Behaviour Research and Therapy 149 (2022) 104011

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Psychopathological networks: Theory, methods and practice*

Laura F. Bringmann [a,b,*], Casper Albers [b], Claudi Bockting [c,d], Denny Borsboom [e], Eva Ceulemans [f], Angélique Cramer [g], Sacha Epskamp [c,e], Markus I. Eronen [h], Ellen Hamaker [i], Peter Kuppens [f], Wolfgang Lutz [j], Richard J. McNally [k], Peter Molenaar [l], Pia Tio [e,m], Manuel C. Voelkle [n], Marieke Wichers [a]

[a] University of Groningen, University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), P.O. Box 30.001 (CC72), 9700 RB, Groningen, the Netherlands
[b] University of Groningen, Faculty of Behavioural and Social Sciences, Department of Psychometrics and Statistics, Grote Kruisstraat 2/1, 9712 TS, Groningen, the Netherlands
[c] Centre for Urban Mental Health, University of Amsterdam, Amsterdam, the Netherlands
[d] Department of Psychiatry, Amsterdam UMC, Location AMC, University of Amsterdam, Amsterdam, the Netherlands
[e] Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands
[f] KU Leuven, Faculty of Psychology and Educational Sciences, Leuven, Belgium
[g] RIVM National Institute for Public Health and the Environment, the Netherlands
[h] Department of Theoretical Philosophy, University of Groningen, the Netherlands
[i] Department of Methodology and Statistics, Utrecht University, the Netherlands
[j] Department of Psychology, University of Trier, Germany
[k] Department of Psychology, Harvard University, Cambridge, USA
[l] Department of Human Development and Family Studies, The Pennsylvania State University, USA
[m] Department of Methodology and Statistics, Tilburg University, Tilburg, the Netherlands
[n] Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

ARTICLE INFO

Keywords:
Psychological networks
Time series models
Clinical theory
Clinical methodology
Psychometrics

ABSTRACT

In recent years, network approaches to psychopathology have sparked much debate and have had a significant impact on how mental disorders are perceived in the field of clinical psychology. However, there are many important challenges in moving from theory to empirical research and clinical practice and vice versa. Therefore, in this article, we bring together different points of view on psychological networks by methodologists and clinicians to give a critical overview on these challenges, and to present an agenda for addressing these challenges. In contrast to previous reviews, we especially focus on methodological issues related to temporal networks. This includes topics such as selecting and assessing the quality of the nodes in the network, distinguishing between- and within-person effects in networks, relating items that are measured at different time scales, and dealing with changes in network structures. These issues are not only important for researchers using network models on empirical data, but also for clinicians, who are increasingly likely to encounter (person-specific) networks in the consulting room.

Since its introduction a decade ago, the network approach to psychopathology (Borsboom, 2008; Borsboom & Cramer, 2013; Cramer, Waldorp, van der Maas, & Borsboom, 2010) has caught the attention of many psychopathology researchers and clinicians (Robinaugh, Hoekstra, Toner, & Borsboom, 2020). In this paper, we discuss possibilities and challenges in applying networks to clinical research and practice.

The introduction of the network approach centered on the substantive claim that mental disorders emerge from causal interactions among symptoms (Borsboom & Cramer, 2013). This idea was concordant with conventional wisdom among cognitive-behavior therapists, who have a longstanding tradition (e.g., Beck, 1967; Ellis, 1956) of viewing mental disorders as arising from systematic causal connections between

* The authors are in alphabetic order, except for the first author.

* Corresponding author. University of Groningen, Faculty of Behavioural and Social Sciences, Department of Psychometrics and Statistics, Grote Kruisstraat 2/1, 9712 TS, Groningen, the Netherlands.
E-mail address: l.f.bringmann@rug.nl (L.F. Bringmann).

https://doi.org/10.1016/j.brat.2021.104011
Received 27 February 2021; Received in revised form 5 November 2021; Accepted 27 November 2021
Available online 1 December 2021
0005-7967/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

cognitions (e.g., negative beliefs about oneself), affect states (e.g., sad mood, feelings of worthlessness), and behavior (e.g., withdrawal from social life, suicide attempts). Framing psychological disorders as networks thus aligned naturally with the dominant mode of thinking of cognitive-behavioral theorists.

In addition, by focusing on psychological symptoms and their interactions, the network approach provides an alternative to monocausal, often brain-based or biology-oriented approaches to mental disorders, which assume that mental disorders have a clear (biological) root cause analogous to medical diseases (Borsboom, Cramer, & Kalis, 2019). As such, the approach draws attention to mental states rather than their neural basis, and creates a conceptual space in which cognitions can play autonomous causal roles in the genesis of mental disorders (Oude Maatman, 2020; Kalis & Borsboom, 2020). This aligns naturally with the notion that interactions among mental states, such as Beck's (1967) negative triad (beliefs about oneself, others and the world), and cognitive representations, such as schemata in anxiety disorders (Beck, 1985, 1988; McNally, 1990, 1994) and depressive disorders (Brouwer et al., 2019, 2020), are central to the genesis and maintenance of mental disorders. In addition, it allows for a genuine role of concrete, everyday problems that are based in the social and physical environment (e.g., debts, housing problems) rather than the result of abnormalities in the brain (e.g., Insel & Cuthbert, 2015), which also matches psychology's long-standing emphasis on the essential role of the situational context in shaping behavior (Mischel, 1968; Mischel & Shoda, 1995).

Following the introduction of the network approach as a theoretical framework, a host of methodological techniques was developed to infer network structures from data in a variety of research designs (Epskamp, Borsboom, & Fried, 2018; van Borkulo et al., 2014; Bringmann et al., 2013; Epskamp, 2020). Although the network approach is not tied to any particular statistical model, and may in fact not involve fitting statistical models at all (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Borsboom & Cramer, 2013; Wittenborn, Rahmandad, Rick, & Hosseinichimeh, 2016; Ruzzano, Borsboom, & Geurts, 2015; Deserno et al., 2020), the use of statistical models quickly became the most popular implementation of the network approach (Robinaugh et al., 2020). Within this statistical approach, two particular modeling traditions have guided the majority of attempts to infer psychological network structures from data.

The first comprises the estimation of conditional association networks based on cross-sectional data gathered from a large number of individuals (e.g., partial correlation networks, if data are multivariate normal; Epskamp & Fried, 2018).[1] In this kind of data, symptoms (i.e., the nodes in the network) are variables assessed at a single assessment occasion, and links between these represent conditional associations between these variables (van Borkulo et al., 2014; Epskamp & Fried, 2018). The most common representation of networks based on cross-sectional data is through an undirected network, in which the absence of a link between two nodes represents the fact that the corresponding variables are conditionally independent given the other variables in the network, although extensions of this approach to directed (causal) relations are possible (McNally, Mair, Mugno, & Riemann, 2017; Richardson & Spirtes, 2002).

The second modeling tradition is based on the idea that network structures may be assessed based on repeated measures that characterize the dynamic evolution of symptoms within the individual over time (i.e., time series; Molenaar, 2004; Hamaker, 2012), possibly extended to a situation in which time series of multiple individuals are assessed simultaneously (Bringmann et al., 2013). In this paper, we refer to such networks as *temporal networks*. The most often used statistical model for inferring these kinds of networks is the vector autoregressive (VAR) model (Bringmann et al., 2013; Epskamp, 2020; Beltz & Gates, 2017). Typically, an arrow from node A to node B in a VAR-based network reflects the variance that can be predicted in variable B from variable A, while controlling for all other nodes in the model. Because of the explicit representation of time order, in contrast to cross-sectional networks VAR-based networks always involve directed edges (arrows). VAR-based network models may also include self-loops (the predictive effect of a symptom on itself from one time point to the next) and representations of contemporaneous effects (i.e., the concurrent partial correlations between variables over time; Epskamp, 2020; Beltz & Gates, 2017).

The determination of a network structure, using methods such as the aforementioned conditional association networks or VAR models, is the first step in network analysis and ideally provides a stepping stone to a second order analysis of the network structure itself (Newman, 2018). In this second order analysis, researchers focus on the topology of the network (e.g., are the edges in a network distributed randomly or do they form a structure like a small world structure?; Watts & Strogatz, 1998) and the role of specific nodes in that topology. The methods of network science for instance may enable detection of network properties such as centrality, that is, which nodes in the network are the most central (interconnected) symptoms potentially sustaining an episode of disorder (Blanken et al., 2018; Rodebaugh et al., 2018). The idea is that researchers and clinicians, using network models, can identify the structurally important symptoms, whose deactivation may hasten a patient's recovery. Moreover, the network approach provides an intuitive story to individuals with mental health problems, and may thereby facilitate their communication with clinicians and lead to insights about mechanisms of their symptoms (Bak, Drukker, Hasmi, & van Os, 2016; Bos, Snippe, Bruggeman, Wichers, & van der Krieke, 2019; Kroeze et al., 2016). These ideas have taken hold so quickly that several authors have voiced criticisms of these ideas; scholars have identified thorny issues in the interpretation of correlation structures as networks (Bringmann, 2021; Bringmann et al., 2019; Dablander & Hinne, 2019), emphasized the limited evidence for the usefulness of topological information (Rodebaugh et al., 2018), and pointed out possible confounds of centralities (Hallquist, Wright, & Molenaar, 2021).

