aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
1–22

DOI: 10.1177/1745691617705892
www.psychologicalscience.org/PPS
SAGE

Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology

Eiko I. Fried and Angélique O. J. Cramer
Department of Psychology, University of Amsterdam

Abstract
Since the introduction of mental disorders as networks of causally interacting symptoms, this novel framework has received considerable attention. The past years have resulted in over 40 scientific publications and numerous conference symposia and workshops. Now is an excellent moment to take stock of the network approach: What are its most fundamental challenges, and what are potential ways forward in addressing them? After a brief conceptual introduction, we first discuss *challenges to network theory*: (1) What is the validity of the network approach beyond some commonly investigated disorders such as major depression? (2) How do we best define psychopathological networks and their constituent elements? And (3) how can we gain a better understanding of the causal nature and real-life underpinnings of associations among symptoms? Next, after a short technical introduction to network modeling, we discuss *challenges to network methodology*: (4) heterogeneity of samples studied with network analytic models, and (5) a lurking replicability crisis in this strongly data-driven and exploratory field. Addressing these challenges may propel the network approach from its adolescence into adulthood and promises advances in understanding psychopathology both at the nomothetic and idiographic level.

Keywords
clinical psychology, dynamic systems, mental disorders, networks, personalized medicine, psychiatry, reproducibility

The last few years have witnessed a revolution in the field of clinical psychology and psychiatry. What has long been common knowledge among clinicians—that psychological problems interact with each other in complex ways—is finally acknowledged among researchers studying mental disorders (Borsboom & Cramer, 2013; Fried, van Borkulo, Cramer et al., 2016). In contrast to the hitherto default assumption that disorders cause their respective symptoms, the network approach conceptualizes mental disorders as networks of symptoms that directly interact with one another.

1. The Network Approach to Psychopathology

From the network perspective, certain symptoms like insomnia, fatigue, and concentration problems in patients with Major Depression (MD) co-occur not because they result from an underlying brain disorder or neurochemical imbalance but because not sleeping well leads to being tired and having concentration problems.

In order to study such symptom-symptom interactions, statistical models were developed and subsequently applied to a number of psychiatric disorders such as MD and psychosis (for a review, see Fried, van Borkulo, Cramer et al., 2016). In just a few years, numerous scientific papers were written, several book chapters published, conference keynotes given, and multiple workshops on network analysis held. This rapid acceleration of network research provides a crucial opportunity for us to pause and summarize some fundamental challenges to the network approach. Some of these derive from the novelties of the statistical methods—for example, how can we safeguard

Corresponding Author:
Eiko I. Fried, Department of Psychology, University of Amsterdam, Niewe Achtergracht 129-B, 1001 NK Amsterdam, The Netherlands
E-mail: eiko.fried@gmail.com

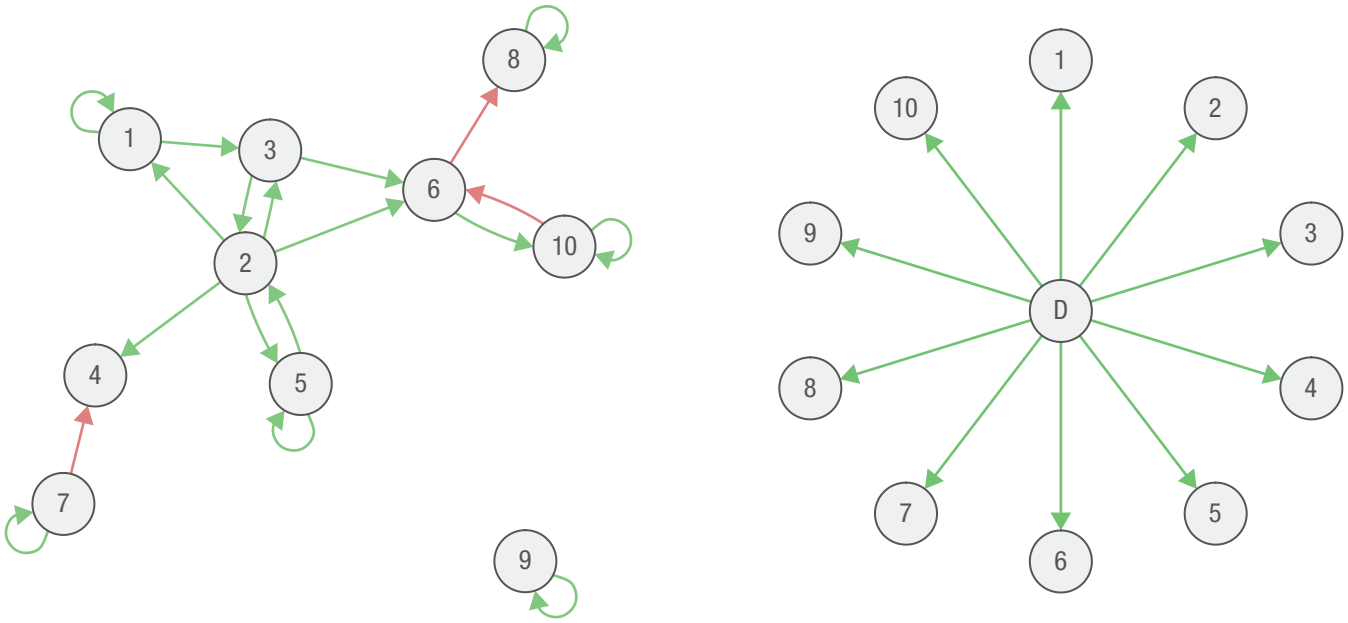

Fig. 1. (Left) The causal network structure for a particular mental disorder. The network consists of 10 nodes (the symptoms) and edges between nodes that depict the causal dependencies; green edges denote positive associations, red edges negative ones. Although networks are often weighted, meaning that edges differ in strength, for sake of simplicity, we display an unweighted network with equal edge strengths. (Right) Common cause model where the underlying disease D causes the symptoms 1–10.

against false positive associations in networks? Challenges to network theory, on the other hand, stem from the complex nature of mental disorders—for example, which disorders are best represented as networks, and which are not? The primary focus of the present paper is to discuss five challenges we deem most urgent if we are to move closer to steering the network approach toward becoming a mature scientific discipline.

1.1 Disorders: Common causes or networks?

Some symptoms co-occur more often than others. This rather simple observation has led to the formation of syndromes and disorders throughout the realm of medicine, including psychiatry. The network approach explains the co-occurrence of such symptoms as resulting from direct interactions between these symptoms (see Fig. 1 left): Insomnia can cause fatigue, psychomotor problems, and concentration problems, and these depression symptoms (APA, 2013) can form vicious circles of problems that are hard to escape (Borsboom, 2017; Borsboom & Cramer, 2013; Fried, van Borkulo, Cramer et al., 2016). Consistent with the network literature, we refer to variables in networks as *nodes* in the remainder of the text and to associations as *edges*.

The idea that symptoms cause each other is not new and has been discussed in the clinical literature in some detail (e.g., Beck, Rush, Shaw, & Emery, 1979). However, the network theory of mental disorders (Borsboom, 2017; Kendler, Zachar, & Craver, 2011; McNally, 2012) was only recently connected to sophisticated psychometric models that allow us to *estimate* such networks for empirical data (Bringmann et al., 2013; Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016b; Epskamp, Borsboom, & Fried, 2017; Epskamp, Maris, Waldorp, & Borsboom, in press; Gates & Molenaar, 2012; Haslbeck & Waldorp, 2015; Schuurman, Ferrer, de Boer-Sonnenschein, & Hamaker, 2016; van Borkulo et al., 2014).

A different approach to explaining covariation among symptoms is the *common cause model* (Fig. 1 right): Symptoms of a disorder *D* co-occur because they have the same underlying cause. An intuitive example of such a model is measles, which is associated with a very specific infectious agent (the common cause) that causes particular symptoms such as fever and Koplik's spots. Treatment of the underlying infection will cure the symptoms because their cause (the disease) disappears; Down's syndrome is another example where the syndrome clearly arises from an underlying chromosomal abnormality: Without the chromosomal abnormality, there would be no symptoms.[1]

The network approach and the common cause model differ fundamentally in their explanations of why symptoms co-occur in syndromes and have been discussed in greater detail elsewhere (Cramer, Waldorp, van der Maas, & Borsboom, 2010; Fried, 2015; Schmittmann et al., 2013). Using MD as an example again, the network perspective hypothesizes that an episode of MD arises from the causal interactions among symptoms such as sadness,

insomnia, and fatigue, whereas the common cause model hypothesizes an underlying cause that resides, for example, in the brain of patients and activates multiple depression symptoms at the same time (Insel et al., 2010). Another important difference is that symptoms of a given disorder are largely interchangeable or equivalent from a common cause perspective, because they are seen as passive indicators of an underlying cause. The network approach, on the other hand, necessitates inquiry into the nature of individual symptoms as well as their causal dynamics (Boschloo, van Borkulo, Borsboom, & Schoevers, 2016; Fried & Nesse, 2015b; Fried, Tuerlinckx, & Borsboom, 2014) and has arguably led to some questions and insights that do not arise from a common cause perspective.

In recent years, papers have often pitted the network approach and common cause models against each other to stress their divergent explanations of why sets of symptoms co-occur and to point out that contemporary research practices often rely on the (implicit) assumption of common causes. The present paper differs from this literature (including some of our own work) in advancing the point that the nature of mental disorders is likely more complex than a simple dichotomy between network and common cause models—both models might contribute to explaining the onset and maintenance of diverse psychopathological conditions, and pure versions of either model may often be unrealistic. A pure version of the common cause model necessitates that the underlying cause can fully explain the covariation among symptoms—that there are no direct causal links between symptoms—which seems implausible for many psychological problems. This is apparent in what is called "residual correlations" (Overall & Porterfield, 1963) in factor models: symptoms that remain correlated after trying to capture the shared variance among all items by one or multiple factors. For instance, in a recent psychometric paper on a depression rating scale (Horton & Perry, 2016), the items "sleep problems" and "fatigue" showed residual correlations after fitting a Rasch model, which is not surprising from a network perspective. A pure form of the network model, on the other hand, posits that the co-occurrence among symptoms is *solely* due to causal interactions among symptoms, which may also be unlikely considering the various factors that can trigger multiple symptoms at the same time.

It is one of the main challenges that the network perspective currently faces: For which disorders is a "pure" network model a promising candidate model, and which disorders might be better understood by taking a common cause perspective? And given the potentially unrealistic extreme versions of both models, how can we reconcile both frameworks in a unifying conceptual model: For which disorders would such *hybrid models* be plausible?

1.2 Overview

In the next chapter that deals with challenges pertaining to *network theory*, we will start out by (1) exploring these questions about the validity of the network perspective on psychopathology and propose hybrid models. Within this chapter, we discuss two further topics: (2) how we should define a psychological system and what are constituent elements of such networks, and (3) how we can gain a better understanding of the causal nature and real-life underpinnings of associations among symptoms. The third chapter starts with a brief introduction to network estimation and discusses challenges for *network methodology*. Specifically, we cover (4) the potential heterogeneity of populations we study with network analysis and (5) how we can avoid a replicability crisis in this emerging field of psychopathological networks, with a focus on stability and generalizability of network models. We conclude by sketching a tentative research program for the coming years. The *R*-syntax for conducting all analyses and generating all figures in the paper is available in the Supplemental Material available online.

2. Challenges for Network Theory

2.1 Validity of the network approach

MD has been the primary target disorder of network studies (e.g., Boschloo, van Borkulo et al., 2016; Bringmann, Lemmens, Huibers, Borsboom, & Tuerlinckx, 2015; Cramer, Borsboom, Aggen, & Kendler, 2013; Cramer et al., 2016; Fried, Bockting et al., 2015; Fried, Epskamp, Nesse, Tuerlinckx, & Borsboom, 2016; Pe et al., 2015; van Borkulo et al., 2015; van de Leemput et al., 2014).

