD–). The 13-item version of the *Beck Depression Inventory* (BDI) was used to measure depressive symptoms in both groups. The depression symptom network was estimated using a statistical method called *fused graphical lasso* (FGL). (As the current article is primarily focused on assessing theoretical and conceptual similarities between network models and critical realism, the statistical methods used to construct the model are beyond the scope of the present discussion.)

A visual representation of the FGL is shown in Figure 3. It shows the estimated network of depressive symptoms of DD– participants on the left and DD+ participants on the right. Each node on

the diagram (b1 to b13) corresponds to one of the symptoms on the BDI. The strength of connections between symptoms is shown by the weight (thickness) of the lines connecting each node. For example, in the DD+ group, pessimism about the future (b2) and loss of pleasure (b4) are strongly connected, while pessimism about the future (b2) and low self-esteem (b3) are only weakly connected. Connections between nodes shaded red are negatively correlated (for example, between nodes b5 and b13 in the DD+ group).

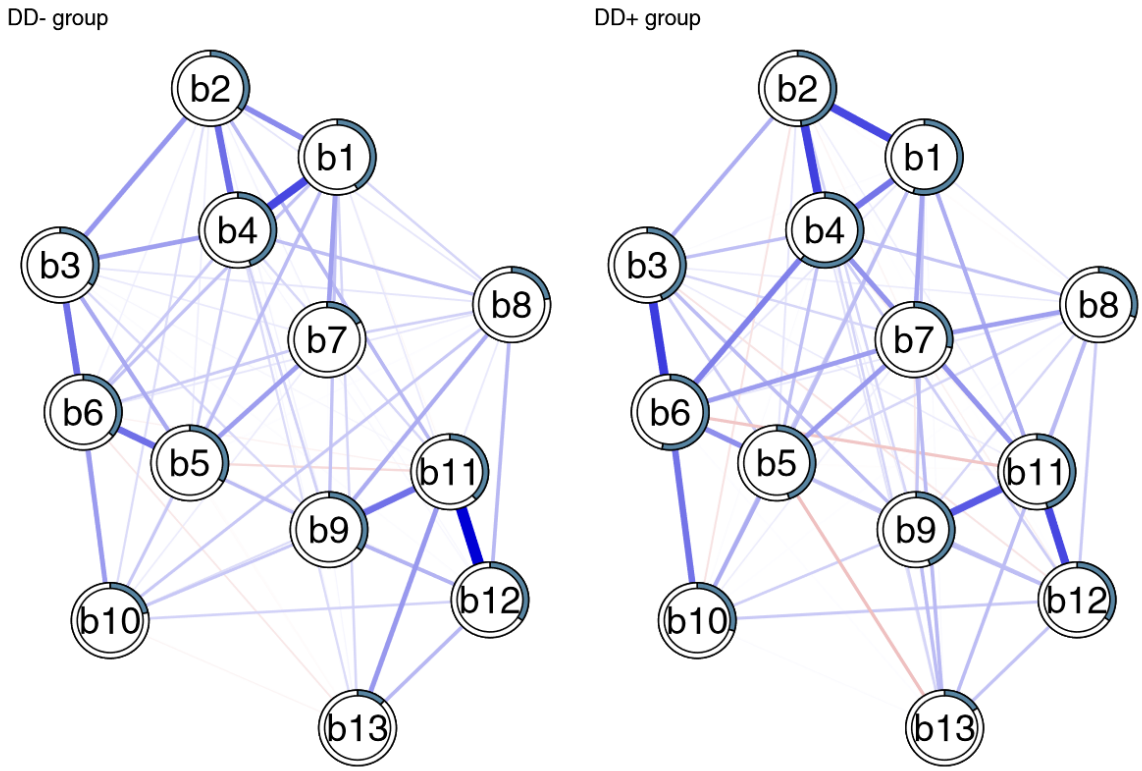

Figure 3. Visualisation of the fused graphical lasso (FGL) estimated networks of depressive symptoms in participants without (DD−) and with (DD+) major depressive disorder or dysthymia.

Note: Symptoms are as follows: b1 = depressed mood/sadness; b2 = pessimistic about the future; b3 = low self-esteem; b4 = loss of pleasure/dissatisfaction; b5 = feeling guilty; b6 = feeling disappointed in oneself/self-dislike; b7 = self harm; b8 = losing interest in other people; b9 = difficulties in decision-making; b10 = feelings of worthlessness; b11 = loss of energy; b12 = tiredness; b13 = loss of appetite.