With a rapid increase of researchers either using network analyses or developing reliable methods to estimate such psychopathological networks from empirical data (Robinaugh et al., 2020), clinicians are eager to apply network techniques in personalized assessment to improve treatment outcomes. Although there certainly is room for optimism about the network approach as one road into personalized clinical research and modeling, it is also important to address some of the methodological and practical challenges we face. Therefore, in this article, we bring together different points of view on psychological networks by methodologists, clinicians, and philosophers, to give a critical overview on these challenges, and to present an agenda for addressing these challenges in clinical research and other fields of psychology. In contrast to previous reviews (see Borsboom et al., 2021; Fried & Cramer, 2017; Fried et al., 2017; McNally, 2021; Robinaugh et al., 2020; Wichers, Riese, Hodges, Snippe, & Bos, 2021), we especially focus on methodological issues related to temporal networks, which are not only important for researchers using network models on empirical data, but also for clinicians, who are increasingly likely to encounter such person-specific networks in the consulting room.

The structure of this paper is as follows. As mentioned above, the network approach is closely aligned with clinical therapy and the idea that symptoms and cognitions influence each other, but it is a large step from this idea to actually selecting the nodes to include in a network. Therefore, we start by zooming in on this issue by introducing and explaining the notion of *node validity*. Next, we discuss in detail how *between-person effects* (e.g., how individuals in different groups react to a drug) and *within-person effects* (e.g., how the effect of a drug unfolds in an individual over time) can and cannot be distinguished in different

[1] In this article, we mainly focus on continuous data. For dichotomous data, Ising models have been popular in the recent network literature. The general term "Markov random fields" can be used to cover both partial correlation networks and Ising models Epskamp, Maris, Waldorp, & Borsboom (2018).

kinds of networks, as knowing how to pull these effects apart is crucial for applying networks in clinical research or practice.

Then we turn to the problem of relating items in a network that are measured at *different time scales*, such as sleep and mood states. As researchers and clinicians increasingly use temporal networks based on time-series data, this issue is becoming very pressing, but has not received much attention in the literature yet. We discuss it thoroughly and provide important starting points for accommodating and dealing with differences in time scale. Furthermore, an important part of the network approach is to study changes in networks, for example when individuals transition into an episode of depression. Accordingly, we go through latest techniques of dealing with *non-stationarity*, that is, changes in network structure over time.

These topics are all focused on how to infer and analyze networks. However, in clinical practice, there is also great interest in network-based interventions, or more generally, how to turn the ideas of the network approach into clinically useful applications. Therefore, we discuss *interventions and networks in clinical research and practice*: The question of how the network approach is applied in practice, and how its usefulness could be further studied. Another issue that is crucial for the success of the network approach in practice is the *reproducibility and replicability* of networks. This topic has been extensively discussed in psychology more generally, and in the context of cross-sectional networks, but we now consider it with regard to both cross-sectional and temporal networks. Finally, in section *testability and falsifiability,* we consider how to actually test the ideas of the network approach in empirical settings, and moving from data-driven to more theory-based approaches to inferring networks; issues which are crucial for improving the network approach and making it more applicable in practice.

Although this article concerns the network approach and we are therefore focusing on networks, most of the methods and models that are currently used in this context are also applied more generally in psychology. For this reason, this article can also be informative and useful for researchers who are working on cross-sectional, time-series or intensive longitudinal data in psychology and are interested in the topics covered here.

1. Node validity

An essential first step to take when clinical researchers and clinicians alike want to use a network model, or any model for that matter, involves the choice of which variables should be included in the model. In clinical practice, a starting point is often to discuss with the patient the main reasons for seeking help and what symptoms and problems are perceived as most burdensome. Network construction thus starts by defining the variables (i.e., nodes) in the network. How should we decide what elements to include in a model? What nodes will best illuminate psychopathological processes? What criteria should adjudicate variable inclusion? To address these questions, we introduce here the concept of *node validity*, which involves two steps. The first is *node selection* and refers to the adequacy of selecting appropriate variables as nodes in a network model. The second concerns *node assessment* and refers to the quality of the operationalizations used for selected variables.

Node selection. Taking the definition of networks broadly, one can include anything as a variable in a network. However, most studies have included only symptoms as nodes (e.g., suicidal thoughts; Robinaugh et al., 2020). This accords with the original idea of the network approach, where the components of psychopathology networks are symptoms currently implemented in diagnostic systems such as DSM-5 (Borsboom, 2017). In contrast, Jones, Heeren and McNally (2017) aim to identify a broader set of components, including theoretically relevant features missing from diagnostic criteria, such as catastrophic misinterpretation of bodily sensations in panic disorder or attentional biases in social anxiety. The Research Domain Criteria system (Insel et al., 2010) furnishes yet another set of putative components. Furthermore, especially in the field of dynamic networks involving time series data, the focus is on momentary states such as thoughts and affective states (e.g., feeling happy or down) which do not directly correspond to symptoms (e.g., because symptoms often identify contextual information, exclusion criteria, and time scales that are not usually sampled in time series data). However, symptoms do not arise out of the blue, but arguably result from dynamical interactions among experiences, feelings, and behaviors in daily life, which means a strong argument for including them can be made (Wichers, 2014). In general, these considerations suggest that although clinical symptoms and affect states are most often the variables of choice for networks in psychopathology, these are not the only possible building blocks for constructing a network.

Crucially, the clinical or theoretical hypothesis of a clinician or clinical researcher, often formulated together with the patient, plays a role in the choice of set of variables or nodes in the network. The chosen nodes should correspond as closely as possible to the hypothesis, which is articulated as unambiguously as possible. That is, ideally the chosen node set should be *minimally complete*, in other words, contain all nodes necessary to model the intended phenomena while excluding superfluous nodes. What "minimally complete" means will differ across contexts; for example, if one holds that depression arises from interactions among DSM symptoms, then taking these as the minimally complete node set would be defensible. But if one holds that these criteria omit important components, then additional measures would be necessary.

An additional important question for researchers to consider, especially when the interest is in causal hypotheses, is whether the nodes in the network are *sufficiently distinct* to act as separate nodes. For example, if items are mere psychometric indicators of a common cause, or (near) synonyms of each other, they should be modeled by using a latent variable model instead of a (causal) network model (Epskamp, Rhemtulla, & Borsboom, 2017; Wichers, Wigman, Bringmann, & de Jonge, 2017). For example, certain positive affect items of the PANAS (e.g., excited and enthusiastic) seem to be near synonyms. If deemed reflective of the construct *positive affect*, then latent variable modelling is justified. More precisely, to be sufficiently distinct to causally affect each other, nodes should be a) *separately identifiable* (i.e., at least in theory, they can be assessed independently of one another), and b) *independently manipulable* (i.e., at least in theory, one should be able to intervene on a node without intervening on other nodes; Woodward, 2015).

Several open problems follow from these ideas. One potentially important issue is whether variables should be included in the network itself or in the "external field" (a term borrowed from physics to identify any factors that impinge on the network from outside; Borsboom, 2017). From a clinical perspective, it is plausible that the context of daily life provides a vital backdrop for understanding and treating psychopathology (Wichers et al., 2011; Myin-Germeys et al., 2009). This can be, for example, daily activities undertaken or external stressors such as spousal loss (Fried et al., 2015). Although in some network studies such factors are included as nodes (e.g., stressful events; Fried et al., 2015; Bringmann et al., 2013, Slofstra et al., 2018), it has also been argued that these kinds of factors should rather be conceptualized as causal variables external to the network itself (Borsboom, 2017; Fried & Cramer, 2017). This also has consequences for modeling. Variables that are part of the network can all potentially influence each other and are therefore treated as endogenous variables (i.e., variables that are influenced by other variables in the model). The vector autoregressive model (VAR) is an example of a model where only such endogenous variables are included. On the other hand, if variables in the external field can influence network variables but not vice versa, they may be conceptualized as exogeneous variables (i.e., variables that receive no influences from other variables in the model; note, however, that there may not always be a clear-cut distinction between endogenous and exogeneous variables, see Antonakis, Bendahan, Jacquart, and Lalive (2010). A network with an external field (exogeneous variables), such as spousal

loss, could be modeled, for example, with a VARX model (where besides endogenous also exogeneous variables can be included; Hamaker & Dolan, 2009). Ultimately, the decision to include variables (e.g., the context of daily life) in the external field or the network itself should be informed by both theoretical and modeling considerations.