There are good reasons for this: (1) *importance*—MD is among the most prevalent disorders and causes considerable impairment and societal burden (Kessler et al., 2003; Kessler, Chiu, Demler, Merikangas, & Walters, 2005); (2) *convenience*—many datasets of MD are readily available for reanalysis; (3) *plausibility*—the network approach appears plausible for MD symptoms, and the complex factorial nature of MD rating scales makes the existence of one underlying common cause highly unlikely (Fried, van Borkulo, Epskamp et al., 2016); and (4) *grounding*—the idea that problems are organized in vicious circles that fuel each other is well-established in the clinical MD literature (e.g., Beck et al., 1979). In this first section on the validity of the network theory, we explore how reasonable and worthwhile network research is for mental disorders other than MD.

Overall, we sketch three general possibilities: (1) The network approach might be a good candidate model; (2) the common cause framework provides a reasonable model; and (3) there are hybrid models where both common causes and networks play a role.

Mental disorders as networks of interacting symptoms. So when might a network be an adequate explanatory model? For panic disorder, the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; APA, 2013) defines a number of symptoms for which direct interactions seem to make sense (Borsboom, 2008): experiencing recurrent panic attacks (Criterion A) that lead to worrying about the consequences of the attack (Criterion B1), which in turn may cause a person to make behavioral changes such as avoiding places that are similar to the place where the panic attack was experienced (Criterion B2) (*panic attacks* → *concerns/worry* → *behavior changes*). The relationship between the symptoms and the disorder would thus not be reflective but mereological: Interactions *constitute* the disorder, and it can be seen as a formative latent variable (Fried, van Borkulo, Epskamp et al., 2016; Van Rooij, Van Looy, & Billieux, 2016). Such interactions, however, are not more than hypothetical for panic disorder as we are not aware of empirical studies.

For bipolar disorder, it seems feasible to conceive of direct relations between symptoms both *within* (e.g., racing thoughts [manic] → distractibility [manic]) and *across* (e.g., inflated self-esteem [manic] → loss of interest [depressive]) the manic and depressive poles of the disorder. The network approach could thus potentially explain two hallmark features of bipolar disorder: (1) distinct poles in which one can get stuck (e.g., having a manic phase because symptoms within this pole keep "infecting" one another) and (2) switching from one pole to another by means of symptom-symptom interactions across poles; identifying symptoms responsible for such switches would be crucial. While we are aware of a network paper on bipolar disorder (Koenders et al., 2015), a focused investigation of this question was not yet performed.

Mental disorders as common causes. For which disorders may the common cause perspective be a viable hypothesis? For PTSD, a plausible shared origin of symptoms is staring us in the face: the trauma itself. The *DSM-5* diagnostic criteria for PTSD clearly reflect the importance of the trauma as implicated in causing the development of symptoms: for example, traumatic nightmares (about the trauma), flashbacks (about the trauma), and intense distress after exposure to traumatic reminders. Without the trauma, these symptoms would not be present. Other criteria, however, do not directly implicate the trauma as the underlying cause, such as negative feelings about oneself or other people, and emotional numbing. Nonetheless, such symptoms are likely at least indirect effects of the trauma and would not have developed without the traumatic event. Interestingly, the relationship between the trauma and symptoms is likely mediated by numerous factors such as cognitive reappraisal (Cavanagh, Fitzgerald, & Urry, 2014) and emotion regulation (Nickerson et al., 2015), and it has been suggested that PTSD becomes persistent when individuals process the trauma in a way that leads to a sense of serious and recurrent perceived threat (Ehlers & Clark, 2000).

Although one can conceptualize a traumatic event as a common cause that explains (some) of the covariance among PTSD symptoms, we cannot ignore recent empirical literature in which PTSD has been conceptualized and estimated as a network of symptoms (Armour et al., 2016; De Schryver, Vindevogel, Rasmussen, & Cramer, 2015; Knefel, Tran, & Lueger-Schuster, 2016; McNally et al., 2015; Mitchell et al., 2017). For example, McNally and colleagues (2015) obtained some clinically plausible direct relations among PTSD symptoms: a strong relation between feeling trauma-related anger and concentration problems, between feeling alienated from others and experiencing loss of interest, and between feeling emotionally numb and having a sense of a foreshortened future. How can we reconcile these two perspectives on PTSD?

Mental disorders as hybrid models. Here we introduce hybrid models that we broadly define as any conceptual model that accommodates both common causes and a network structure between symptoms (Fig. 2). A simple example is one where the *onset* of a disorder is governed by a common cause, while its *maintenance* is fueled by direct interactions between symptoms (Fig. 2a). For PTSD, a trauma may be responsible for the initial development (i.e., onset) of PTSD symptoms, whereas these symptoms may directly interact with one another over time (e.g., anger that results in continuous concentration problems) such that the disorder remains present (i.e., maintenance). Substance abuse might also be a hybrid condition where symptom-symptom interactions are responsible for maintenance (e.g., *withdrawal* → *substance use*, or *substance use* → *legal problems* → *substance use*; Rhemtulla et al., 2016), but onset of repeated use may have its roots in imbalances in dopaminergic circuits (i.e., a shared underlying cause that leads individuals to exhibit specific behaviors). To complicate things further, this imbalance could keep activating certain symptoms even after disorder onset, which may then trigger other symptoms (e.g. Fig. 2d).

MD is a second example of a disorder that may best be described by a hybrid model. The majority of patients who develop a first episode of MD were previously exposed to an adverse life event or experienced chronic stress (Brown & Harris, 1989; Hammen, 2005), and such stressors and events might act as a common cause for depressive symptoms, while a network structure between these symptoms might lead to chronicity.

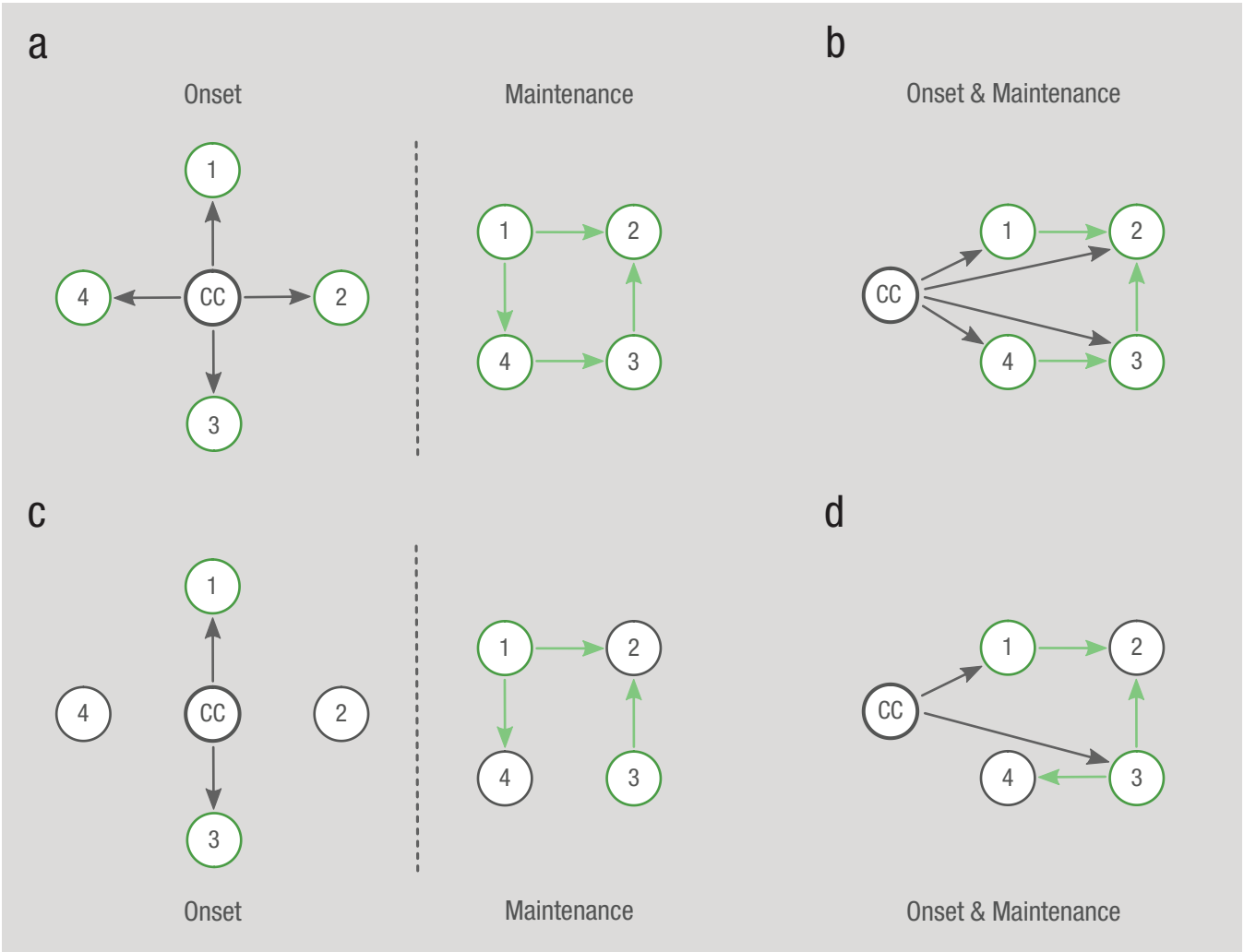

Fig. 2. Four different possibilities of a hybrid model; each node depicts a symptom; green nodes are nodes presently active. (a) A common cause (CC) that occurs only once—such as an adverse life event—triggers all four symptoms and leads to the onset of a disorder; then a network model with mutual interactions among symptoms is responsible for the maintenance of the disorder. (b) A chronic stressor acts as CC for all four symptoms and keeps activating them across time; at the same time, these symptoms interact with each other causally. (c) Same as panel a, but this time there is a local CC that only activates two of the four symptoms. (d) Same as panel c, but this time the chronic stressor is a local CC that only activates two of the four symptoms.

Immediately, several conceptual extensions of the hybrid model come to mind (Fig. 2 b–d). Unlike the situation shown in Figure 2a, an adverse or traumatic experience such as a divorce may not instantiate *all* symptoms of PTSD or MD, and the model shown in Figure 2c may be somewhat more realistic where the disorder serves as what we call a *local* common cause and only triggers specific symptoms (in this case, symptoms 1 and 3). For MD, the type of symptoms affected by a stressor may depend on the *type* of adverse event experienced (Fried, Bockting et al., 2015; Fried, Nesse, Guille, & Sen, 2015; Keller, Neale, & Kendler, 2007; Keller & Nesse, 2005): The death of a loved one can trigger different depressive symptoms than a divorce or losing a job. What panels a and c have in common is that the onset is governed by a (local) common cause, whereas the maintenance of an episode is primarily governed by direct symptom-symptom interactions.

Chronic stress might also lead to psychopathology—as opposed to a single adverse event with a clear ending. In this case, onset and maintenance may be difficult to distinguish, because the (local) common cause (i.e., chronic stress; Figs. 2b and 2d) keeps reactivating the symptoms, which then interact with each other in a network. Prospective studies on the experience of adverse life events or chronic stressors in populations at risk (e.g., medical residents; Sen et al., 2010) may offer a promising opportunity to investigate the validity of hybrid models. This is

especially so for common causes with clear material referents such as losing a job or perceived stress, which differs from more diffuse or abstract conceptualizations of common causes such as "depression" for depression symptoms (McNally, 2016).

Idiographic aspects of psychopathology. We have discussed how specific disorders might be conceptualized as either networks, common causes, or hybrids. However, we know that patients often differ dramatically in their etiology and symptomatology, and this holds for various disorders (Fried & Nesse, 2015a; Galatzer-Levy & Bryant, 2013; Olbert, Gala, & Tupler, 2014). Therefore, an equally interesting and possibly more complicated question is: Which of the three models described above fits the psychopathology *of a given person* best? MD, for instance, could stem from a common cause (e.g., brain pathology), a network model (e.g., vicious circles between negative thoughts and emotions; Beck et al., 1979), or a hybrid model (e.g., a network following severe adversity), depending on the specific individual and her or his specific circumstances. Or take anxiety disorders: For some people, vicious circles of negative emotions may describe the psychopathology best, whereas these negative emotions may simply be passive indicators of an underlying negative emotional disposition for others. Or take the causal chain *stimulus* → *worry* → *avoidance* common in phobias: Not every person fears dogs or mice, which implies that the appraisal of stimuli could mediate certain associations (in this case, *stimulus* → *worry*). This view stresses an idiographic perspective on mental health research and acknowledges that only embracing the heterogeneity of diagnostic categories will enable us to make true progress toward personalized medicine (Kramer et al., 2014; Molenaar, 2004).