Source: C. Hakulinen et al. (2020, p. 1276). Published under CC-BY license. http://creativecommons.org/licenses/by/4.0/

Hakulinen et al. (2020) found the most central symptoms among DD+ participants were loss of pleasure (b4), low self-esteem (b3), depressed mood/sadness (b1), and loss of energy (b11). They report that other studies have found similar results, suggesting that sadness, loss of pleasure, low self-esteem and lack of energy may be the most central symptoms of MDD. However, they found no difference in overall network connectivity between the DD– and DD+ groups. This result was unexpected and is inconsistent with the hypothesis that individuals with more highly connected symptom networks are likely to be more vulnerable to MDD. These findings, therefore, do not provide support for the hypothesis that vulnerability towards MDD is related to more highly connected symptom networks.

It should be noted that this is just one example of a network model in action, and there is significant variation between network models in terms of data collection and statistical methodologies. In a review of the literature, Robinaugh et al. (2020) find dozens of articles that have used networks model to understand the symptoms of depression, including Afzali et al. (2017), Beard et al. (2016), Bekhuis et al. (2018), Bos et al. (2018), Curtiss and Klemanski (2016) and Frewen et al. (2013). In contrast, the network approach appears

to have received no attention from the fields of critical psychiatry or sociology. This represents a significant gap in the literature.

Comparing the philosophical assumptions of the network approach and critical realism

As discussed in the previous section, the network model opposes reductionist explanations of mental disorders. Furthermore, because network influences extend beyond the brain and out into the external world, explanations of mental disorders must be context-dependent (Borsboom et al., 2019). Some connections between symptoms are likely to be stable across cultures because they are based on biological mechanisms. For example, insomnia is likely to cause fatigue and concentration problems no matter the cultural context. However, connections between some symptoms are likely to vary between cultures. Borsboom et al. (2019) point out that the connection between feelings of guilt and suicidal ideation is likely to be mediated by cultural factors. For example, a Japanese soldier in World War II who has brought shame to his battalion is more likely to experience suicidal ideation due to feelings of guilt than a Catholic priest who has committed a cardinal sin. This is because the cultural background of Catholicism inhibits suicidal ideation to a greater extent than Japan's honour culture does. Borsboom et al. (2019, p. 8) state that "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."

Borsboom et al. (2019) go further and claim that because the social and cultural environment is part of the network structure, the ontology of mental disorders extends beyond the brain and into the environmental context that causally influences symptoms of mental distress. They use gambling addiction as an example. Gambling addiction is not simply a mental disorder which exists entirely within a person's brain; it involves behaviour that mediates between the operational specifications of gambling machines and economic conditions that produce financial desperation. Therefore, according to the network approach, "not only cultural and historical features, but also the environment itself may become part of the network structure, and hence part of the disorder" (Borsboom et al., 2019, p. 8). This approach draws on the *extended mind hypothesis* within the philosophy of mind, which views the mind as extending beyond the brain and into the social environment (Clark & Chalmers, 1998). This, however, does not mean that psychopathology is just a social construction. While arbitrary labelling of mental disorders is a problem within psychiatry, and boundaries between disorders are unclear, symptoms associated with particular disorders do cluster together in highly reliable ways. Borsboom (2017a, p. 89) states that:

> Although the current classification schemes may not cut reality precisely at its joints, the allocation of these symptoms to disorders or disorder groups … does capture an important objective fact about reality, which transcends purely arbitrary labeling. Thus, mental disorders seem too real to be fully arbitrary, but do not correspond to neatly identifiable (sets of) essential properties.

The network approach, therefore, rejects essentialism of any kind with respect to mental disorders. Instead, the network approach has a more pragmatic approach to nosology: it simply asks whether classification systems are useful for understanding the aetiology, causation and treatment of mental disorders. A construct like depression is useful if it improves our ability to predict, understand and treat the symptoms associated with depression. It does not pick out some kind of essential, unchangeable and context-independent biological feature within a person. If the concept of depression is so poorly understood that it cannot be used for reliable diagnosis and treatment, then it should be abandoned.

The network approach quite clearly shares many ontological assumptions about mental disorders with CR. Both agree that mental distress is real and not merely a social construction. Both agree that reality is complex and multi-layered. Both see the world as an interconnected whole in a state of constant flux. Both view symptoms of psychopathology as an emergent property that results from dynamic interactions between multiple components within a system. Both agree that reductionist attempts to establish fixed, permanent and context-independent categories of mental disorders are likely to fail. Both are sceptical of the discrete categorisation of mental disorders. Both reject the biomedical model. Both adopt a fairly pragmatic approach to understanding mental distress. And both agree that attempts to understand causality require epistemic humility and multifactorial explanations. The network approach quite clearly assumes ontological realism about mental disorder and, in fact, appears to make the same ontological assumptions about the world as CR.