Importantly, researchers and clinicians, together with patients, not only gather data on psychological variables, but also biological data such as actigraphy measurements, electroencephalography (EEG) or social interactions. Should such variables of different levels (e.g., biological or social variables) be included in the same network with psychological variables (Wittenborn, Rahmandad, Rick, & Hosseinichimeh, 2016; Jones et al., 2017; Fried & Cramer, 2017), or should they appear in distinct networks studied in parallel with psychological networks (Blanken et al., 2021)? Inherently, this depends on what is meant by "levels" and how the variables at different levels are related (Eronen, 2021; Riese & Wichers, 2021). For example, if a biological variable is the underlying neural basis of a psychological variable, the two variables are conceptually overlapping, and therefore it may be problematic to include both in the same network as separate nodes as this would violate the distinctiveness criterion identified above. However, if levels are understood in a different sense, for example, in terms of time scales, it is easier to see how variables of different levels (e.g., personality traits and psychological symptoms) could be related to each other and represented in a broader network (Lunansky, van Borkulo, & Borsboom, 2020).

Node assessment. The above considerations highlight the need for delicate and precise assessment of nodes. However, in practice the above criteria are often difficult to assess. If clinicians or clinical researchers measure the variables *loss of libido*, *loss of appetite* and *loss of interest*, are they sufficiently distinct to warrant their inclusion as distinct nodes in a network, or are these all just manifestations of underlying anhedonia, their distinct names notwithstanding?

For approaching questions like these, the classic psychometric criteria such as reliability and validity form the fundamental basis for assessment (AERA, APA, & NCME, 2014). Ideally, researchers can assess reliability and provide evidence of validity that helps to demonstrate that the variables are indeed distinct and identifiable from one another. However, one issue on which network node assessment might diverge from the classic psychometric approach is the number of items. Although psychometric tradition would typically prescribe a sizeable number of items to measure one variable (e.g., Nunnally & Bernstein, 1994; Simms, 2008), practical considerations often make this unsuitable for network modeling.

For instance, psychopathological network models are increasingly fit on time-series data that allow for the study of temporal relations within a single individual, often a patient. For such data, only a limited number of questions can be asked per measurement occasion due to considerations of participant burden (Eisele, Vachon, et al., 2020). In practice, this often conflicts with the requirements of theoretical completeness. Technological developments can be potentially helpful by enabling automatic assessment of network nodes (see e.g., mobile sensing or wearables), thereby complementing questionnaire data while reducing participant burden (Harari, Müller, Aung, & Rentfrow, 2017; Insel, 2017; Jacobson, Lekkas, Huang, & Thomas, 2021). This, however, again raises the question how variables of differing nature or different levels, including those that cannot be measured automatically (e.g., suicidal thoughts), can be combined into one network.

A further issue in times-series research is that psychometrically validated scales or evidence-based guidelines on what to measure in these studies are still under development (e.g., Eisele, Lafit, et al., 2021). Currently, most researchers choose their own item set for their study, and little is therefore known about the psychometric measurement quality of the resulting data (Wright & Zimmermann, 2019). Moreover, due to the personalized nature of idiographic research (Wright & Woods, 2020), it may also be preferable to include different variables per person (von Klipstein, Riese, Servaas, & Schoevers, 2020). Finally, also here, the consideration of the timescale on which different nodes are measured is a crucial one, involving questions about how and when different timescales can be combined. This is discussed in more detail in the section on *Different time scales*.

2. Between- and within person effects

Clinical research has predominantly focused on between-person effects, such as different responses to treatment. The network approach, however, emphasizes the importance of individual heterogeneity, and thus, the importance of studying within-person effects, such as how the process of major depressive disorder unfolds in a single individual over time (Borsboom & Cramer, 2013; Fried & Cramer, 2017). How to infer between- and within-person effects, and what kind of data is needed to distinguish between them, is an issue that is important for psychological methodology in general (Brose, Voelkle, Lövdén, Lindenberger, & Schmiedek, 2015; Hamaker, 2012; Molenaar, 2004; Voelkle, Brose, Schmiedek, & Lindenberger, 2014), and has recently surfaced in the network literature.

Importantly, the data that a clinical researcher or clinician has available constrains the inferences that can be drawn. If there is only one time point available per person, it is hard to say much about within-person effects. On the other hand, if the data consist of a time series of just one person, it is impossible to infer between-person effects from the fit of a statistical model. In general, we may distinguish between three types of data used in estimating network structures: (1) cross-sectional data of many individuals measured once, (2) many repeated measures of only one individual ($N = 1$), known as time series data, and (3) repeated measures of many individuals (Borsboom et al., 2021). In this section, we discuss different ways of distinguishing between- and within-person effects based on such data, (Epskamp, 2020; Epskamp, Borsboom, & Fried, 2018; Hamaker, Asparouhov, Brose, Schmiedek, & Muthén, 2018), as well as the pressing question of how these levels of analysis should be conceptually related in network analyses.

Cross-sectional data. Most psychopathological networks are based on cross-sectional data (Robinaugh et al., 2020), frequently with the aim to study between-person (i.e., interindividual) differences, and sometimes with the aim of generating hypotheses on within-person (i.e., intra-individual) dynamics, for instance by using the Gaussian Graphical Model (Epskamp, Maris, Waldorp, & Borsboom, 2018). However, it is important to note that cross-sectional analyses do not statistically separate between-person variability (e.g., typically stable trait-like features) from within-person variability (e.g., more transient, state-like features). For example, the state of participants when they provide self-report ratings has been shown to affect cross-sectional results in particular cases (Brose, Lindenberger, & Schmiedek, 2013). Thus, analyses based on cross-sectional data might result in networks with edges that reflect a mix of between and within effects (Hamaker, 2012).

Furthermore, measurement instruments are usually tailored to capture either between-person or within-person variability. The difference is not just in how frequently the questions are asked (e.g., just once vs. several times per day), but also in the way they are formulated. For example, in cross-sectional studies participants are often asked to average over long periods of time (e.g., how sad have you felt in the past two weeks), whereas in intensive longitudinal studies they are asked about the current state (e.g., how sad are you feeling at this moment). These conceptual differences hamper statistical comparisons between cross-sectional and temporal networks. In a similar vein, Brose, Schmiedek, Gerstorf, and Voelkle (2020) have recently pointed out that the common practice of using measurement instruments that were established at the between-person level to assess affect at the within-person level is problematic, as these instruments are likely to have different psychometric properties at the within-person level.

Moreover, cross-sectional data usually cannot be straightforwardly used to identify within-person dynamics or the development of state-like features (Hamaker, 2012; Molenaar, 2004). As pointed out by Molenaar (2004), only under very specific circumstances known as *ergodicity* are

the results obtained from studying cross-sectional data directly representative of intraindividual processes. Ergodicity requires that individuals are independent and that the same data generating process applies to all individuals (homogeneity). Furthermore, the characteristics of the data (e.g., means, variances, auto- and cross-covariances) are not allowed to change over time, which means there cannot be any trends in the data; this is known as the assumption of stationarity (see Molenaar, 2004 for details, and also the section *(Non)-stationarity* below).

In psychological research, however, these assumptions are seldom realistic and for particular cases have been demonstrated to be contradicted by empirical data (Brose et al., 2015; Schmiedek, Lövdén, von Oertzen, & Lindenberger, 2020). As such, strict ergodicity can typically be ruled out as a reasonable possibility. Although this precludes a direct generalization from the cross-sectional results to the within-person level, it is important to note that lack of ergodicity does not imply that *nothing* can be learned from cross-sectional data about within-person structures. For example, by controlling for factors that introduce heterogeneity across individuals or across time, we may establish "conditional equivalence" that allows (conditional) inferences across levels (Adolf & Fried, 2019; Voelkle et al., 2014).[2]

Even with conditional equivalence, however, the challenge is to identify these factors, and great care is required to adequately control for them. Such control should be exerted experimentally (by random assignment), or when that is not feasible, by statistically controlling for the factors that induce heterogeneity, while avoiding issues such as conditioning on colliders (for an introduction to this topic, see Rohrer, 2018). In general, our theoretical understanding of how such factors affect ergodicity is limited, so the question of which aspects of inferences can be supported by various designs represents an important challenge for the network approach. Nevertheless, while network hypotheses often concern within-person dynamical processes, researchers will often have to rely on (partly) cross-sectional data due to the subject matter, as many relevant causal effects cannot realistically be studied in a time series design (e.g., the effect of child abuse on adult psychopathology; Isvoranu et al., 2016; see also the discussion below).

Time series data. When inferring network models based on time series data of a single subject (e.g., an episode of depression in one patient), all variance involves fluctuations in (emotional) states over time; the only stable, trait-like feature is the person's mean value for each variable. Accordingly, a time series $N = 1$ design enables us to model the dynamics of lagged effects within and between variables over time (e.g., the vector autoregressive model; Lütkepohl, 2005), and to model contemporaneous (i.e., lag 0) relations between variables to determine which phenomena tend to co-occur (e.g., using the graphical vector autoregressive model and GIMME; Gates & Molenaar, 2012; Beltz, Wright, Sprague, & Molenaar, 2016). One disadvantage of this approach when inferring networks from clinical data is that there may not always be sufficient fluctuation in the variables: For example, if levels of suicidality and hopelessness always remain high, it is not possible to say much about the effects these variables have on other symptoms and thus the central role and urgency of high suicidality may be overlooked as well. A further downside of single-subject time series data is that the findings are individual-specific and cannot be generalized to other individuals (unless we assume that all individuals have the same within-person dynamics). This may, however, not always be problematic for clinical practice, as the focus there is often on a single individual.