The validity of diagnostic categories. It is difficult to critically discuss the network approach to mental disorders without acknowledging the elephant in the room: debates about the validity of diagnostic categories. *DSM-5* diagnoses such as PTSD and MD are highly heterogeneous phenotypes: Patients with the same diagnosis can exhibit very different problems. A recent study identified 1,030 unique depression symptom profiles in 3,703 depressed patients (Fried & Nesse, 2015a; see also Olbert et al., 2014; Zimmerman, Ellison, Young, Chelminski, & Dalrymple, 2014); for PTSD, there are 636,120 symptom combinations that all qualify for the same diagnosis (Galatzer-Levy & Bryant, 2013) (although not all of these may be clinically plausible). Many *DSM* disorders fail to meet orthodox criteria for validity such as a clear clinical presentation, precise diagnostic boundaries, treatment specificity, and temporal stability (Fried, 2015; Insel, 2013; Kupfer, First, & Regier, 2002; Parker, 2005), and the *DSM-5* field trials have documented questionable reliability coefficients for numerous mental disorders (Regier et al., 2013). For these reasons, a growing chorus of voices has suggested to investigate symptoms instead of syndromes (Costello, 1993; Fried, 2015; Fried & Nesse, 2015b; Persons, 1986).

Like many other clinical disciplines such as resilience research, genetics, or neuroimaging, many prior network papers were written for single disorders, because these syndromes arguably provide a reasonable starting point to investigate associations among symptoms. Given the high comorbidity rates among disorders and the central tenet of the network approach that problems attract problems—both within and across diagnostic boundaries—this calls for more transdiagnostic work. Although traditional models understand the co-occurrence between disorders such as MD and GAD as the result of two distinct etiologies, network models hypothesize that comorbidities arise due to shared symptoms between disorders. These symptoms can act as *causal bridges* and influence symptoms of both MD and GAD at the same time (Cramer et al., 2010). In this sense, the network approach naturally accommodates comorbidities as a central part of its theory.

In recent years, more network research has focused on comorbidity among two or more disorders (e.g., Afzali et al., 2017; Beard et al., 2016; McNally, Mair, Mugno, & Riemann, 2017; Robinaugh, Leblanc, Vuletich, & McNally, 2014), and several papers have looked into the network structures of psychopathology in general (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Boschloo et al., 2015; Boschloo, Schoevers et al., 2016; Tio, Epskamp, Noordhof, & Borsboom, 2016), as reviewed in detail elsewhere (Fried, van Borkulo, Cramer et al., 2016). For this paper, we understand mental disorders such as MD or PTSD not as reliable and valid phenotypes but as reasonable starting points for clinical investigations of the network structure of symptoms.

2.2 Constituent elements of psychopathological networks

The Merriam-Webster dictionary defines a *symptom* as "something that indicates the existence of something else" (*Symptom*, 2015). This means that the most important property of a symptom is that it is an indication of the presence *of* something else. The Cambridge dictionary includes this causal aspect even more clearly: A *symptom* is "any feeling of illness or physical or mental change that is caused by a particular disease" (*Symptom*, 2016)—for example, weight loss or nausea might point to the presence of a malignant tumor. Without an underlying condition or disease, however, the term *symptom* is meaningless. And although a

person can certainly have a medical disorder without a symptom (e.g., the beginning stages of certain cancers), this is quite impossible to envision for mental disorders (e.g., having schizophrenia without schizophrenia symptoms).

This traditional conceptualization of the relation between disorders and symptoms has granted symptom variables a certain importance above and beyond other clinical variables in psychopathology research. Network research has so far not been immune to this: Most studies have investigated the mutual interaction among symptoms as defined by the *DSM* while ignoring other variables that might be just as relevant in someone's problem economy. From a network perspective, however, symptoms are indicators not of an underlying disease, but rather of problems that are interacting over time. For this reason, different researchers have suggested that a better term for "symptom" may be "element" (McNally, 2012; Robinaugh et al., 2014; Snaith, 1993), because usage of the term "symptom" implies that the true model is a reflective latent variable model in which the common cause explains the covariance among symptoms (Fried, 2015). This semantic relabeling alone does not resolve the problem of the current limited focus on symptom networks, of course, but stresses that variables beyond symptoms may play a crucial role in psychopathological systems.

A working definition of a *dynamical system* in psychopathology may help structure the question of what other variables, apart from symptoms, may be relevant. We understand elements of systems here as a set of variables that can change over time, and additionally can both influence other variables and be influenced. Gender does not fluctuate and is immutable and thus would not be a sensible element in a dynamical system, whereas changes in mood make for a plausible addition. This leads to two conceptual types of variables we can explore: elements that are part of the system, and variables in the so-called *external field* that influence the system from the outside.

Problems beyond the *DSM*-defined symptoms are important candidates for inclusion in psychopathological systems. For example, the *DSM* diagnosis of depression does not list problems such as anxiety and irritability that are common, clinically relevant, and central symptoms in (networks of) depressed populations (Fava et al., 2008; Fried, Epskamp et al., 2016; Judd, Schettler, Coryell, Akiskal, & Fiedorowicz, 2013; ten Have et al., 2016). Apart from clinically relevant non-*DSM* problems (including emotions), impairment of functioning—for example, impairment at work, in social activities, or in a relationship—may be a crucial variable of interest. Although prior work has shown that depression symptoms may impact differentially on impairment (Fried & Nesse, 2014), it is unclear to what degree impairment feeds back into symptoms. Cognitive processes such as self-esteem or a sense of self-efficacy may be relevant, too, and distress as well as approach/avoidance behaviors could play an important role in anxiety disorders. Other promising constituent elements of psychopathological systems might include variables such as positive or negative social interactions per day, rejection events, physical activities, or substance abuse.

Life events provide an example of variables in the external field that can influence a psychopathological system from the outside: A traumatic experience can activate a number of PTSD symptoms, and adverse life events such as going through a divorce or losing a loved one can trigger symptoms of depression. As explained by Borsboom (2017), such external factors need not necessarily be outside the person. For example, well-studied risk factors for psychopathology include age, gender, intelligence, coping strategies, cognitive styles, and personality traits. *How* could such variables in the external field (e.g., losing a spouse) influence a system (e.g., a system of three connected symptoms—insomnia, fatigue, and depressed mood)? One straightforward possibility is that losing a spouse directly influences a specific system variable: losing a spouse $\rightarrow$ insomnia. A second possibility is that losing a spouse lowers the *threshold* for developing insomnia. A lower threshold means that insomnia can now more easily be activated by other nodes (e.g., depressed mood). Although such reductions in thresholds may be temporary when individuals undergo stress, there may also be people who have dispositionally lower thresholds for certain (a few, some, all) symptoms (e.g., neuroticism that may lower thresholds of depressive symptoms; Kendler, Kuhn, & Prescott, 2004; van Os & Jones, 2001).

We would also consider variables that change much slower over time than other elements in the system to be part of the external field, such as attributional styles or negative cognitions in depressed patients (Beck et al., 1979), or biased attention to specific stimuli in patients with social anxiety disorder (Heeren & McNally, 2016) that are more trait- than state-like. Note that the separation between elements within and outside the system is purely conceptual; depending on the time frame that is studied, it might make sense to consider the attributional style as an element of a patient's system, the same way it might make sense to consider a general emotional disposition (some people may have a stronger disgust response in general) as external. Additionally, impairment of functioning, which we argued to be a system variable, might alternatively be conceptualized as an *outcome* of pathological interactions between system variables and, as such, not as a part of the system itself. Empirical work is needed to answer these questions.

Topological overlap and missing nodes. We see two remaining challenges pertaining to the topic of constituent elements: (1) What if important variables are missing from a system, and (2) what do we do with nodes that are highly correlated and may measure the same construct (such as "sad mood" and "feeling blue")?

First, if a node that is strongly associated with a number of other nodes is removed, the network structure is likely to change substantially. If insomnia is strongly associated with both fatigue and feeling blue and fatigue and feeling blue are conditionally independent given insomnia (i.e., they show no partial correlation), removing insomnia from the network will lead to a strong spurious connection between fatigue and feeling blue. This implies that failing to incorporate all "relevant" variables (defined as those that covary with others) may lead to a misrepresentation of the network structure. Although erring on the side of including rather too many than too few variables may seem an easy solution, current network studies are often underpowered (Epskamp et al., 2017)—with too few persons for the number of parameters we model. Another concern when including many nodes is conditioning on a collider (Elwert & Winship, 2014). Novel statistical approaches may allow us to test which variables belong to the same causal system but are presently unsuited to handle the number of variables that are common in psychopathology (Sugihara et al., 2012). For now, it is up to the researcher to think beforehand about relevant variables to include for a given construct and to be careful not to overinterpret results of network analyses as representing reality.

The second question is the extent to which two nodes in a psychopathological network really represent different things. This problem does not arise for social networks where nodes are often individual people (Wasserman & Faust, 1994) or physical networks where nodes represent clear entities: computers that are connected via the Internet or airports that are connected via airplane routes. For a psychopathological network, however, nodes may not be that separable. Consider the two insomnia symptoms of "trouble falling asleep" and "early morning awakening" that co-appear in numerous rating scales for depression. These symptoms are strongly correlated with each other, which can mean two things: (1) They measure the same construct "sleep problems," in which case "sleep problems" should be a node in the network and not the two separate symptoms, or (2) they measure two different constructs—similar to height and weight, which are highly correlated yet different things—and should thus both be modeled. How can we know which of these options is most likely for a given pair of strongly correlated variables? This is a considerable challenge because rating scales in clinical psychology that are often used in network models were in many cases constructed to measure one underlying disorder. The Center of Epidemiological Studies Depression Scale (Radloff, 1977), for instance, features "sad mood," "depressed mood," "feeling blue," and "feeling happy"—all of which could be argued to be multiple measurements of one node in a network; the Hamilton Rating Scale (Hamilton, 1960), on the other hand, encompasses three different insomnia items. If these items would in fact measure one latent variable, this would bias centrality estimates (the items would be strongly interconnected, increasing their centrality estimates).

One potential way forward is to investigate *topological overlap* (Costantini, 2014; Oldham et al., 2008; Zhang & Horvath, 2005) (Fig. 3)[2]: If two highly correlated variables such as the two insomnia items measure the same construct, they should have very similar associations to all other nodes in the network. A solution in this case is to combine overlapping variables into one node (e.g., via a latent variable) (Fig. 3 left). In the second case, two variables would show differential associations with other nodes in the network and should not be collapsed into one node (Fig. 3 right). This latter option might be plausible for the insomnia items discussed above: Early morning awakening is more common among patients with melancholic depression, whereas trouble falling asleep might play an important role in relation to anxiety. Future studies will be required to test whether topological overlap presents an opportunity to guide decisions about what nodes to model in psychopathological networks. Investigations into the topic seem highly relevant, given the similarity of many items currently modeled as separate nodes in network analyses that may lead to spurious causal claims between symptoms in case they measure the same construct.

2.3 What is the stuff that networks are made of?

Psychopathological networks consist of associations among variables. There are two common assumptions about such associations: that they are causal in nature and that they are not just statistical parameters but reflect biological/psychological processes with material referents in the world. This section covers challenges pertaining to these topics.