Critical realism, the biopsychosocial model, and the network approach

While critical realism is highly critical of the biomedical model, it has a more ambiguous relationship with the *biopsychosocial (BPS) model* of mental disorders. The BPS model views mental disorders as the product of a combination of biological, psychological and social causes. However, unlike the biomedical model, it does not necessarily give explanatory priority to biological factors. Strictly speaking, the BPS model is not a model in the mathematical sense but a mental model or theoretical framework. This lack of explanatory reductionism makes it more aligned with critical realism and the network approach.

This lack of explanatory reductionism also makes the BPS model more aligned with the network approach. Pilgrim (2015) argues that the biopsychosocial model (BPS) is compatible with a CR perspective, although it is limited by its tendency towards epistemological naivety. Given the similarities between the BPS model and the network approach, the limitations identified by Pilgrim may also apply to the network approach. Pilgrim (2015) notes five main limitations of the BPS model from a CR perspective:

1. There is a tendency to reify constructs like 'depression' as a real thing. While proponents of the BPS model talk about a variety of causal mechanisms that may affect mood, they often take the concept of depression as self-evident or as a stable medical fact.
2. Proponents of the BPS model are often biased towards biological factors that cause mental distress.
3. Related to (2), proponents of the BPS model tend to drift towards biological reductionism.
4. The history of the BPS model suggests that it was not developed as a genuine critique of biological reductionism but was developed to protect medical knowledge and defend its legitimacy in the face of evidence that social conditions affect medical outcomes.
5. The BPS model is not a reflexive theory about how social context and power structures shape medical knowledge and discourse.

Network models share limitation (5) with the BPS model, plausibly share limitation (2), and mostly avoid limitations (1) and (3). As far as I can tell, there has been no sociological research into the social context and intra-professional pressures that may have influenced the development and growing popularity of network models, so it is unclear whether network models also share limitation (4). More sociological research in this area is required.

Starting with limitation (5), network models are not a self-reflexive analysis of the social context that led to the creation of the concept of depression. Therefore, network models share limitation (5) with the BPS model. Regarding limitation (2), this is described within the network literature as the problem of node validity (Bringmann et al., 2022). Network analysis starts with the choice of variables (nodes) to include

within the system. However, as Bringmann et al. (2022) note, most network analyses have only included symptoms as nodes and have excluded environmental factors. In a review of theoretical problems associated with network models, Fried (2020) also notes that applications of network models rarely include environmental factors within the nodes of a network model. This suggests that limitation (2) may apply to network models. However, without more systematic research, this can only be tentatively claimed.

Even if network models are biased towards biological factors and tend to exclude environmental or social factors, I do not think the label of biological reductionism can be applied. As shown in Figure 2, in principle, the network approach includes environmental factors from the external field as part of the disorder. According to this view, mental disorders cannot be reduced to a latent biological entity or set of biochemical processes. Therefore, while network models may share limitation (2) with the BPS model, it avoids limitation (3). Network models also avoid limitation (1). Network theorists are rather adamant in their view that depression should not be reified as a self-evident medical fact. Instead, depression is a fuzzy, ambiguous, but sometimes useful construct that can be used to describe the tendency for a specific group of psychopathological symptoms to cluster together.

In addition to these rather general philosophical concerns, specific criticisms of the network approach have been made. For example, the falsifiability of network models has been questioned as it is unclear how hypotheses generated from network models can be tested (Bringmann et al., 2022). Similar concerns have been raised about the reproducibility of network models. Bringmann et al. (2022) note that it may be difficult for researchers to reproduce network models due to software availability and methodology changes over time. For example, the original code used to produce a particular model may not be available, may not be well documented or may contain bugs. Critical realists are likely to share these more specific concerns that have already been raised within the network model literature.

Discussion

In principle, both the network approach and CR share the same ontological assumptions about the nature of mental distress, and they both accept judgemental rationality—that is, that people have the capacity to weigh up evidence within a particular context. The final question is whether the network model is committed to naïve positivism or epistemological relativism. The exact requirement for avoiding naïve positivism from a CR perspective is a little unclear. However, given the fact that the network approach is not a self-reflexive theory of how social context shapes medical discourse and knowledge, I think the network model would still be seen as a form of naïve positivism, albeit significantly less naïve than the biomedical model or the BPS model.