Multiple persons time series data. So far, we have discussed data of several individuals at one time point and data of a single individual over many time points. When these two are combined, we have time series data of multiple individuals, providing a wealth of information for clinical researchers. With such data, we can disentangle the within-person, state-like variance—akin to that in the $N = 1$ setting—from the between-person, trait-like variance characterized by the stable differences in means between individuals.

Disentangling the within- and between-person variance can be accomplished by using a multilevel version of the VAR model in which the predictors are within-person centered (Bolger & Laurenceau, 2013; Bringmann et al., 2013; Hamaker, Kuiper, & Grasman, 2015). Such multilevel VAR models have been very popular in estimating psychopathological network structures (e.g., Bringmann et al., 2016; Pe et al., 2015; Wigman et al., 2015). However, in these kinds of time series models, the observed mean scores are often used to center the predictor, which can lead to biases (McNeish & Hamaker, 2020; Nickell, 1981). Bias can be avoided by decomposing the observed scores into two components via latent person-mean centering, for example, by combining multilevel vector autoregressive modeling with structural equation modeling (Hamaker & Muthen, 2018; cf. Anderlucci & Viroli, 2015; Ernst, Albers, Jeronimus, & Timmerman, 2020). As these are time series-based models, they only work well with intensive longitudinal data. However, networks that disentangle the within and between structure can also be estimated when only few repeated measurements are available (i.e., panel data), for example through a cross-lagged panel model with random intercepts (Hamaker et al., 2015; Epskamp, 2020; Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020).

Once the within-person part and a between-person part have been separated, these different components can be further investigated when studying psychological networks. On the one hand, the within-person part offers the opportunity to study how the momentary fluctuations in one aspect (e.g., stress) are related to momentary fluctuations in another aspect (e.g., worry). Moreover, it allows us to use lagged regressions (e.g., using vector autoregressive based models) to see whether, for instance, increases in stress *precede* or *follow* increases in worrying; such temporal relations strongly adhere to the notion of a process that unfolds over time, with symptoms, experiences, thoughts, and behaviors triggering or attenuating each other (Borsboom & Cramer, 2013). On the other hand, the between-person part reflects the stable differences between individuals, with some individuals scoring consistently higher (or lower) than others on the measured phenomena. If one seeks clues for long term clinical intervention, then it may be better to focus on stable between-person mean differences on measures of symptoms, behavior, and so forth, rather than short-term, transient fluctuations modeled at the within-person level.

Epskamp, Waldorp, Mõttus, & Borsboom (2018) argue that both parts may be of interest in network analysis, and that what is needed is a more thorough investigation of what causal processes are reflected by each level, and how one may be able to intervene on these. One important factor is the *timescale* at which a process unfolds. Some interventions will be associated with the mechanism that is operating at a timescale captured by the within-person fluctuations (e.g., the effect of a minor setback on an individual's mood during a period of two weeks). However, other interventions will be only visible at the timescale associated with the between-person variability (e.g., the effect of major life stressors on affect), and perhaps should be seen as part of the external field of the network (see the subsection on *Node selection* above).

Importantly, often variables associated with between-person variability cannot reasonably be measured repeatedly. For instance, significant childhood experiences, such as the divorce of one's parents, may have profound effects on the development of children and on the mental health of an individual. Although these experiences themselves were

[2] Furthermore, under the less restrictive assumption of measurement invariance, some aspects of within-person dynamics do produce identifiable patterns in the cross-sectional data (Adolf, Schuurman, Borkenau, Borsboom, & Dolan, 2014; also see; Epskamp, 2020). Importantly, however, before attempting any generalization across levels, the degree of equivalence needs to be established and critically evaluated, which requires longitudinal data. Moreover, discussions of (conditional) ergodicity have mainly focused on contemporaneous (lag 0) relationships instead of dynamic or temporal (e.g., lag 1) relationships, which temporal networks consist of.

processes that evolved over time, often we can only study the effect of such experiences by comparing individuals *with* and *without* them, while trying to account for possible confounders. These stable between-person differences in such experiences can then be used to study other individual differences later in life (e.g., the ability to maintain an intimate relationship). Translated to the network context, although some variables are measured only once and give information about between-person variability, they can nevertheless also help to explain differences in temporal networks. As a hypothetical example, childhood trauma as a between-person variable could be used to explain why some individuals (who experienced childhood trauma) have, for instance, stronger lagged and contemporaneous relationships between variables in their individual network than others (who did not experience childhood trauma). In other words, individuals can also differ not just with respect to mean levels of, for example, symptoms, but also with regard to the short-term, transient fluctuations modeled at the within-person level. In order to study these kinds of differences, temporal network models and thus time series data (in addition to cross-sectional data) are needed.

Thus, complex (within-person) network dynamics cannot be directly determined from cross-sectional data, nor from between-person differences concerning experiences in the past. Instead, it is important to carefully think about the timescale and timing at which a mechanism occurs, and to align the *timing*, *frequency*, and *duration* of measurements with the *time* and *timescale* at which the process of interest unfolds. A fruitful starting point would be clinical practice, for example, case conceptualization in cognitive behavioral therapy. In case conceptualization, therapists (together with the patient) discuss how the mental health problems manifested or evolved over time and relate to one's beliefs, behavior and life events, resulting in clinical theory for the specific patient (Beck, 1979). Such theories can give important clues on the time scale at which a certain clinical process unfolds (Burger et al., 2020; von Klipstein et al., 2020).

3. Different time-scales

Closely related to the previous section, an important challenge for the analysis and interpretation of psychological networks is how to deal with different time-scales. As we have seen in the previous section, clinically relevant variables or nodes are measured at various times scales. For example, sleep is usually measured daily, whereas affect is often measured several times a day (see, e.g., Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017). Typically, only nodes measured at the same time scale appear in a network, whereas other highly important nodes are not included because they are measured on a different time scale (Bastiaansen et al., 2020).

In order to tackle the problem concerning different time-scales it is useful to distinguish between two different, but related, problems: First, how to analyze, interpret, and communicate results from networks that contain nodes that are *assessed* at different time scales and second, how to analyze, interpret, and communicate results from networks that contain nodes that *operate* at different time scales.

Regarding the first problem of measurements at different time points, consider a patient's bodyweight and physical well-being. Bodyweight could be assessed every two weeks, whereas physical well-being could be assessed once a week. In addition, the time intervals between measurements of each construct could differ from one measurement occasion to the next and could be different for each patient. For example, a patient could have skipped several assessments (e.g., due to holidays, physical well-being may not be reported two weeks in a row) and individuals could differ when they skip assessments (not all individuals take their holidays at the same time). This can easily result in a highly unbalanced data structure, where each network node and each patient is observed at a unique point in time that is not shared by any other node or patient. How can we still analyze and interpret the development and connectivity among nodes in such a situation?

Regarding the second problem of how to analyze, interpret, and communicate results from networks that contain nodes that *operate* at different time scales, it is important to realize that even if variables such as bodyweight and positive affect are assessed at the same time points (e.g., once a day), they are likely to operate on different time scales. In a typical population of healthy adults, meaningful changes in bodyweight from one day to the next seem less likely than meaningful changes in positive affect from one day to another. Yet, there may be good reasons to expect a relationship between the two variables.

In existing psychological network approaches to psychopathology, the two problems are currently unresolved. However, in other fields, such as time series analysis, dynamical systems and control theory, researchers have developed ways to deal with these problems that may also prove useful for psychological networks (Kirk, 2004; Lütkepohl, 2005; Mansell & Marken, 2015; van Montfort et al., 2018). In our opinion, key to the resolution of both problems is to separate the measurement process, which is always discrete, from the underlying process of substantive interest, which often develops continuously over time. An "unemployment index does not cease to exist between readings, nor does Yule's pendulum cease to swing" (Bartlett, 1946, p. 31). Perhaps a psychological network also does not cease to operate when it is not observed. Many variables, such as bodyweight or positive affect operate continuously and do not operate at *one* true time scale. However, other variables, such as drug use or panic attacks, may involve discrete events that may truly be absent, and not just unobserved, between measurement occasions.