Networks as causal systems. Despite the face validity of many psychopathological symptom pathways in networks, such as *insomnia* → *fatigue* → *concentration problems* → *psychomotor problems* in depressed patients, building a stronger case for the causal nature of these relationships is crucial if we want to truly advance them as *causal* systems. After all, many network papers have estimated undirected networks in cross-sectional data,

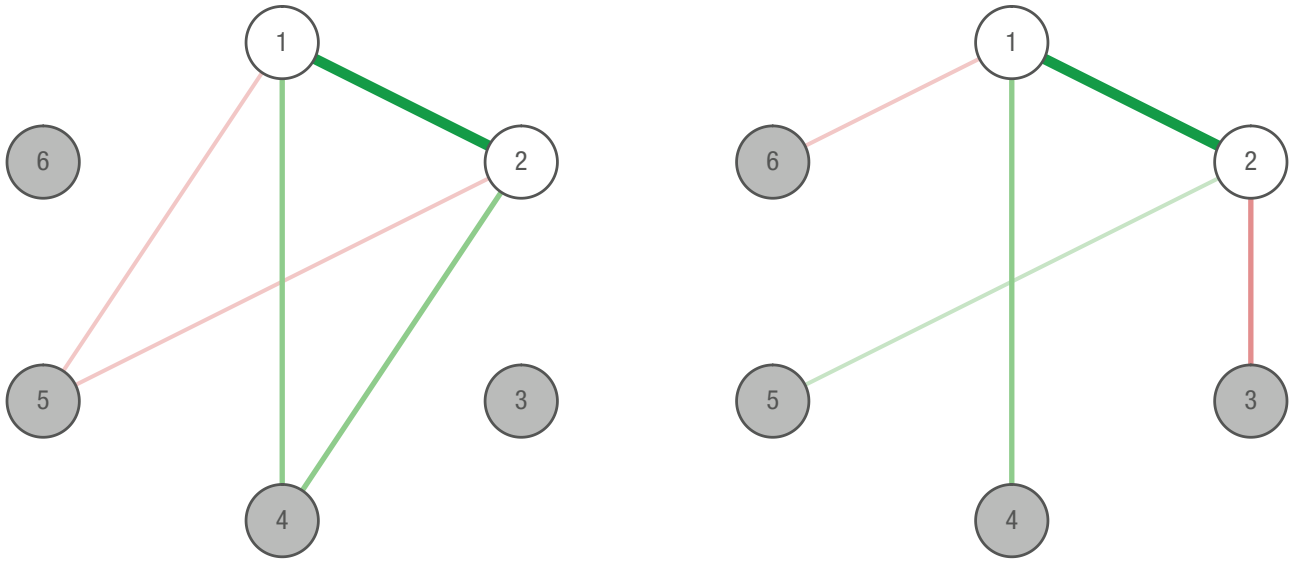

Fig. 3. (Left) The two white nodes 1 and 2 are highly correlated and exhibit similar relationships to other nodes (topological overlap); they may best be combined into one node. (Right) The nodes 1 and 2 are highly correlated but exhibit differential relationships (no topological overlap); they may best be both included in the network.

and even directed networks derived from time-series data constitute at best Granger causality (i.e., forecasting; Granger, 1969).

Both common cause and network models can give rise to the same correlations among symptoms (Epskamp, Maris et al., in press; Molenaar, 2003; Molenaar, van Rijn, & Hamaker, 2007; van der Maas et al., 2006), and experimental manipulations provide an excellent opportunity to test the causal hypothesis of networks, because networks and common cause models differ fundamentally in their prediction for such experiments: If *insomnia* → *fatigue* is the true model underlying the observed correlation between insomnia and fatigue, intervening on insomnia should reduce subsequent fatigue. In contrast, this intervention will not be successful if a common cause underlies the two symptoms, in which case *only* intervening on the common cause will successfully cure both symptoms.

Answering such questions about causality will have severe implications. In biological psychiatry, mental disorders are commonly conceptualized as the result of brain dysfunctions: For example, one of the most-cited and well-funded recent research frameworks, the NIMH's Research Domain Criteria (RDoC), explicitly states that all mental disorders are to be understood as brain disorders (Abbott, 2016; Insel et al., 2010). Such an assumption implies a common cause hypothesis about the etiology of psychopathology, a hypothesis that seems widely shared in the field (e.g., the neurotrophic hypothesis of depression; Schmaal et al., 2015). Establishing strong causal connections between symptoms would greatly limit the utility of such research (why should we investigate hippocampal volume as a common cause of depressive symptoms if MD is best represented by a network structure without a common cause), whereas identifying common causes for symptoms will make the network approach obsolete (at least in terms of symptom onset, not necessarily maintenance).

The psychology and biology of network parameters. Early papers about the network approach in psychology (Cramer et al., 2012; Cramer et al., 2010) elicited responses in which one common thread concerned the question, "Where is the biology and/or psychology in the proposed network models?" After all, network parameters such as edges and thresholds are by themselves statistical descriptions of psychological and biological processes and come from material referents in the real world. In a social network in which connections represent friendships, for instance, edges are a meaningful representation of actual social processes. The link between insomnia and fatigue, on the other hand, likely describes a host of intricate processes in a person's physiological system. The question arises as to what a low threshold for depressed mood in a person's network or a strong edge between fatigue and low self-esteem actually describes—what are the potential psychological and/or biological underpinnings of these parameters, and are they amenable to change so we can develop novel clinical strategies?

These questions imply the study of the real-world *mechanisms* that underlie network parameters. Such *mechanistic explanations* (Machamer, Darden, & Craver, 2000) of phenomena are, almost per definition, powerful

in terms of prediction: If we understand how a car works, we do not need statistical models to generate the prediction that it will not move if it does not have an engine. To date, only few studies have elucidated such potential mechanisms. One example concerns biological moderators between lack of sleep, on the one hand, and daily activities, concentration problems, fatigue, and alertness, on the other (Achermann, 2004; Borbély & Achermann, 1999); however, this work has so far not been connected to the psychopathological network literature. The report of Smeets, Lataster, Viechtbauer, and Delespaul (2014) provides another example: The authors showed that for early psychosis, the risk of developing delusions after experiencing hallucinations (i.e., the connection between delusions and hallucinations) is moderated by both genetic and environmental factors (Smeets et al., 2014), which may provide some leverage for thinking about novel prevention strategies.

Ways forward. An important first step toward exploring both the causal nature and mechanisms of symptom associations is to identify connections that appear consistently *across* many people. This could then inspire future research to test causal hypotheses and search for underpinnings of these pathways. Although *insomnia* → *fatigue* likely generalizes across the majority of both healthy individuals and people suffering from mental illness, many other pathways might only hold in patient samples, whereas others only hold for specific diagnoses. Additionally, the strength of these associations may be *moderated* by certain biopsychosocial variables and dispositions: Demographic characteristics like gender and age could influence symptom associations such as *anhedonia* → *suicidal ideation* or *hallucinations* → *delusions,* and so could biological processes such as glutamate neurotransmission that has been implicated in the etiology of numerous mental disorders (Grados, Specht, Sung, & Fortune, 2013; Riaza Bermudo-Soriano, Perez-Rodriguez, Vaquero-Lorenzo, & Baca-Garcia, 2012; Sanacora, Treccani, & Popoli, 2012; Schwartz, Sachdeva, & Stahl, 2012).

3. Challenges to Network Methodology

The following chapter about challenges to network methodology covers (1) heterogeneity of networks and (2) stability and replicability issues in network research. To facilitate the discussion of statistical challenges, we introduce the two commonly used network models and some basic terminology.

Researchers have predominantly used two types of models to study interactions among symptoms, emotions, and/or daily experiences. The first is used for cross-sectional between-person data (i.e., the network is estimated for a particular *sample* of participants that were measured at one time point; e.g., Fig. 6 left), and the second for within-person time-series data (i.e., a network is constructed for one or more people measured several times per day for multiple weeks; e.g., Fig. 1 left). A common way to estimate between-person networks is to use regularized partial correlation networks (Epskamp & Fried, 2016) that are available for binary, metric, or mixed data; we will refer to these networks as Pairwise Markov Random Fields (PMRFs) in the remainder of the text.[3] In such PMRFs, edges can be understood as partial correlations, and an edge between A and B in a network implies a relationship between these two variables that remains after controlling for all other nodes in the network. Likewise, the absence of an edge between two nodes means that these two variables are *conditionally independent* given all other nodes in the network. PMRFs are *undirected* and thus feature no arrows in their visualization (e.g., Fig. 6 left)—edges represent associations or connections and should not be misunderstood as causal. PMRFs entail a series of regressions in which each variable serves as the dependent variable with all other variables as potential predictors, which means that PRMFs are exploratory and data-driven and explains why they require larger samples because they estimate a large number of parameters (Epskamp & Fried, 2016; van Borkulo et al., 2014). Of note, estimating PMRFs in *R* uses *regularization techniques* with the goal to avoid estimating false positive associations. Simplified, regularization means that the partial correlations between nodes are estimated in a very conservative way: All edges are shrunk and small edges set to zero, resulting in *sparse* networks. This safeguards against erroneously concluding that a particular edge is present while it is not.[4] In contrast to latent variable models where we model the shared variance of a set of items, we estimate the unique variance of items in PMRFs.

The second class of models are used for intra-individual time-series data where an individual is measured multiple times a day for day, weeks, or months. A commonly used model is the vector autoregressive model (VAR model; Chatfield, 2003; Lutkepohl, 2005; e.g., Fig. 1 left). Here, associations among nodes for a particular person are estimated both *within time* and *across* time: We obtain an undirected network for the contemporaneous (within time) connections and a directed network for the associations across time (Epskamp, van Borkulo et al., 2016). It is also possible to estimate VAR models for a group of individuals. These so-called multilevel VAR models (Bringmann et al., 2013; Schuurman et al., 2016) allow for separating the within-person dynamics from the stable between-person differences; they result in a directed network for each individual person, a group-level network across all persons, and a variability network that shows to which degree all individuals in the group differ in their networks (see Fig. 4 bottom). Regularization methods are not implemented for VAR models

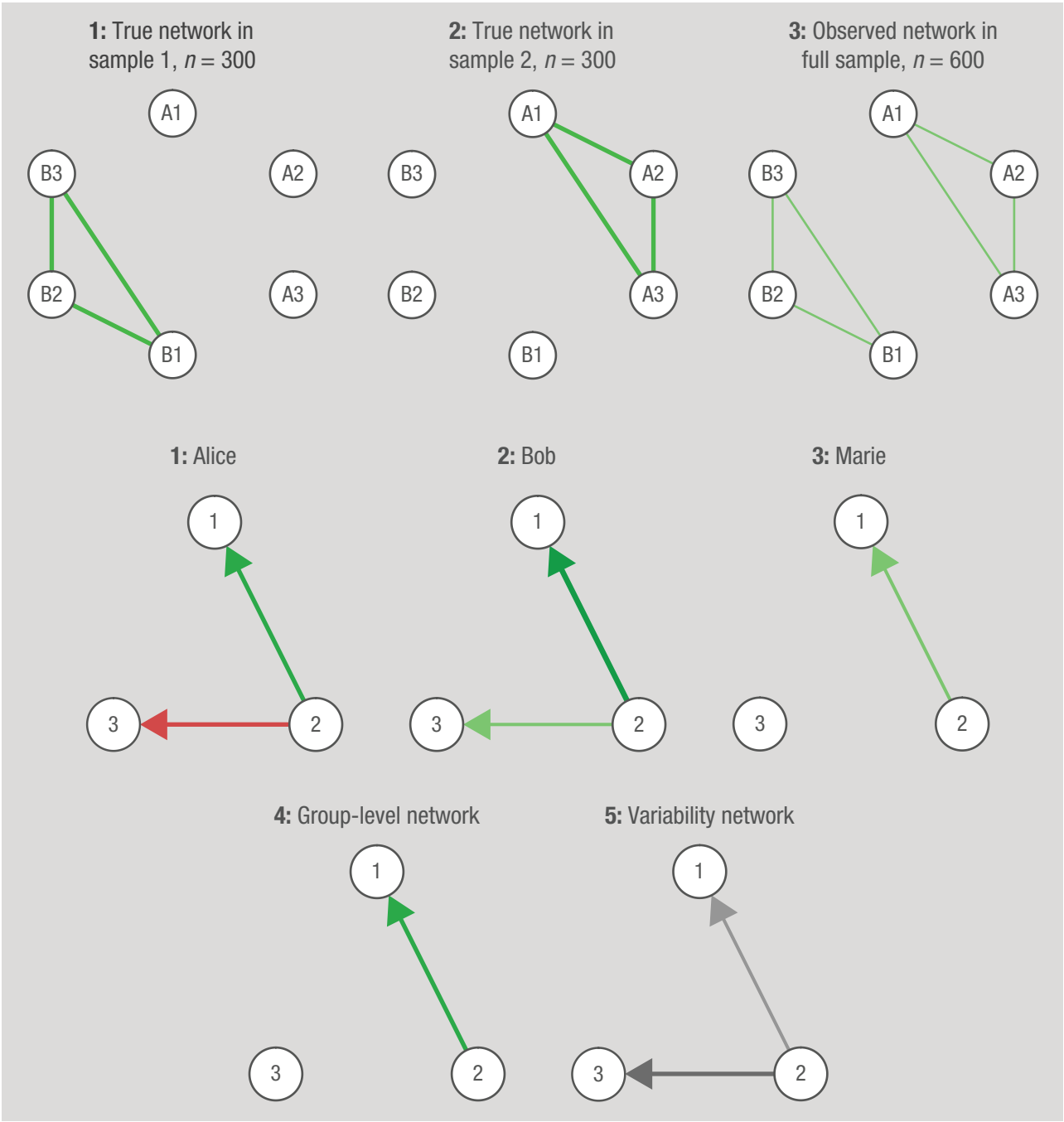

Fig. 4. (Top) When we are unaware of two underlying populations with different networks 1 and 2, the observed population network 3 will be the average of the two underlying networks. (Bottom) Alice, Bob, and Marie have different temporal networks (1–3). While the group level network 4 conceals these differences, the variability network that depicts the strength of differences of each edge offers insights into difference across participants.