Pilgrim (2015, p. 177) concludes that the BPS model "does a fairly good job about the ontology of disease", but it would be unwise to accept it uncritically. He suggests that critical realists could attempt to reform the BPS model to include a political and economic critique of the medical system, in this way combining the biopsychosocial model with a CR perspective. However, Pilgrim acknowledges that models are necessary simplifications of a complex reality, and the BPS model is still a useful simplification. It does one job reasonably well, but it cannot be expected to do everything. An attempt to reform the model from a CR perspective may increase the complexity of the model while decreasing its usefulness.

Pilgrim (2015, p. 178) suggests that "alternatively, maybe preferably, we could simply leave [the biopsychosocial model] be, accept its limited usefulness, but then knowingly augment it with a separate critical realist orientated sociology of professional and lay knowledge." This would allow critical realists to use the BPS model as a useful simplification when appropriate while also providing a division of labour where critical realists can focus on understanding the complex ways social factors and medical discourse shape our understanding of mental disorders.

A similar sentiment could be expressed towards the network approach. The network approach provides a high degree of ontological sophistication and avoids many of the epistemological pitfalls of earlier theories; however, thorough sociological research into the context in which the network approach was created has not been undertaken. Therefore, from a CR perspective, it would also be unwise to uncritically accept the network approach. Following Pilgrim's (2015) suggestions, one could attempt to reform network models to include sociological factors that shape our understanding of mental disorders, but it is currently unclear how these factors could be included, and this would add complexity to already complex models. It is also unclear—and in my view, unlikely—that this approach would increase the usefulness of current network models. Alternatively—and for this author, preferably—critical realists could accept the usefulness of network models as a necessary simplification while augmenting them with a separate critical realist analysis of how medical power and discourse has shaped the development and growing use of these models.

In practice, network models of depression appear to be less well aligned with the principles of CR. Like biopsychosocial models, current applications of network models appear to focus on biological factors and neglect environmental and social factors. Currently, there is little research into the extent to which social factors are neglected within the network approach, and overall, sociologists have paid little attention to the rise of network models within psychiatry. However, despite the practical limitations of current network models, there is considerable conceptual overlap between the network approach and CR. A combination of these approaches has the potential to provide an insightful analysis of MDD and other forms of mental distress.

Conclusion

The network approach provides an ontologically sophisticated and empirically grounded theory of depression. Interest in network models has been growing rapidly over the last decade, resulting in the publication of hundreds of papers defending, criticising and applying these models to analyse mental disorders. None of these critiques of the network approach have come from critical psychiatry or critical sociology. This represents a significant gap in the literature. Of the various theories within sociology, CR appears to be particularly well-placed to constructively critique the network model. The network approach shares fundamental ontological assumptions about the nature of depression with CR, while CR offers a useful meta-theoretical framework to understand the potential limitations of the approach. The network approach to depression informed by a CR perspective, therefore, offers a promising new approach for sociological theorising about the nature and causes of the symptoms associated with depression.

References

Afzali, M. H., Sunderland, M., Teesson, M., Carragher, N., Mills, K., & Slade, T. (2017). A network approach to the co-morbidity between posttraumatic stress disorder and major depressive disorder: The role of overlapping symptoms. *Journal of Affective Disorders*, *208*, 490–496. https://doi.org/10.1016/j.jad.2016.10.037

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). American Psychiatric Association.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

American Psychiatric Association. (n.d.). *What is depression?* https://www.psychiatry.org/patients-families/depression/what-is-depression

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*(16), 3359–3369. https://doi.org/10.1017/S0033291716002300

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*(2), 121–123. https://doi.org/10.1159/000486793

Bos, F. M., Fried, E. I., Hollon, S. D., Bringmann, L. F., Dimidjian, S., DeRubeis, R. J., & Bockting, C. L. (2018). Cross-sectional networks of depressive symptoms before and after antidepressant medication treatment. *Social Psychiatry and Psychiatric Epidemiology*, *53*(6), 617–627. https://doi.org/10.1007/s00127-018-1506-1

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

Borsboom, D. (2017a). Mental disorders, network models, and dynamical systems. In K. Kendler & J. Parnas (Eds), *Philosophical issues in psychiatry IV: Psychiatric nosology* (pp. 80–97). Oxford University Press.