In cases where discrete measurements are imposed on an underlying continuous process, instead of trying to "hit" the right time scale and then proceed with analyzing the data at that time scale, we can alternatively start with a model of how the process develops in real, continuous time and then seek evidence for or against this model based on discrete measurement occasions. One way to achieve this is by means of continuous time dynamic models that have recently gained increasing attention in psychological research (e.g., Haslbeck, Ryan, Robinaugh, Waldorp, & Borsboom, 2021; Oud & Jansen, 2000; van Montfort, Oud, & Voelkle, 2018; Ryan, Kuiper, & Hamaker, 2018; Ryan & Hamaker, 2021; Voelkle, Oud, Davidov, & Schmidt, 2012; Voelkle et al., 2012). By means of stochastic differential equations (SDEs), continuous time models formulate how change takes place over an infinitesimally small interval (i.e., in continuous time). By solving the SDE for a given starting point and a given discrete measurement interval, the model parameters can be estimated irrespective of the time scale the assessment took place. Importantly, the time intervals may differ across constructs (and individuals), so that variables with different sampling schemes, such as sleep and affect, may be included in the same model without the need to preprocess the data or deal with "missing" observations as it is often the case in discrete time modeling. By deriving the dynamics for many different intervals, we can study how nodes operate at different time scales. Moving beyond simple first order continuous time dynamic models, higher order models offer the necessary flexibility to study rather complex behavior (Oud, Voelkle, & Driver, 2018). While continuous time models based on linear SDEs can usually be solved analytically, nonlinear dynamic models may offer yet additional flexibility, for example, in terms of studying different equilibria a network may converge to over time. However, naturally, such models come with additional challenges.

Once the model is estimated, one can study, for instance, node centrality as a function of the time scale of substantive interest (Ryan & Hamaker, 2021). For example, one could determine the time interval at which an effect is large (maximal) given a continuous time model, rather than hoping that the discrete time measurement interval was optimally chosen to capture a (large) effect. Especially for network models with many nodes, the direction and size of effects between nodes will often depend on a complex nonlinear fashion on the time-scale under consideration (e.g., Voelkle, Gische, Driver, & Lindenberger, 2018).

Importantly, however, the increased flexibility of measuring and

studying phenomena at different time scales does not free the researcher from theoretical considerations regarding the "optimal" time scale. Indeed, study design will remain crucially important, because to obtain information about variables operating at different time scales, we still need at least some measurements at appropriate time scales (Bolger, Davis, & Rafaeli, 2003; Bolger & Laurenceau, 2013; Mehl & Conner, 2012; Shrout et al., 2018). Ideally, assessments should occur at a sampling rate appropriate to the process of interest (do symptoms and feelings affect one another within seconds, minutes, hours, or days; Hamaker et al., 2015; Trull, Lane, Koval, & Ebner-Priemer, 2015). More specifically, if the aim is to infer short term behavior of a network based on large intervals, such interpolation may be accompanied by high uncertainty (see, e.g., Adolf, Loossens, Tuerlinckx, & Ceulemans, 2021). For example, if affect is measured in all studies only once per day, not even the most advanced continuous time model can capture how it fluctuates at a different time scale, such as from hour to hour. However, most intensive longitudinal or time series studies use three to ten assessments per day, and the choice of measurement frequency is usually not merely influenced by the expected time scale of the processes of interest, but often even more so by considerations of feasibility and compliance (Vachon, Rintala, Viechtbauer, & Myin-Germeys, 2018). In addition, some variables, such as panic attacks, psychotic experiences, or binge eating episodes, may be essentially discrete events that are not simple functions of an underlying continuous process.

To date, most clinical theories are not very precise in terms of formulating the exact "optimal" time scale at which an effect will take place. Often effects are simply divided into "short term" and "long term", but it remains unclear what this refers to in quantifiable units such as hours, days, or years. While more careful theoretical considerations regarding the time scale are essential, by adopting a continuous time framework, part of the burden of finding the optimal time scale can be left to the model (cf. Dormann & Griffin, 2015; Voelkle & Oud, 2012). In the end, it is up to the researcher or clinician to get the time scales *approximately* right and to obtain data, preferably at a high frequency (Helmich et al., 2021), that allow her to identify the presumed effects.

Finally, although continuous time models are potentially a good solution for the issues of aligning different time scales, we still must confront the challenge of estimating continuous time network models from intensive longitudinal data. For example, the relatively large number of nodes in networks means that the number of time points needed to reliably infer a continuous time model will often be much higher than is feasible in practice (Ryan, 2020). This is thus an important topic for future research, where Bayesian approaches or regularization techniques may prove beneficial (Driver & Voelkle, 2018; Epskamp & Fried, 2018; Oravecz, Tuerlinckx, & Vandekerckhove, 2016).

4. (Non)-stationarity

An issue receiving increasing attention in psychological networks is non-stationarity. Most extant network models are stationary, meaning that certain characteristics – e.g., the global structure of the network – are assumed to be stable over time. However, change is integral to psychopathological networks. One hypothesis, for instance, that has been much discussed in the network literature is that the connections (edges) between symptoms are less strong in a healthy state than during an episode of depression (Borsboom, 2017; Wichers & Groot, 2016). In other words, a night of bad sleep may not trigger any feelings of sadness and anxiety when one is doing well, but when one is approaching a depressive episode, it may result in a vicious circle of symptoms that keep on activating each other. Moreover, changes in symptom dynamics may provide better early warning signals of a depressive episode than changes (i.e., increases) in the mean levels of symptoms (see also section *Networks and interventions in clinical practice*).

To study changes in network structure, we need multivariate statistical approaches that make it possible to study symptom dynamics that themselves are subject to change (i.e., non-stationary). However, when handling such non-stationarity, the most progress so far has been made in univariate, not multivariate, models. Therefore, we discuss both univariate and multivariate approaches to modeling changes in time series data. Even though univariate models are not strictly speaking network models, they may still be used to test certain network hypotheses related to change.

Changes in the network can take many forms: For example, the change can be gradual or abrupt. One approach to screen multivariate time series for *abrupt* long-lasting changes is KCP-RS (Kernel Change Point detection on the Running Statistics), a non-parametric multivariate change point detection toolbox and associated R-package (Cabrieto et al., 2018a, 2018b, 2018c, Cabrieto, Adolf, Tuerlinckx, Kuppens, & Ceulemans, 2019). The basic idea of KCP-RS is simple: Users select the statistics of interest (e.g., means, variances, correlations, autocorrelations, or others). For the network approach, changes in cross-lagged effects would be particularly important, that is, how one symptom influences another symptom at the next time point (e.g., how the effect of poor sleep on sadness changes from a negligible to a strong effect).

Next, the user extracts running versions of these statistics by sliding a time window across the time series and computing the statistic of choice in each time window. Finally, Kernel Change Point detection (Arlot, Celisse, & Harchaoui, 2012) is applied to the obtained running statistics. The KCP-RS toolbox includes significance tests to decide whether one or more change points are present in each of the running statistics and a tuning procedure to assess the exact number and location of change points. Validation studies on simulated and empirical data showcase good performance in case of longer time series and moderately large effect sizes. However, regarding cross-lagged effects, no validation studies have been conducted yet. Hence, more research is needed to determine how many time points are required and how large the moving window must be to reliably detect change points in the cross-lagged effects of the network.

Closely related to the change point detection approach is the threshold autoregressive (TAR) model (Tong & Lim., 1980) or the fixed moderated time series model (Adolf, Voelkle, Brose, & Schmiedek, 2017). In these models, the breaks or changes in the dynamics of one (e.g., changes in the autoregressive effect) or more (e.g., changes in the cross-lagged effects) nodes can be identified by using another observed variable (i.e., a predictor), such as another node in the network. In doing so, these models can model non-linear effects of a predictor: Through estimating separate slopes for different values of this predictor (e.g., below or above a specific threshold value), the effect of a one-unit increase becomes dependent on the concurrent predictor score. Such models have been used to investigate whether inertia (i.e., autoregressive effect) of an individual differs depending on whether one is scoring low or high on momentary positive (or negative) affect (De Haan-Rietdijk, Gottman, Bergeman, & Hamaker, 2016) and to study the dynamics of affect and stress and how these co-vary with daily events (Adolf et al., 2017). This approach implies that individuals can switch back and forth between two or more different regimes that are characterized by different parameters (e.g., intercepts, autoregressive coefficient, cross-lagged coefficients, and innovation variance), and that can be identified through the predictor (see, e.g., Hamaker, 2009).[3] Thus, if such a model detects two regimes, two regime-specific networks can be estimated based on a predictor that can be within or outside of the network, with the individual switching back and forth between these regimes (i.e., networks).

The above models assume rather sudden and abrupt changes which are associated with a break or switch between regimes. Alternatively, time-varying (V)AR models that use splines or kernels allow for

[3] Another option is to use regime-switching models in which the switching is governed by a hidden Markov model (Kim & Nelson, 1999), rather than an observed variable. Then, in addition to estimating the regime specific parameters, one also estimates the switching probabilities between any two regimes.

processes that are characterized by *gradual* changes over time, in the intercept, autoregression, cross-lagged effects and the mean (Bringmann et al., 2017, Bringmann, Ferrer, Hamaker, Borsboom, & Tuerlinckx, 2018). Such gradual changes thus imply that the change in the underlying network structure is also gradual (Bringmann et al. 2017, 2018; Haslbeck & Waldorp, 2020; Haslbeck, Waldorp & Bringmann, 2021). These time-varying models are discrete time models, but extensions to continuous time models (see also the models discussed in the section *Different time-scales*) have been proposed (Chen, Chow, & Hunter, 2018). Modeling heterogeneity or change over time in the innovation (i.e., residuals) variance is impossible with most of these models (but see Adolf et al., 2017).