(the False Discovery Rate has been used to control for multiple testing; Bringmann et al., 2015), although there is currently work in progress on the topic.[5] Another promising approach to model both group- and individual-level relations in time series data is the Group Iterative Multiple Model Estimation (GIMME; Gates & Molenaar, 2012).[6] Although both VAR models and GIMME cannot establish causality in a strict sense, they meet the requirements for Granger causality: In case a variable predicts another across time, we can conclude one Granger causes the other (Granger, 1969).

Answering crucial research questions that try to capture the causal nature of psychopathological processes will require the collection and analysis of temporal data,

which is also reflected in the shift from between-person to within-person publications (Fisher, 2015; Wichers, Groot, Psychosystems, ESM Group, & EWS Group, 2016; Wichers, Wigman, & Myin-Germeys, 2015; Wright & Simms, 2016). Such temporal data will enable us to investigate if and how networks change across time (Bringmann et al., 2016) and could lead to better prediction of psychopathology onset, treatment response, and relapse.

3.1 Heterogeneity of networks

As we discussed in more detail above, there is considerable heterogeneity within diagnoses such as PTSD and MD, and patients may exhibit very different problems (Fried & Nesse, 2015a; Galatzer-Levy & Bryant, 2013; Olbert et al., 2014). What challenges does heterogeneity imply for the network approach to psychopathology?

Heterogeneity in cross-sectional network models. For cross-sectional between-person networks, we can turn to the SEM literature that has long acknowledged this problem. In SEM, so-called *mixture models* like latent class analysis or cluster analysis assign people to subgroups based on their symptom profiles (Hagenaars & McCutcheon, 2002). This can result in more homogeneous and informative classes of patients that may exhibit similar problems, show similar responses to specific treatments, or have more similar biomarker profiles (ten Have et al., 2016; Wardenaar, Monden, Conradi, & de Jonge, 2015).

Although it is easy to conceive of a similar situation in which multiple network structures are present in one population, such mixture models are not yet available for network models. Imagine a cross-sectional study of 600 patients with the same diagnosis: It is possible that half of the patients show pronounced associations only among symptom cluster A, whereas the other half of the sample exhibits associations only among cluster B (see Fig. 4 top). Averaging over these two subpopulations in one model leads to an amalgam network that does not reflect the true population networks and would likely lead to unwarranted clinical conclusions. This also implies that we should be extremely careful when drawing intra-individual inferences (e.g., we should target the central symptom depressed mood in therapy) from the results of cross-sectional network models. Cross-sectional network models are capable, however, of generating hypotheses at the *group* level: for example, the potential hypothesis that women—as a group—have a more strongly connected depression network than men—as a group. *Network mixture models* would allow us to identify such groups of people in a data-driven way that are more homogeneous in their respective group-level networks.

The main challenge here is the relationship between sample size and parameters. To estimate between-person network models, we commonly use PMRFs that require a large number of estimated parameters; with 20 nodes in a network, we need to estimate 190 edges, and 1,225 edges in a network with 50 nodes.[7] Although there are no clear guidelines yet as to how many participants we need per parameter, a rule of thumb put forward was at least three people per parameter; however, recent work on network stability (Epskamp et al., 2017) has shown that this may not be sufficient to estimate networks accurately, implying that we may need an even higher observation-to-parameter ratio. Mixture models would further increase the number of estimated parameters—about twice as many parameters for two subgroups of people—requiring samples much larger than the size of many psychopathological datasets.

Heterogeneity in time-series network models. Heterogeneity is equally relevant for intra-individual time-series networks as it is for cross-sectional between-person networks (Fig. 4 bottom) because most time-series papers so far have focused on the group-level networks in multilevel VAR models and not on the intra-individual networks (e.g., Bringmann et al., 2015; Bringmann et al., 2013). Although certain edges may only differ slightly across participants, other pathways may differ substantially. In Figure 4 bottom, node 2 activates node 3 at the next time point for Alice, but the opposite holds for Bob; for Marie, the two problems are unconnected. For all participants, nodes 1 and 3 are unconnected, and node 2 triggers node 1, but in varying degrees (in decreasing order for Alice, Bob, and Marie). The group-level network resulting from a multilevel VAR model results in an empty edge between nodes 2 and 3 (because the average of Marie's positive, Bob's negative, and Alice' absent edge is 0) and a moderately strong positive edge for $2 \rightarrow 1$, obfuscating important differences across participants.

A way to investigate unobserved heterogeneity in the realm of VAR models is the estimation of a *variability network* to identify which edges vary considerably across participants and which edges are similar (Fig. 4 bottom) (Bringmann et al., 2013). This allows us to identify symptom pathways that generalize in the population (i.e., nomothetic in contrast to idiographic symptom associations), along with pathways with large inter-individual variability, providing an important step towards uncovering heterogeneity. In our case, there is only small variability in the coefficient from $2 \rightarrow 1$ that differs somewhat across participants and large variability for $2 \rightarrow 3$. Bulteel et al. (2016a) recently proposed a data-driven method to group participants in VAR models according to their VAR regression weights while simultaneously fitting a shared

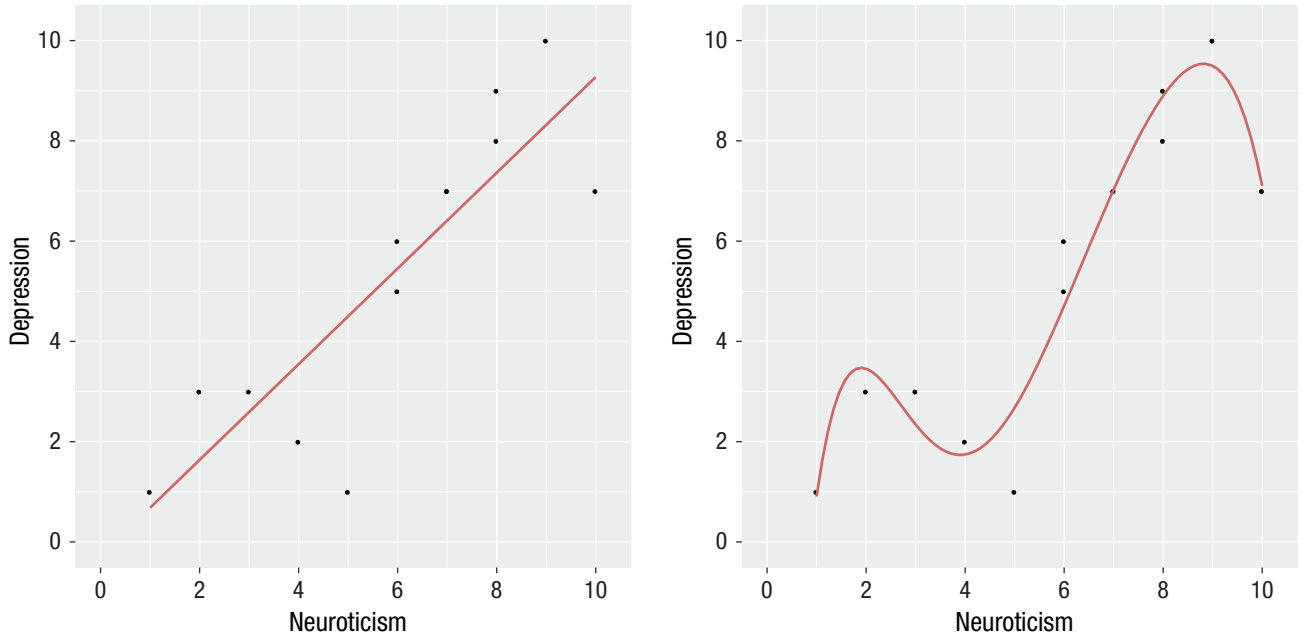

Fig. 5. Two increasingly complex models that explain the relationship between the two variables of depression and neuroticism. (Left) A linear regression. (Right) A complex polynomial function that explains the present dataset better but may not generalize well to other datasets.

VAR model to all persons within a group; this allows for detecting latent clusters of people with similar dynamics.

3.2 Stability and generalizability of psychopathological networks

Numerous scientific disciplines suffer from what has recently been called the *replicability crisis* or *reproducibility crisis*. In cancer research, an investigation showed that only 6 out of 53 landmark papers could be replicated (Begley & Ellis, 2012), and a recent study concluded that the irreproducibility of preclinical research in the life sciences surpasses 50%, leading to about US$28,000,000,000 spent on research in the United States every year that does not replicate (Freedman, Cockburn, & Simcoe, 2015). The replication crisis also hit psychology: An investigation of the Open Science Collaboration showed that the majority of 100 social psychology experiments were not replicable (Open Science Collaboration, 2015). In this section, we will discuss challenges to the generalizability and stability of psychopathological network research. In other words, how do *we* avoid a replicability crisis?

Generalizability of psychopathological networks. Network models in psychopathological research are data-driven and exploratory. This pertains to both cross-sectional network analyses (e.g., via PMRFs) and time-series modeling (e.g., via VAR models). A general problem with any exploratory model is that it is not necessarily the best model that will generalize to other datasets (Yarkoni & Westfall, 2016). Consider a very simple example in which we want to understand the relationship between neuroticism and depression. Figure 5 shows the results of two models, and a complex polynomial describes the data much better than a simple linear regression. However, this complicated model describes not only the relationship between depression and neuroticism but also measurement error, which leads to overfitting (Babyak, 2004). As a result, the model we chose for explanation may not generalize to other data, and the regression model that fits our dataset worse may constitute an excellent model for the relationship between depression and neuroticism in other datasets.