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

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

Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., Epskamp, S., Eronen, M. I., Hamaker, E., Kuppens, P., Lutz, W., McNally, R. J., Molenaar, P., Tiao, P., Voelkle, M. C., & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy*, *149*, 104011. https://doi.org/10.1016/j.brat.2021.104011

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Curtiss, J., & Klemanski, D. H. (2016). Taxonicity and network structure of generalized anxiety disorder and major depressive disorder: An admixture analysis and complex network analysis. *Journal of Affective Disorders*, *199*, 99–105. https://doi.org/10.1016/j.jad.2016.04.007

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26, 57–63. https://doi.org/:10.1001/archpsyc.1972.01750190059011

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, *6*, 309. https://doi.org/10.3389/fpsyg.2015.00309

Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry*, *31*(4), 271–288. https://doi.org/10.1080/1047840X.2020.1853461

Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 1–11. https://doi.org/10.1186/s12916-015-0325-4

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

Frewen, P. A., Schmittmann, V. D., Bringmann, L. F., & Borsboom, D. (2013). Perceived causal relations between anxiety, posttraumatic stress and depression: Extension to moderation, mediation, and network analysis. *European Journal of Psychotraumatology*, *4*(1), 20656. https://doi.org/10.3402/ejpt.v4i0.20656

Horwitz, A. V. (2011). Creating an age of depression: The social construction and consequences of the major depression diagnosis. *Society and Mental Health*, *1*(1), 41–54. https://doi.org/10.1177/2156869310393986

Hakulinen, C., Fried, E. I., Pulkki-Råback, L., Virtanen, M., Suvisaari, J., & Elovainio, M. (2020). Network structure of depression symptomology in participants with and without depressive disorder: The population-based Health 2000–2011 study. *Social Psychiatry and Psychiatric Epidemiology*, *55*(10), 1273–1282. https://doi.org/10.1007/s00127-020-01843-7

Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, *348*(6234), 499–500. https://doi.org/10.1126/science.aab2358

Isvoranu, A. M., van Borkulo, C. D., Boyette, L. L., Wigman, J. T., Vinkers, C. H., Borsboom, D., & Group Investigators. (2016). A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms. *Schizophrenia Bulletin*, *43*(1), 187–196. https://doi.org/10.1093/schbul/sbw055

Jorm, A. F., Patten, S. B., Brugha, T. S., & Mojtabai, R. (2017). Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. *World Psychiatry*, *16*(1), 90–99. https://doi.org/10.1002/wps.20388

McNally, R. J. (2019). The network takeover reaches psychopathology. *Behavioral and Brain Sciences*, *42*, e15. https://doi.org/10.1017/S0140525X18001073

McNally, R. J., Robinaugh, D. J., Wu, G. W., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*(6), 836–849. https://doi.org/10.1177/2167702614553230

Pilgrim, D. (2015). The biopsychosocial model in health research: Its strengths and limitations for critical realists. *Journal of Critical Realism*, *14*(2), 164–180. https://doi.org/10.1179/1572513814Y.0000000007

Pilgrim, D. (2017). Critical realism. In B. M. Z. Cohen (Ed.). *Routledge international handbook of critical mental health* (pp. 24–30). Taylor & Francis.

Pilgrim, D., & Bentall, R. (1999). The medicalisation of misery: A critical realist analysis of the concept of depression. *Journal of Mental Health*, *8*(3), 261–274. https://doi.org/10.1080/09638239917427

Robinaugh, D. J., Hoekstra, R. H., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine*, *50*(3), 353–366. https://doi.org/10.1017/S0033291719003404

Scull, A. (2021). American psychiatry in the new millennium: A critical appraisal. *Psychological Medicine*, *51*(16), 2762–2770. https://doi.org/10.1017/S0033291721001975

Strickland, C. M., & Patrick, C. J. (2014). Biomedical model. In R. Cautin & S. Lilienfield (Eds). *The encyclopedia of clinical psychology*. https://doi.org/10.1002/9781118625392.wbecp476

Szasz, T. S. (1974). *Ideology and insanity: Essays on the psychiatric dehumanization of man*. Anchor Books.

White, K. (2017). Social constructionism. In B. M. Z. Cohen (Ed.) *Routledge international handbook of critical mental health* (pp. 24–30) Taylor & Francis.

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), 551–562. https://doi.org/10.1017/S0033291715002044

World Health Organization. (2021). *Depression*. https://www.who.int/news-room/fact-sheets/detail/depression