Other approaches can detect and model *both* gradual and abrupt changes, but most are univariate and therefore cannot be directly used to study changes in symptom dynamics in a network. A recent such approach is presented by Albers and Bringmann (2020), who introduce a time-varying change point autoregressive (TVCP-AR) model. This approach uses the time-varying AR model as a starting point (for modeling gradual changes) and combines it with a change point (CP) modelling approach (for modeling abrupt changes; Bringmann et al., 2017; Hamilton, 1994). The change point part of the model is especially well suited for data when one or a few abrupt changes happen, but is not suitable if the dynamics frequently shift between different regimes for short times (in the latter case, the models discussed above are more appropriate).

Although these options for modeling non-stationarity in network dynamics are both promising and complementary, several challenges remain, yielding a research agenda for the future. First, to retrieve changes, simulation studies have shown that abundant data from the same person are needed (e.g., 200 time points, but often more; Cabrieto et al., 2018b, 2018c Bringmann et al., 2018), especially when changes are more subtle. Therefore, future research may inspect how much can be gained in signal to noise ratio by selecting or creating the most informative nodes possible (see subsections on *Node selection* and *assessment*), or by aggregating information from different individuals that go through the same changes, such as studying the changes in the networks of participants following the same type of therapy.

Second, whereas changes in means are relatively easy to detect, first results on changes in other statistics or model parameters, such as the cross-lagged effects that are of main interest for the network approach, indicate that they are often much more subtle and therefore more difficult to pinpoint (Cabrieto et al., 2018b). Especially in methods for detecting changes in multiple features simultaneously, results will frequently be dominated by mean changes that obscure other changes. These findings call on the one hand for clear theoretical hypotheses on which statistics change when for instance a depressive episode develops (Wichers, Schreuder, Goekoop, & Groen, 2019); such hypotheses can then be used to build and apply parsimonious and targeted modeling strategies. For instance, in a regime switching model one can allow all parameters (i.e., intercepts, auto and cross-lagged effects, innovation covariances) to vary across regimes, or one can restrict some to be the same across regimes. In addition to theoretical hypotheses, more work is needed on model selection criteria, and tests and procedures to assist researchers in applying and evaluating these model building strategies.

5. Networks and interventions in clinical research and practice

Ultimately, the success of the network approach will depend on whether it can add value to evidence-based predictions and enhance effects of interventions for mental disorders (McNally, 2016). In other words, will the complex computational methods of network analysis enhance the ability of clinicians to reduce emotional suffering? Network approaches have already been applied to guide personalized, optimal treatment strategies (Bak et al., 2016; Fisher et al., 2017; Kroeze et al., 2016; Lutz et al., 2018; Rubel, Fisher, Husen, & Lutz, 2018). These early applications show promise as a means of translating network analysis into clinical practice. However, until now no study has demonstrated that these network-based treatment strategies are beneficial above and beyond existing treatments for mental disorders, and many important questions remain.

Let us first turn to the question whether the network approach can lead to better predictions of who will develop (or relapse into) a mental disorder. So far, one of the most promising approaches has been connecting network models to dynamic systems theory (Cramer et al., 2016; Wichers, 2014) whereby early warning signals enable predictability (Cramer et al., 2016; van de Leemput et al., 2014). The idea is that mental disorders such as depression behave like complex dynamic systems, where fragility of the system develops gradually until a tipping point is reached, and the system abruptly shifts from one state (e.g., healthy state) to another (e.g., depressed state). Early warning signals may precede such tipping points. For example, it has been proposed that an increase in autocorrelation or an increase in network density (i.e., increase in the strength of associations among symptoms) may forecast such transitions between states (van de Leemput et al., 2014; Wichers et al., 2011). Preliminary results from two small pilot studies with one (Wichers et al., 2011) and six patients, respectively (Wichers, Smit, & Snippe, 2020), suggested that increase in network density preceded the onset of depression, but larger studies are needed to see if this is a robust phenomenon.

The predictability of networks has also been studied in a different way by Lutz et al. (2018), who explored the predictive power of centrality measures derived from individual dynamic networks. In this study, initial evidence suggested that centrality measures can have added value in predicting patient dropout. In general, although network-based predictability shows early promise for clinical practice, research on it has been limited, and more large-scale intensive longitudinal studies are needed (e.g., Helmich, Olthof, et al., 2021; Schreuder, Groen, Wigman, Hartman, & Wichers, 2020). More specifically, the predictive value of network connections and centrality on the course of mental health conditions needs to be studied systematically.

Another issue with current application of network structures in clinical practice is how to determine which specific edges or node centralities provide optimal target points for *interventions*. As mentioned in the introduction, it is an appealing idea that identifying the most central nodes (symptoms) in the network can also help to identify the most important intervention targets. Although the general theoretical idea of centrality seems clear, there are many different centrality measures, and it has been hotly debated whether centrality measures (and if so which ones) are suitable for psychological networks, and if they indeed indicate clinically important intervention targets (Bringmann et al., 2019; Dablander & Hinne, 2019; Hallquist et al., 2021). Whereas Bringmann et al. (2019) argue that the conceptualizations and operationalizations of centrality measures are not transparent in psychological networks, and it is not clear what they measure, Dablander & Hinne (2019) argue that central nodes are in general not necessarily causally important nodes in the network. That is, it is unclear whether deactivating a central node will result in the deactivation of its neighbors. Thus, these theoretical ideas, for example that nodes with many connections to other nodes might serve as efficient treatment targets, still need more mathematical and theoretical scrutiny (e.g., based on control theory; Henry, Robinaugh, & Fried, 2021) as well as empirical studies. Such empirical studies could include, for example, experimental designs using micro randomized controlled trials (including intense repeated daily assessments over time) and case series designs (Brouwer et al., 2020; Holmes et al., 2018.; Slofstra et al., 2018).

A related issue is that even if we assume that we have successfully identified the most clinically important nodes (symptoms) in the network, it is not always clear whether and if so how we could intervene on those specific nodes. Importantly, it is often difficult to see how one could intervene on one psychological symptom in a network without at the same time influencing other symptoms. For example, the prescription benzodiazepine for insomnia not only reduces sleeping problems,

but is also likely to reduce anxiety and worrying as well. Similarly, interventions based on cognitive behavioral therapy typically target various aspects of sleep at the same time, for instance by combining sleep restriction intervention with attempts to change cognitive structures (Blanken et al., 2019; van der Zweerde, Van Straten, Effting, Kyle, & Lancee, 2019). In other words, is it possible to "surgically" target a single symptom in isolation, or are psychological interventions always "fat-handed", changing several symptoms at the same time (Eronen, 2020)?

It can also be challenging in practice to distinguish between an intervention that effectively targets a single symptom, whose reduction in severity then diminishes the severity of other symptoms shortly thereafter (which is what network-inspired clinical practice would aim to achieve) from an intervention that simultaneously affects several symptoms (i.e., a fat-handed intervention). Moreover, the absence of improvement after an intervention does not necessarily mean that the network model was incorrect or that a network hypothesis (e.g., that a certain symptom is central in the network) is falsified, as it may be the case that the intervention was insufficiently efficacious to change the network. Thus, more research and empirical studies are needed that assess how networks change in response to an intervention and how this change relates to clinical improvement. Besides interventions that target specific nodes, this can also include other kinds of interventions, such as studying how intervening on the environment or external factors changes the network. For example, Bringmann et al. (2013) compared the affect networks of individuals with residual depressive symptoms before and after therapy (see also Snippe et al., 2017).

Finally, it is also important to study whether personalization of treatment guided by specific network characteristics also leads to better outcomes than current evidence-based treatments and/or traditional personalized assessments by clinicians. We may expect, if we can assume that the network structure is informative, that people will improve more strongly if the intervention is based on the person's network information, compared to a randomly generated network. However, even if the network structure is valid and informative, it may not lead to enough clinical gain to be useful in practice. This could be tested in randomized controlled trials comparing the effect of network-based treatment and traditional clinician-based treatment.

6. Reproducibility and replicability of networks

Another issue that is crucial for the clinical success of the network approach and that has received much attention recently is replicability (obtaining the same conclusions from new data) and reproducibility (obtaining the same conclusions from the same data) of network models. In this section, we will discuss the replicability and reproducibility of both cross-sectional and dynamic network models based on time series. Recently, many modelling techniques have been developed to check for robustness of results for cross-sectional data (e.g., Epskamp, Borsboom, & Fried, 2018). In contrast, for time series or intensive longitudinal data, this is largely uncharted territory, except for cross-validation approaches to check for overfitting (Bulteel, Mestdagh, Tuerlinckx, & Ceulemans, 2018).