In the case of network models, overfitting is an especially severe challenge because we investigate relationships among a large number of variables, which means there is danger of overfitting a large number of parameters. One way to mitigate this problem somewhat is to *regularize* networks, a procedure that leads to *sparse* networks that we have discussed above in some detail (Friedman, Hastie, & Tibshirani, 2008; Tibshirani, 1996). Regularization techniques try to explain the covariance among symptoms with as few connections as possible and reduce the danger of overfitting by shrinking all connections and by setting small coefficients exactly to zero (Epskamp & Fried, 2016). This will result in network models with a lower fit to the data (less explanation) but may increase prediction (replicability in other datasets). However, it is unclear at present to what degree

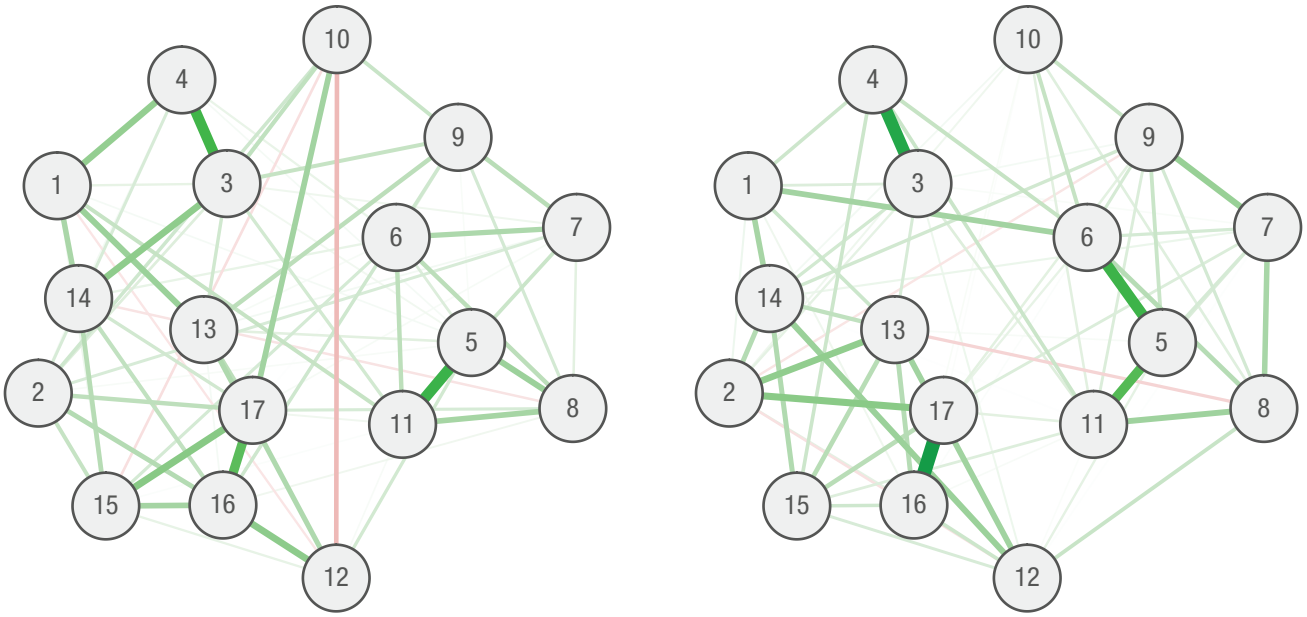

Fig. 6. Two networks of 17 PTSD symptoms in two different samples. (Left) A network estimated in a sample of 180 female participants with PTSD. (Right) A network estimated in a different sample of 179 female participants with PTSD.

regularization techniques increase the generalizability of network research.

Stability and accuracy of psychopathological networks. When we analyze a particular psychopathological dataset, we usually obtain one estimated network model and visualize the model in order to depict the multiple dependencies among variables (such as symptoms). The main question we discuss in this paragraph is how *stable* such network structures are—that is, how *accurately* are the parameters estimated, and how likely are they to *replicate* in a different dataset?

So let us write a quick paper together to see why stability matters. We estimate a network of 17 PTSD symptoms in a sample of 180 women with posttraumatic stress disorder[8] (Fig. 6 left): A strong edge emerges between 3 and 4, representing a clinically plausible association between *being startled easily* and *being overly alert* (for example). We also observe a negative edge between symptoms 10 and 12 and conclude that people who, for instance, do not remember the trauma are less likely to have trouble sleeping (and vice versa). In a second step, we investigate the centrality (connectedness) of nodes (Opsahl, Agneessens, & Skvoretz, 2010). In our example network, node 17 has the highest degree of centrality (1.25) and node 7 the lowest (0.65). We now finalize the paper and suggest that future studies should pay specific attention to edges 3–4 and 10–12 and that targeted treatment of node 3 may achieve the greatest benefits for patients. Success!

But are these clinical conclusions really warranted—how likely is it that another study with similar data of female PTSD patients would result in a similar network structure in which the same edges play the most important role and the same symptoms are the most central symptoms? To answer this question, we obtain a second dataset of similar size (179 female PTSD patients) and estimate a second network in this dataset (Fig. 6 right).[9] The two resulting networks look somewhat similar, but there are also differences; for instance, the negative edge between 10 and 12 that we pointed out as clinically relevant in our hypothetical paper above disappears. Furthermore, in contrast to the first network, the most central symptom is now node 6 (1.20), the least central one node 10 (0.45), fundamentally different from the previous results, and the correlation of centrality estimates between the two networks is only 0.48.

Little research has been conducted on the topic of how stable or accurate network parameters such as edge weights and centrality estimates are. This is problematic, because current routine practices may be prone to chance findings and vulnerable to interpretations that are not as generalizable as one might hope. A way forward is to investigate the accuracy of network parameters such as edges and centrality estimates, which will help us answer whether a very strong edge such as 3–4 (edge weight 0.38) in Figure 6 left is significantly different from the barely visible edge 3–11 (edge weight 0.09); bootstrapping 95% confidence intervals reveals that this is not the case here, implying that we should not interpret the first edge as substantially stronger than the second.[10] Several tools for investigating the accuracy of network parameters are available in a novel *R*-package *bootnet* (Epskamp et al., 2017). To our knowledge, this is the first approach of tackling the challenge of reproducibility of psychopathological networks, and we are looking forward to seeing more conceptual and methodological

developments with the aim of estimating accuracy and stability of networks. In case the CIs of many edges overlap (i.e., edges look differently strong in the graph, but we cannot reliably distinguish between them statistically), a way forward may be to turn a weighted network as in Figure 6 into an unweighted network (where connections are either absent or present but do not differ in strength). This may be a more accurate visual representation of the output: We may often not be able to reliably distinguish between stronger and weaker edges, but regularization techniques will often reliably distinguish absent from present edges.

In general, estimating and reporting the accuracy of network parameters in scientific publications is at best a first step of tackling the challenge of replicability. To move the field forward, we require cross-validation across similar samples to investigate whether network models of, for instance, MD replicate in different datasets. If the sample is large enough, another approach to cross-validation would be to fit a network model to half of the sample and then test to what extent that model holds in the other half of the sample. A more general recommendation to enhance replicability is to develop methods for confirmatory network modeling: to impose a specific network structure, instead of current data-driven approaches—and test the absolute fit of that model to the data.

4. Discussion

The network perspective on psychopathology is starting to mature from an intuitive idea into a scientific discipline. And because it is the new kid on the block in psychopathology research, it faces some challenges that either pertain to network *theory* (e.g., the validity of the network perspective) or to network *methodology* (e.g., the stability of network models). Naturally, the distinction between theory and methodology that we have maintained throughout the paper is not absolute: Future methodological advances may help in advancing theoretical issues about the validity of the network perspective for a given disorder, and progress in network theory (e.g., what constitutes a system) will likely inspire the development of novel methodology (e.g., method to determine which elements belong to a system and which not).

Future directions

In only a few years, the network approach has renewed an arguably much-needed focus on the *individual* and his or her specific psychological problems (Molenaar, 2004). If the next few years can generate solutions to the challenges that we have outlined, we see a promising future for personalized clinical psychology/psychiatry in general and for the network perspective in particular.

But how to tackle these challenges? In addition to some directions discussed in the prior sections, we sketch a few tentative possibilities here. First of all, this paper serves as a call to action for *methodologists*. Specifically, the network perspective will benefit from the following methodological advances: (1) confirmatory network modeling—that is, models with which we can confirm hypotheses about network structure instead of exploratory, data-driven, network analyses; (2) mixture network modeling—that is, models with which we can test the existence of subgroups with different network models underlying one population network; (3) methods to statistically compare latent variable models to network models; and (4) power recommendations for network analysis—that is, how many participants do we need to reliably model the association among k nodes in a cross-sectional network model and how many time points do we need to reliably estimate the associations among k nodes for n persons in a time-series model. (5) Finally, it is important that empirical researchers gain a better understanding of network models and their assumptions. And because such insights are critically dependent on nontechnical and accessible explanations that are presently scarce (Costantini et al., 2014; Epskamp et al., 2017; Epskamp & Fried, 2016), we call for tutorial papers on network estimation and interpretation in the realm of psychopathology.

Second, we provide some conceptual guidelines for *empirical researchers*. (1) We need a better conceptual understanding of mental disorders as networks, and clinical theory may help guide a priori decisions on which the model is the most accurate account of a particular disorder. For example, do we believe that the etiology of a particular patient with PTSD resembles a pure network model or a hybrid model? This affects which models we use for the network analysis. (2) We should let our research questions guide the decision of what kind of data are most suited for answering it. For instance, many interesting research questions can be investigated at the level of the individual in a clinical setting (e.g., will cognitive-behavioral therapy be the optimal intervention for Patient X with Disorder Y), whereas others are best examined at the level of the population (e.g., why are depression rates higher in women than men). The network perspective has both idiographic and nomothetic sides, and both hold promising potential. (3) Researchers should test and report the stability of their network models. This would strengthen the robustness of empirical research in this emerging field, safeguard against false-positive results, and also help us to identify consistent pathways that are highly reliable across studies. To that end, we can cross-validate networks in confirmatory analyses, compare results of network analysis with those

reported in similar datasets, and use statistical tools to ascertain the accuracy of estimated network parameters.

Final thoughts

We want to conclude by listing some challenges that go beyond the size limitations (but not necessarily the scope) of this paper:

1. We did not discuss the validity of the network approach for various disorders such as schizophrenia, autism, ADHD, social anxiety disorder, or personality disorders on which network literature has been published very recently (e.g., Heeren & McNally, 2016; Wright & Simms, 2016; for a review, see Fried, van Borkulo, Cramer, et al., 2016).
2. We omitted the important discussion on how between-person networks relate to within-person networks (E. H. Bos & de Jonge, 2014; E. H. Bos & Wanders, 2016) because it is largely unresolved: Does the network of 500 people with a given disorder relate to the way the 500 individual networks look like? Multilevel VAR models that allow the estimation of both idiographic and nomothetic networks may offer possibilities to explore this question (F. M. Bos et al., 2017).
3. Network analysis relies on estimating associations among individual symptoms or emotions. This means we model connections among single-item indicators that may have substantial measurement error—something that is less of a problem in latent variable models. Future studies should aim to investigate the *reliability* of these single-item measurements used in network analysis. From a measurement perspective, it may be advantageous to query participants about a given item (e.g., a symptom) using multiple questions and then model them as latent variables in networks. The Inventory of Depression and Anxiety Symptoms provides a good example for a rating scale that could be used for such a purpose: It uses multiple questions per symptom domain and for instance taps suicidal tendencies with six different items (Watson et al., 2007) that could be combined into a latent variable for subsequent analyses. Another possibility is to incorporate measurement error in network models (Schuurman, Houtveen, & Hamaker, 2015) or to estimate a network model on the residual covariance in an SEM framework (Epskamp, Rhemtulla, & Borsboom, 2016).
4. The temporal character of symptoms and emotions is unresolved, which is a crucial topic when choosing the sampling-scheme for within-subjects studies: How many time-points per day or week should one plan? Do symptoms or emotions evolve in a time frame of minutes, hours, or days? Is this time frame different for different items or associations? And could it be that the temporal association of A $\rightarrow$ B only appears after A has continuously occurred: One night of bad sleep may be sufficient to trigger fatigue, but one day of sad mood is likely not sufficient to trigger suicidal ideation. Thinking about these questions before designing studies is of crucial importance because modeling processes at the wrong time frame may lead to erroneous estimates of the associations. One promising development for VAR models that could help remediate this challenge is that we do not only investigate the lagged effects among variables but can also estimate an undirected contemporaneous network (what are the associations among variables in the last assessment period; Epskamp, van Borkulo et al., 2016). This contemporaneous network captures associations that occur at the same measurement point.

In sum, we have sketched the most pertinent challenges the network perspective currently faces. Dealing effectively with these challenges might propel this relatively novel perspective from its adolescence into adulthood. We deem it in the best interest of clinical psychology and psychiatry to try and meet these challenges because we believe that conceptualizing mental disorders as networks of interacting problems might offer an important inroad to understanding psychopathology. Given the number of young, gifted, and passionate researchers learning network analyses presently, we are optimistic that the network perspective is not far from a critical transition into a mature state.

Acknowledgments

We would like to express our sincerest thanks to Gary Brown, Kiara Timpano, Alex Millner, Donald Robinaugh, and Aidan Wright for feedback on prior versions of this manuscript. We also thank the anonymous reviewers for their invaluable feedback on the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

E.I.F. is supported by European Research Council (ERC) Consolidator Grant 647209 A.O.J.C. by the Netherlands Organization for Scientific Research (NWO) Innovational Research Grant VENI (451-14-002).

Supplemental Material

Additional supporting information may be found at http://journals.sagepub.com/doi/suppl/10.1177/17456916177058.

Notes

1. With a common cause for mental disorders, we mean a variable such as a traumatic brain injury that causally explains at least a *large proportion of the shared variance* among most symptoms of a disorder. Of note, a symptom in a network (e.g., Fig. 1 left) may predict numerous other symptoms: Concentration problems and fatigue may both be predicted by insomnia. We do not understand insomnia in this case as a common cause, because it is unlikely to explain the majority of the covariance among concentration problems and fatigue.

2. Note that a similar concept in the sociological literature is called *structural equivalence* (Lorraine & White, 1971).

3. PMRFs are used to estimate regularized partial correlation networks. For binary data, the appropriate PMRF is the Ising Model (van Borkulo et al., 2014), which can be readily estimated with the R-package *IsingFit* (van Borkulo & Epskamp, 2014). The PRMF for metric data is called Gaussian Graphical Model (Lauritzen, 1996) and can be estimated via the R-package *qgraph* (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). For mixed variables, so-called Mixed Graphical Models are available that can be estimated via the R-package *mgm* (Haslbeck & Waldorp, 2015).

4. PMRFs are often regularized using the "least absolute shrinkage and selection" (LASSO; Friedman et al., 2008) that shrinks edges and sets small edges exactly to zero, meaning the estimated network is a sparse/parsimonious network: Only a few edges in the network are used to explain the correlations among items. Details on regularized partial correlation networks in psychopathology are available elsewhere (Epskamp & Fried, 2016). Of note, although regularization implies that surviving edges are likely nonzero, it does not ensure that these edges are reliably estimated (Epskamp et al., 2017).

5. The R-package *graphicalVAR* (https://cran.r-project.org/web/packages/graphicalVAR) allows the estimation of regularized VAR models, but only for $n = 1$ networks.

6. https://cran.r-project.org/web/packages/gimme/index.html

7. With k nodes, a PRMF results in $(k * k - 1)/2$ estimated edges; additionally, k threshold parameters are estimated.

8. We estimated the network structure using a Gaussian Graphical Model with the graphical lasso regularization, as implemented in the R-package *qgraph* (Epskamp et al., 2012). Edges depict estimations of regularized partial correlations (Epskamp & Fried, 2016). Data come from the study of Hien et al. (2009) and are publicly available at the Data Share Website of the National Institute on Drug Abuse (https://datashare.nida.nih.gov/study/nida-ctn-0015). Syntax for the analyses is available in the Supplementary Materials.

9. Note that data for both networks come from the same dataset (https://datashare.nida.nih.gov/study/nida-ctn-0015); we split participants in two groups of $n = 180$ and $n = 179$ (total $n = 359$). We performed the split only once; the syntax is available in the Supplementary Materials.

10. Note that bootstrapped CIs are difficult to interpret in *regularized* partial correlation networks such as the one we estimated here, and the conclusions drawn regarding differences between edge weights should be done with care because the 95% CIs cannot be understood as a significance test for differences. A detailed explanation of this is available elsewhere (Epskamp et al., 2017).

References

Abbott, A. (2016). US mental-health chief: Psychiatry must get serious about mathematics. *Nature*, *539*, 18–19. doi:10.1038/nature.2016.20893

Achermann, P. (2004). The two-process model of sleep regulation revisited. *Aviation, Space, and Environmental Medicine*, *75*(3 Suppl), A37–A43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15018264

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author. Retrieved from http://books.google.com/books?id=S8p3lwEACAAJ&pgis=1

Armour, C., Fried, E. I., Deserno, M. K., Tsai, J., Pietrzak, R. H., Southwick, S. M., & Pietrzak, R. H. (2016). A network analysis of DSM-5 posttraumatic stress disorder symptoms and clinically relevant correlates in a national sample of U.S. military veterans. *Journal of Anxiety Disorders*, *45*, 49–59. doi:10.1016/j.janxdis.2016.11.008

Babyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, *66*, 411–421. doi:10.1097/01.psy.0000127692.23278.a9

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

Beck, A. T., Rush, A. J., Shaw, F. S., & Emery, G. (1979). *Cognitive therapy of depression*. New York, NY: Guilford Press.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533. doi:10.1038/483531a

Borbély, A. A., & Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *Journal of Biological Rhythms*, *14*, 559–570. doi:10.1177/074873099129000894

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

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

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. doi:10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE*, *6*(11), e27407. doi:10.1371/journal.pone.0027407

Bos, E. H., & de Jonge, P. (2014). "Critical slowing down in depression" is a great idea that still needs empirical proof. *Proceedings of the National Academy of Sciences, USA*, *111*, E878–E878. doi:10.1073/pnas.1323672111

Bos, E. H., & Wanders, R. B. K. (2016). Group-level symptom networks in depression. *JAMA Psychiatry, 73,* 411. doi:10.1001/jamapsychiatry.2015.3103

Bos, F. M., Snippe, E., de Vos, S., Hartmann, J. A., Simons, C. J. P., van der Kriege, L., . . . Wichers, M. (2017). Can we jump from cross-sectional to dynamic interpretations of networks? Implications for the Network Perspective in Psychiatry. *Psychotherapy and Psychosomatics*, *86*(3), 175–177. doi:10.1159/000453583

Boschloo, L., Schoevers, R. A., van Borkulo, C. D., Borsboom, D., Albertine, J., & Oldehinkel, A. J. (2016). The network structure of psychopathology in a community sample of preadolescents. *Journal of Abnormal Psychology*, *125*, 599–606. doi:10.1037/abn0000150

Boschloo, L., van Borkulo, C. D., Borsboom, D., & Schoevers, R. A. (2016). A prospective study on how symptoms in a network predict the onset of depression. *Psychotherapy and Psychosomatics*, *85*, 183–184. doi:10.1159/000442001

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*(9), e0137621. doi:10.1371/journal.pone.0137621

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2016). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods*. doi:10.1037/met0000085

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*, *45*, 747–57. doi:10.1017/S0033291714001809

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., . . . Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, *8*(4), e60188. doi:10.1371/journal.pone.0060188

Brown, G. W., & Harris, T. O. (1989). Depression. In G. W. Brown & T. O. Harris (Eds.), *Life events and illness* (pp. 49–93). New York, NY: Guilford Press.

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016a). Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics. *Frontiers in Psychology*, *7*(1540), 1–15. doi:10.3389/fpsyg.2016.01540

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016b). Using raw VAR regression coefficients to build networks can be misleading. *Multivariate Behavioral Research*, *51*(23), 330–344. doi:10.1080/00273171.2016.1150151

Cavanagh, S. R., Fitzgerald, E. J., & Urry, H. L. (2014). Emotion reactivity and regulation are associated with psychological functioning following the 2011 earthquake, tsunami, and nuclear crisis in Japan. *Emotion*, *14*, 235–240. doi:10.1037/a0035422

Chatfield, C. (2003). *The analysis of time series: An introduction* (5th ed.). Boca Raton, FL: Chapman & Hall/CRC.

Costantini, G. (2014). *Network analysis: A new perspective on personality psychology*. Unpublished doctoral dissertation, University of Milan-Bicocca, Milan, Italy.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (2014). State of the {aRt} personality research: A tutorial on network analysis of personality data in {R}. *Journal of Research in Personality*, *54*, 13–29. doi:10.1016/j.jrp.2014.07.003

Costello, C. (1993). The advantages of the symptom approach to depression. In C. Costello (Ed.), *Symptoms of depression* (pp. 1–21). New York, NY: John Wiley.

Cramer, A. O. J., Borsboom, D., Aggen, S. H., & Kendler, K. S. (2013). The pathoplasticity of dysphoric episodes: Differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychological Medicine*, *42*, 957–965. doi:10.1017/S003329171100211X

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*(12), e0167490. doi:10.1371/journal.pone.0167490

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., . . . Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*, 414–431. doi:10.1002/per.1866

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *The Behavioral and Brain Sciences*, *33*(2–3), 137–150. doi:10.1017/S0140525X09991567

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*, 1–10. doi:10.3389/fpsyg.2015.01896

Ehlers, A., & Clark, D. M. (2000). A cognitive model of post-traumatic stress disorder. *Behaviour Research and Therapy*, *38*, 319–345.

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology, 40,* 31–53. http://doi.org/10.1146/annurev-soc-071913-043455

Epskamp, S. (2017). *Network psychometrics*. Unpublished doctoral dissertation, University of Amsterdam, The Netherlands. Retrieved from http://hdl.handle.net/11245.1/a76273c6-6abc-4cc7-a2e9-3b5f1ae3c29e

Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *Behavioral Research Methods*. doi:10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18.

Epskamp, S., & Fried, E. I. (2016). A tutorial on regularized partial correlation networks. *Arxiv Preprint (ID 1607.01367)*, 1–21. Retrieved from https://arxiv.org/abs/1607.01367

Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (in press). Network psychometrics. In P. Irwing, D. Hughes,

& T. Booth (Eds.), *Handbook of psychometrics*. New York, NY: Wiley.

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2016). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*. Retrieved from http://arxiv.org/abs/1605.09288

Fava, M., Rush, A. J., Alpert, J. E., Balasubramani, G. K., Wisniewski, S. R., Carmin, C. N., . . . Trivedi, M. H. (2008). Difference in treatment outcome in outpatients with anxious versus nonanxious depression: A STAR*D report. *The American Journal of Psychiatry*, *165*, 342–351. doi:10.1176/appi.ajp.2007.06111868

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology*, *83*, 825–836.

Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. (2015). The economics of reproducibility in preclinical research. *PLoS Biology*, *13*(6), e1002165. doi:10.1371/journal.pbio.1002165

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, *6*(306), 1–11. doi:10.3389/fpsyg.2015.00309

Fried, E. I., Bockting, C. L. H., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, A. O. J., . . . Stroebe, M. (2015). From loss to loneliness: The relationship between bereavement and depressive symptoms. *Journal of Abnormal Psychology*, *124*, 256–265. doi:10.1037/abn0000028

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

Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS ONE*, *9*(2), e90311. doi:10.1371/journal.pone.0090311

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

Fried, E. I., & Nesse, R. M. (2015b). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(72), 1–11. doi:10.1186/s12916-015-0325-4

Fried, E. I., Nesse, R. M., Guille, C., & Sen, S. (2015). The differential influence of life stress on individual symptoms of depression. *Acta Psychiatrica Scandinavica*, *131*, 465–471. doi:10.1111/acps.12395

Fried, E. I., Tuerlinckx, F., & Borsboom, D. (2014). Mental health: More than neurobiology. *Nature*, *508*, 458. doi:10.1038/508458c

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2016). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *1*, 1–32. doi:10.1007/s00127-016-1319-z

Fried, E. I., van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time . . . or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*, *28*, 1354–1367. doi:10.1037/pas0000275

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*, 432–441. doi:10.1093/biostatistics/kxm045

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science*, *8*, 651–662. doi:10.1177/1745691613504115

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*, 310–319. doi:10.1016/j.neuroimage.2012.06.026

Grados, M. A., Specht, M. W., Sung, H.- M., & Fortune, D. (2013). Glutamate drugs and pharmacogenetics of OCD: A pathway-based exploratory approach. *Expert Opinion on Drug Discovery*, *8*(12), 1515–1527. doi:10.1517/17460441.2013.845553

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438. doi:10.2307/1912791

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis* (J. A. Hagenaars & A. L. McCutcheon, Eds.). Cambridge, UK: Cambridge University Press.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=495331&tool=pmcentrez&rendertype=abstract

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology*, *1*, 293–319. doi:10.1146/annurev.clinpsy.1.102803.143938

Haslbeck, J. M. B., & Waldorp, L. J. (2015). Structure estimation for mixed graphical models in high-dimensional data. *Journal of Statistical Software, VV*(January). Retrieved from http://arxiv.org/abs/1510.05677