Replicability of cross-sectional networks Recently several papers have been published that investigate expected replicability of cross-sectional networks (Borsboom, Robinaugh, Psychosystems Group, Rhemtulla, & Cramer, 2018; Fried et al., 2018), making use of simulation methods. For example, the *bootnet R* package contains a function *replicationSimulator* that can be used to generate two datasets from the same model and assess how well an estimated network for the first dataset will replicate in the second dataset. Furthermore, the *bootnet* package can be used to assess the stability and accuracy of results (Epskamp, Borsboom, & Fried, 2018). Additionally, van Borkulo et al. (2017) present a method to test if differences in observed network structures are larger than what would be expected if there actually is no difference at the population level.

In order to reduce the false positive rate and increase replicability, many cross-sectional network estimation methods include built-in regularization techniques. Regularization techniques, such as the well-known least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996), set very small coefficients exactly to zero, thereby reducing the likelihood of spurious edges (Epskamp, Waldorp, et al., 2018). In this way, many network estimation methods aim to be conservative. Therefore, it is to be expected that at low sample sizes the networks estimated contain only a subset of the true edges. This means that, for example, if we estimate a network based on two subsets of the same data and the sample size is low, it is likely that not all edges from the first network replicate in the second network. Additionally, if the two data sets differ in sample size, it is to be expected that the network based on the dataset with a higher sample-size will contain more edges. A final important consideration is that samples can also feature genuine heterogeneity in their underlying models (e.g., due to differences in populations), and as such it may not be expected that a single network model is retrieved from two datasets even when these have high sample sizes (Isvoranu, Epskamp, & Cheung, 2021). The replicability of cross-sectional networks remains an important topic in current discussions (see, e.g., Forbes, Wright, Markon, & Krueger, 2019 and Jones, Williams, & McNally, 2021), and novel approaches have been proposed, such as Bayesian model selection methods to demonstrate replicability (Williams, Rast, Pericchi, & Mulder, 2020), as well as meta-analytic methods for aggregating multiple studies (Epskamp, Isvoranu, & Cheng, 2021).

Replicability of dynamic networks For dynamic networks based on time series models, replicability remains a largely understudied topic. However, one way in which it has been studied is from the angle of cross-validation and predictive accuracy, assessing how well model estimates for training parts of the data allow to predict test parts of the data (Bulteel, Mestdagh, et al., 2018; Bulteel, Tuerlinckx, Brose, & Ceulemans, 2018). First, results suggest that predictive accuracy is low when a complex network model (e.g., a multilevel VAR model) with many edges is fitted on a relatively small data set (i.e., low number of measurement occasions), thus resulting in overfitting. Although it is likely that the amount of overfitting will decrease with larger (multilevel) data sets and with fewer nodes in the network (Lafit, Meers, & Ceulemans, 2021), it is not possible to formulate one-size-fits-all rules of thumb regarding how large the data set should be. Recommendations will depend on the number of variables, effect size, type of model, etc. (Lafit et al., 2021; see also; Liu, 2017). Therefore, careful sample size planning (e.g., power analyses; Lafit et al., 2020) is important for increasing the likelihood of replicable results.

Software availability In addition to the above threats to replicability, the network approach also faces threats to reproducibility, most importantly, threats due to software availability and due to methodological developments (Epskamp, 2019). Recent years have seen a surge of applied publications in which state-of-the-art network methods are applied. Usually, these new methods have been introduced in methodological journals and supported by evidence in the form of mathematical proofs as well as extensive simulations studies. However, it may not be easy or feasible for many applied researchers to implement such a method correctly for several reasons. When code is made available, this may come in various levels of quality, ranging from poorly documented non-flexible code to transparent and well-documented R packages on the comprehensive R archive network (R Core Team, 2020). Even when code is implemented in a documented software package, it may still contain bugs, be still in development, or may change over time. The problem may further be exacerbated by having such potentially changing software packages depending on other packages that may change (dependency trees; e.g., Figure 2 of Epskamp, 2019). Therefore, there is the risk that applied studies do not implement methods correctly or that the results of network studies do not reproduce due to software issues.

Methodological development In addition to software developing and

changing over time, methodological recommendations may also develop and change over time. It is typical for methodological papers to be thoroughly reviewed, leading to publication processes that may well take over a year. This means that when applied researchers use a method based on a preprint, recommendations for using that method may still considerably change from the preprint to the final published version, and the results of the applied article may end up being published before the methodological article. In addition, a fast-paced methodological world alone already dictates that the current state-of-the-art may not be next year's state-of-the-art, and new choices in the analyses may lead to different conclusions. As such, while a research method (e.g., an estimated network structure) may be reproducible, its generated conclusions may not be replicable per se.

This is not a problem exclusive to psychological network analysis, as it may also apply to numerous novel methodologies (e.g., advances in machine learning). It may never be clear when a methodology is sufficiently crystalized such that further changes are unlikely. Hence, there is no straightforward solution to this problem. However, it is important that researchers make as much of the analysis code and data available openly, in order to allow future researchers to study the same question by using updated methods. Moreover, it is also helpful if methodological researchers provide their simulation codes and data online, to make it easier to conduct benchmarking studies that clarify which methods work best for which type of data (Doove, Wilderjans, Calcagnì, & Van Mechelen, 2017). Additionally, openness should not be restricted to sharing code and data, but should also include transparency. In other words, it is important that code and data are well annotated and documented and thus understandable for other researchers.

In general, considerations of replicability and reproducibility of psychological networks are crucially important, not just for network research, but also when applying network methods and results to clinical practice.

7. Testability and falsifiability

The network approach connects broad conceptual frameworks and statistical models with empirical psycho(patho)logical phenomena, and has led to novel hypotheses, such as that the episodes of depression are trigged by causal interactions between symptoms (Cramer et al., 2016). However, it is still far from clear how exactly these hypotheses, or the theories and models underlying them, could be falsified or tested. For the network approach to develop further, it is therefore important to discuss how network hypotheses, theories and models can be empirically tested.

It is first of all important to note that unlike social networks or railroad networks, psychological networks are not directly observable from the raw data, but are rather based on statistical models (e.g., the Ising model, the Rasch model, or the VAR model). These statistical models as such are not normally considered to be subject to falsification, because they are merely abstract families of probability distributions (Bringmann & Eronen, 2018). One can only start testing theories and thus falsifying them once these models are given an empirical *interpretation*, that is, when the abstract symbols "X" and "Y" in the formulation of a statistical model are replaced by concrete components such as symptoms (e.g., "insomnia" or "concentration problems") and the abstract statistical regression of Y on X is replaced by a substantively meaningful dynamic or causal relation between these symptoms (e.g., if insomnia persists for days or weeks, concentration problems start to arise). This move from abstract schemes or statistical models to falsifiable interpreted theories requires three important theoretical steps: (a) identifying the *components* of the network (i.e., what are the Xs and Ys?), (b) identifying the *dynamic relations* between them (i.e., how do these real-world components – represented by Xs and Ys – influence one another over time?), and (c) deriving empirical implications from the so-constructed dynamical model.

In the section *Node validity* above, we have discussed topic (a) identifying the components or variables of the network, which has already gained attention in the literature. However, in most currently used statistical network models, (b) the *dynamical relations* are not explicitly represented. In population or cross-sectional models, dynamical relations are absent (Epskamp, Waldorp, et al., 2018), whereas in regression models fitted on highly intensive time series (e.g., VAR-based models; Bringmann et al., 2013) the model is typically chosen for convenience rather than derived from a theory that says how network components influence each other over time. However, if one wants to move from network models to network theories, one must specify how changes (e.g., the component "insomnia" becomes active) may propagate through the whole network (e.g., how "insomnia" may trigger the development of other symptoms and, as such, trigger a depressive episode).

There are only a few examples of theoretical network models that have made this step, even provisionally: Van der Maas et al.'s (2006) mutualism model of intelligence, Cramer et al.'s (2016) proposal for a minimal working theoretical model for major depression, Dalege, Borsboom, van Harreveld, and van der Maas (2018) theory of attitude change, and Robinaugh et al.'s (2019) computational model for Panic Disorder. For example, in Robinaugh et al.'s (2019) formalization of the standing theory on Panic Disorder, the component *arousal schema* (a set of beliefs about the implications of bodily sensations like increased heart rate) controls the strength of the feedback loop between the components *arousal* and *perceived threat* that generates panic attacks (Beck, 1985). Because this theory has (a) identified the relevant components, and (b) specified the relations between these components, testable consequences can be derived: for instance, if we observed no association between arousal schema and panic attacks, the theory would have to be rejected or adapted.