Heeren, A., & McNally, R. J. (2016). An integrative network approach to social anxiety disorder: The complex dynamic interplay among attentional bias for threat, attentional control, and symptoms. *Journal of Anxiety Disorders*, *42*, 95–104. doi:10.1016/j.janxdis.2016.06.009

Hien, D. A., Wells, E. A., Jiang, H., Suarez-Morales, L., Campbell, A. N. C., Cohen, L. R., . . . Nunes, E. V. (2009). Multi-site randomized trial of behavioral interventions for women with co-occurring PTSD and substance use disorders. *Journal of Consulting and Clinical Psychology*, *77*, 607–619. doi:10.1037/a0016227

Horton, M., & Perry, A. E. (2016). Screening for depression in primary care: A Rasch analysis of the PHQ-9. *BJPsych Bulletin*, *40*, 237–243. doi:10.1192/pb.bp.114.050294

Insel, T. R. (2013). *Transforming diagnosis*. Retrieved from http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml

Insel, T. R., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, *167*, 748–751. doi:10.1176/appi.ajp.2010.09091379

Judd, L. L., Schettler, P. J., Coryell, W., Akiskal, H. S., & Fiedorowicz, J. G. (2013). Overt irritability/anger in unipolar

major depressive episodes: Past and current characteristics and implications for long-term course. *JAMA Psychiatry*, *70*, 1171–1180. doi:10.1001/jamapsychiatry.2013.1957

Keller, M. C., Neale, M. C., & Kendler, K. S. (2007). Association of different adverse life events with distinct patterns of depressive symptoms. *The American Journal of Psychiatry*, *164*, 1521–1529. doi:10.1176/appi.ajp.2007.06091564

Keller, M. C., & Nesse, R. M. (2005). Is low mood an adaptation? Evidence for subtypes with symptoms that match precipitants. *Journal of Affective Disorders*, *86*(1), 27–35. doi:10.1016/j.jad.2004.12.005

Kendler, K. S., Kuhn, J., & Prescott, C. A. (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *American Journal of Psychiatry*, *161*, 631–636. Retrieved from http://ajp.psychiatryonline.org/doi/abs/10.1176/appi.ajp.161.4.631

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

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., . . . Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *The Journal of the American Medical Association*, *289*, 3095–3105. doi:10.1001/jama.289.23.3095

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & 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*, 617–627. doi:10.1001/archpsyc.62.6.617

Knefel, M., Tran, U. S., & Lueger-Schuster, B. (2016). The association of posttraumatic stress disorder, complex posttraumatic stress disorder, and borderline personality disorder from a network analytical perspective. *Journal of Anxiety Disorders*, *43*, 70–78. doi:10.1016/j.janxdis.2016.09.002

Koenders, M. A., de Kleijn, R., Giltay, E. J., Elzinga, B. M., Spinhoven, P., & Spijker, A. T. (2015). A network approach to bipolar symptomatology in patients with different course types. *PLoS ONE*, *10*(10), e0141420. doi:10.1371/journal.pone.0141420

Kramer, I., Simons, C., Hartmann, J. A., Menne-Lothmann, C., Viechtbauer, W., Peeters, F., . . . Wichers, M. (2014). A therapeutic application of the experience sampling method in the treatment of depression: A randomized controlled trial. *World Psychiatry*, *13*, 68–77. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/wps.20090/full

Kupfer, D. J., First, M. B., & Regier, D. A. (2002). *A research agenda for DSM V*. Washington, DC: American Psychiatric Association.

Lauritzen, S. L. (1996). *Graphical models*. Oxford, UK: Clarendon Press.

Lorraine, F., & White, H. C. (1971). Structural equivalence of individuals in social networks. *The Journal of Mathematical Sociology*, *1*, 49–80.

Lutkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin, Germany: Springer.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*, 1–25.

McNally, R. J. (2012). The ontology of posttraumatic stress disorder: Natural kind, social construction, or causal system? *Clinical Psychology: Science and Practice*, *19*, 220–228. doi:10.1111/cpsp.12001

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

McNally, R. J., Mair, P., Mugno, B., & Riemann, B. (2017). Co-morbid obsessive-compulsive disorder and depression: A Bayesian network approach. *Psychological Medicine*, *47*(7), 1204–1214. doi:10.1017/S0033291716003287

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

Mitchell, K. S., Wolf, E. J., Bovin, M. J., Lee, L. O., Green, J. D., Raymond, C., . . . Marx, B. P. (2017). *Journal of Abnormal Psychology* network models of DSM – 5 posttraumatic stress disorder. *Journal of Abnormal Psychology, 126,* 355–366.

Molenaar, P. (2003). *State space techniques in structural equation modeling: Transformation of latent variables in and out of latent variable models*. Retrieved from http://www.hhdev.psu.edu/hdfs/faculty/docs/StateSpaceTechniques.pdf

Molenaar, P. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*. doi:10.1207/s15366359mea0204

Molenaar, P., van Rijn, P., & Hamaker, E. (2007). A new class of SEM model equivalences and its implications. In M. Wenger & S. Boker (Eds.), *Data analytic techniques for dynamical systems* (pp. 189–211). Mahwah, NJ: Erlbaum.

Nickerson, A., Bryant, R. A., Schnyder, U., Schick, M., Mueller, J., & Morina, N. (2015). Emotion dysregulation mediates the relationship between trauma exposure, post-migration living difficulties and psychological outcomes in traumatized refugees. *Journal of Affective Disorders*, *173*, 185–192. doi:10.1016/j.jad.2014.10.043

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*, 452–462. doi:10.1037/a0036068

Oldham, M. C., Konopka, G., Iwamoto, K., Langfelder, P., Horvath, S., & Geschwind, D. H. (2008). Functional organization of the transcriptome in human brain. *Nature Neuroscience*, *11*, 1271–1282. doi:10.1038/nn.2207l

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716–aac4716. doi:10.1126/science.aac4716

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*, 245–251. doi:10.1016/j.socnet.2010.03.006

Overall, J., & Porterfield, J. (1963). Powered vector method of factor analysis. *Psychometrika*, *28*, 415–422.

Parker, G. (2005). Beyond major depression. *Psychological Medicine*, *35*, 467–474. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15856717

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., . . . Gotlib, I. H. (2015). Emotion-network

density in major depressive disorder. *Clinical Psychological Science*, *3*, 292–300. doi:10.1177/2167702614540645

Persons, J. B. (1986). The advantages of studying psychological phenomena rather than psychiatric diagnoses. *The American Psychologist*, *41*, 1252–1260. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3813184

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401. doi:10.1177/014662167700100306

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. *The American Journal of Psychiatry*, *170*, 59–70. doi:10.1176/appi.ajp.2012.12070999

Rhemtulla, M., Fried, E. I., Aggen, S. H., Tuerlinckx, F., Kendler, K. S., & Borsboom, D. (2016). Network analysis of substance abuse and dependence symptoms. *Drug and Alcohol Dependence*, *161*, 230–237. doi:10.1016/j.drugalcdep.2016.02.005

Riaza Bermudo-Soriano, C., Perez-Rodriguez, M. M., Vaquero-Lorenzo, C., & Baca-Garcia, E. (2012). New perspectives in glutamate and anxiety. *Pharmacology, Biochemistry, and Behavior*, *100*, 752–774. doi:10.1016/j.pbb.2011.04.010

Robinaugh, D. J., Leblanc, N. J., Vuletich, H. A., & McNally, R. J. (2014). Network analysis of persistent complex bereavement disorder in conjugally bereaved adults. *Journal of Abnormal Psychology*, *123*, 510–22. doi:10.1037/abn0000002

Sanacora, G., Treccani, G., & Popoli, M. (2012). Towards a glutamate hypothesis of depression: An emerging frontier of neuropsychopharmacology for mood disorders. *Neuropharmacology*, *62*(1), 63–77. doi:10.1016/j.neuropharm.2011.07.036

Schmaal, L., Veltman, D. J., van Erp, T. G. M., Sämann, P. G., Frodl, T., Jahanshad, N., . . . Hibar, D. P. (2015). Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA Major Depressive Disorder working group. *Molecular Psychiatry*. doi:10.1038/mp.2015.69

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*(1), 43–53. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0732118X1100016X

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*, 206–221. doi:10.1037/met0000062

Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology*, *6*, 1038. doi:10.3389/fpsyg.2015.01038

Schwartz, T. L., Sachdeva, S., & Stahl, S. M. (2012). Glutamate neurocircuitry: Theoretical underpinnings in schizophrenia. *Frontiers in Pharmacology*, *3*, 195. doi:10.3389/fphar.2012.00195

Sen, S., Kranzler, H. R., Krystal, J. H., Speller, H., Chan, G., Gelernter, J., & Guille, C. (2010). A prospective cohort study investigating factors associated with depression during medical internship. *Archives of General Psychiatry*, *67*, 557–565. doi:10.1001/archgenpsychiatry.2010.41

Smeets, F., Lataster, T., Viechtbauer, W., & Delespaul, P. (2014). Evidence that environmental and genetic risks for psychotic disorder may operate by impacting on connections between core symptoms of perceptual alteration and delusional ideation. *Schizophrenia Bulletin*. doi:10.1093/schbul/sbu122

Snaith, P. (1993). What do depression rating scales measure? *The British Journal of Psychiatry*, *163*, 293–298. doi:10.1192/bjp.163.3.293

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*, 496–500. doi:10.1126/science.1227079

Symptom. (2015). Retrieved from http://www.merriam-webster.com/dictionary/symptom

Symptom. (2016). Retrieved from http://dictionary.cambridge.org/dictionary/english/symptom

ten Have, M., Lamers, F., Wardenaar, K., Beekman, A., de Jonge, P., van Dorsselaer, S., . . . de Graaf, R. (2016). The identification of symptom-based subtypes of depression: A nationally representative cohort study. *Journal of Affective Disorders*, *190*, 395–406. doi:10.1016/j.jad.2015.10.040

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B: Methodological*, *58*, 267–288.

Tio, P., Epskamp, S., Noordhof, A., & Borsboom, D. (2016). Mapping the manuals of madness: Comparing the ICD-10 and DSM-IV-TR using a network approach. *International Journal of Methods in Psychiatric Research*. doi:10.1002/mpr.1503

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, *4*(5918), 1–10. doi:10.1038/srep05918

van Borkulo, C. D., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry*, *72*, 1219. doi:10.1001/jamapsychiatry.2015.2079

van Borkulo, C. D., & Epskamp, S. (2014). *IsingFit*. Retrieved from http://cran.r-project.org/web/packages/IsingFit/index.html

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., . . . Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences, USA*, *111*(1), 87–92. doi:10.1073/pnas.1312114110

van der Maas, H. L. J., Dolan, C. V, Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861. doi:10.1037/0033-295X.113.4.842

van Os, J., & Jones, P. B. (2001). Neuroticism as a risk factor for schizophrenia. *Psychological Medicine*, *31*(6), 1129–1134. doi:10.1017/S0033291701004044

Van Rooij, A. J., Van Looy, J., & Billieux, J. (2016). Internet gaming disorder as a formative construct: Implications

for conceptualization and measurement. *Psychiatry and Clinical Neurosciences*. doi:10.1111/pcn.12404

Wardenaar, K. J., Monden, R., Conradi, H. J., & de Jonge, P. (2015). Symptom-specific course trajectories and their determinants in primary care patients with Major Depressive Disorder: Evidence for two etiologically distinct prototypes. *Journal of Affective Disorders*, *179*, 38–46. doi:10.1016/j.jad.2015.03.029

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Structural Analysis in the Social Sciences, Vol. 8). Cambridge, UK: Cambridge University Press. doi:10.1093/infdis/jir193

Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E. A., . . . Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment*, *19*, 253–268. doi:10.1037/1040-3590.19.3.253

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*, 114–116. doi:10.1159/000441458

Wichers, M., Wigman, J. T. W., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review*, *7*, 362–367. doi:10.1177/1754073915590623

Wright, A. G. C., & Simms, L. J. (2016). Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology*, *125*(4), 1–16. doi:10.1037/abn0000169

Yarkoni, T., & Westfall, J. (in press). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*.

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*, Article 17. doi:10.2202/1544-6115.1128

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2014). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry*. doi:10.1016/j.comppsych.2014.09.007