Thus, after both (a) network components and (b) dynamical relations between them have been specified, the resulting computational model can be used to (c) *derive testable hypotheses* concerning the behavior of the system as a whole. For example, Cramer et al. (2016) discovered that their simulated major depression network implies the existence of hysteresis, that is, the phenomenon that the amount of stress reduction needed to get out of depression is much higher than the amount of stress needed to trigger an episode of depression. Dalege et al.'s (2018) attitude model implies the existence of an inverse mere thought effect, and Robinaugh et al.'s (2019) model implies several distinct pathways to generate panic attacks. These implications can then be tested in a relevant experimental or quasi-experimental design; importantly, this may, but need not, involve fitting network models statistically. For example, if Cramer et al.'s (2016) depression model is correct, reducing insomnia should reduce other symptoms; and although this prediction can be tested with a network model (Blanken et al., 2019), it can also be tested in other ways (Ballesio et al., 2018; van der Zweerde et al., 2019). However, if statistical network models are indeed used in testing substantive network theories, current exploratory algorithms may be complemented by confirmatory network analysis techniques (Epskamp et al., 2017; Epskamp, 2020).

This less data-driven and more theory-based approach can also help to make the network approach more applicable to clinical practice. Whereas currently the focus in applying the network approach to clinical practice has been very much on data-driven and exploratory techniques such as VAR-based models, approaches that are more focused on confirmation and falsification are better suited to identify flaws in network theories or models. This can be an important step in developing models that are helpful for clinical practice, and should receive more attention in the clinical field.

8. Conclusions

In the past decade, the network approach has sparked much debate and has had a significant impact on the way how mental disorders are perceived in the field of clinical psychology, namely as a complex system

where symptoms and mechanisms are interacting with each other, giving symptoms a central role in understanding how mental disorders develop and are sustained. However, there are still important challenges in moving from the theory to empirical research and clinical practice and vice versa. In this article, we have discussed these methodological and practical challenges that the network approach still faces.

Importantly, many of the challenges come back to the question of timescales. First, timescales are crucial for thinking about how different levels, such as the psychological and physiological (or even affect states vs. symptoms), interact with each other. Second, to know which nodes should be included in a network and how they relate to each other, timescale should be considered, for example, to distinguish between between-person and within-person effects. The question of timescales also has important methodological consequences, not just for data collection or for node selection, but also for modeling, whether to use a continuous time model or a discrete time model. It is also important to study which timescales (e.g., hours, days, weeks or months) are the most relevant ones for capturing the most crucial fluctuations for understanding psychopathology, as increase in the number of time points collected can also lead to increase in the burden for the patients.

Another important challenge for the network approach is that change is crucial for psychopathology, for example for understanding how individuals develop or recover from a mental health condition. This involves change in the outcome variables of interest, as well as change in the model parameters that captures such changes. However, most standard models to infer networks, such as Gaussian graphical models or VAR, assume stationarity, meaning that the dynamics of the network do not change over time. Yet change is integral to the psychotherapeutic enterprise. In this article we have discussed the various emerging options for dealing with nonstationarity, such as time-varying VAR models and change point analysis.

As in all other fields of science, reproducibility and replicability are crucial for the researchers and clinicians using network techniques. We have laid out the latest developments in this regard, emphasizing the importance of making the code and methods clear and transparent, and discussed new ways of assessing the replicability of network models. We have also argued that in order to test network hypotheses and to identify flaws in network theories and models, data-driven and exploratory techniques such as VAR based models are not sufficient, but also confirmatory and theory-driven approaches are needed.

In the end, the success of the network approach will depend on the extent to which it can add value to the prediction of onset and maintenance of mental health conditions and even more so the added value of improving the effect of treatment above current evidence-based treatments. We hope that in this paper we have not only laid out the theoretical and methodological challenges to making the network approach clinically applicable, but also have shown pathways and avenues to move forward to address these challenges.

CRediT authorship contribution statement

Laura F. Bringmann: Conceptualization, Writing – original draft, Writing – review & editing, Project administration. **Casper Albers:** Conceptualization, Writing – original draft, Writing – review & editing. **Claudi Bockting:** Conceptualization, Writing – original draft, Writing – review & editing. **Denny Borsboom:** Conceptualization, Writing – original draft, Writing – review & editing. **Eva Ceulemans:** Conceptualization, Writing – original draft, Writing – review & editing. **Angélique Cramer:** Writing – original draft, Writing – review & editing. **Sacha Epskamp:** Conceptualization, Writing – original draft, Writing – review & editing. **Markus I. Eronen:** Writing – original draft, Writing – review & editing. **Ellen Hamaker:** Conceptualization, Writing – original draft, Writing – review & editing. **Peter Kuppens:** Conceptualization, Writing – original draft, Writing – review & editing. **Wolfgang Lutz:** Conceptualization, Writing – original draft, Writing – review & editing. **Richard J. McNally:** Conceptualization, Writing – original draft, Writing – review & editing. **Peter Molenaar:** Writing – original draft, Writing – review & editing. **Pia Tio:** Writing – original draft, Writing – review & editing. **Manuel C. Voelkle:** Conceptualization, Writing – original draft, Writing – review & editing. **Marieke Wichers:** Conceptualization, Writing – original draft.

Declaration of competing interest

None.

Acknowledgments

This paper is based on the symposium “Psychological networks and time series models: Improving the analyses of complex clinical data”, which took place in Groningen (The Netherlands) on September 14–15, 2018. We are very grateful to Peter de Jonge and Tom Snijders for their contributions to this symposium, as well as for their feedback on earlier versions of the manuscript. The research that led to this paper was funded by the Netherlands Organization for Scientific Research Veni Grant (NWO-Veni 191G.037; Laura Bringmann) and by the German Research Foundation (DFG, grant number LU 660-10/1; Wolfgang Lutz).

References

Adolf, J. K., & Fried, E. I. (2019). Ergodicity is sufficient but not necessary for group-to-individual generalizability. *Proceedings of the National Academy of Sciences, 116*, 6540–6541. https://doi.org/10.1073/pnas.1818675116

Adolf, J., Schuurman, N. K., Borkenau, P., Borsboom, D., & Dolan, C. V. (2014). Measurement invariance within and between individuals: A distinct problem in testing the equivalence of intra-and inter-individual model structures. *Frontiers in Psychology, 5*, 883. https://doi.org/10.3389/fpsyg.2014.00883

Adolf, J. K., Loossens, T., Tuerlinckx, F., & Ceulemans, E. (2021). Optimal sampling rates for reliable continuous-time first-order autoregressive and vector autoregressive modeling. *Psychological Methods*. https://doi.org/10.1037/met0000398

Adolf, J. K., Voelkle, M. C., Brose, A., & Schmiedek, F. (2017). Capturing context-related change in emotional dynamics via fixed moderated time series analysis. *Multivariate Behavioral Research, 52*, 499–531. https://doi.org/10.1080/00273171.2017.1321978

AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Albers, C. J., & Bringmann, L. F. (2020). Inspecting gradual and abrupt changes in emotion dynamics with the time-varying change point autoregressive model. *European Journal of Psychological Assessment, 36*, 492–499. https://doi.org/10.1027/1015-5759/a000589

Anderlucci, L., & Viroli, C. (2015). Covariance pattern mixture models for the analysis of multivariate heterogeneous longitudinal data. *Annals of Applied Statistics, 9*(2), 777–800. https://doi.org/10.1214/15-AOAS816

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Arlot, S., Celisse, A., & Harchaoui, Z. (2012). *Kernel change-point detection*. Retrieved from https://arxiv.org/pdf/1202.3878v1.pdf.

Bak, M., Drukker, M., Hasmi, L., & van Os, J. (2016). An n=1 clinical network analysis of symptoms and treatment in psychosis. *PLoS One, 11*(9), e0162811. https://doi.org/10.1371/journal.pone.0162811

Ballesio, A., Aquino, M. R. J. V., Feige, B., Johann, A. F., Kyle, S. D., Spiegelhalder, K., et al. (2018). The effectiveness of behavioural and cognitive behavioural therapies for insomnia on depressive and fatigue symptoms: A systematic review and network meta-analysis. *Sleep Medicine Reviews, 37*, 114–129. https://doi.org/10.1016/j.smrv.2017.01.006

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Journal of the Royal Statistical Society, 8*(Supplement), 27–41. https://doi.org/10.2307/2983611

Bastiaansen, J. A., Kunkels, Y. K., Blaauw, F. J., Boker, S. M., Ceulemans, E., Chen, M., et al. (2020). Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research, 137*, 110211. https://doi.org/10.1016/j.jpsychores.2020.110211

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A. T. (Ed.). (1979). *Cognitive therapy of depression*. Guilford press.

Beck, A. T. (1985). *Anxiety disorders and phobias: A cognitive perspective*. NY: Basic Books.

Beck, A. T. (1988). Cognitive approaches to panic disorder: Theory and therapy. In S. Rachman (Ed.), *Panic: Psychological perspectives* (pp. 91–110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research, 52*(6), 789–804. https://doi.org/10.1080/00273171.2017.1373014

Beltz, A. M., Wright, A. G., Sprague, B. N., & Molenaar, P. C. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment, 23*(4), 447–458. https://doi.org/10.1177/1073191116648209

Blanken, T. F., Bathelt, J., Deserno, M. K., Voge, L., Borsboom, D., & Douw, L. (2021). Connecting brain and behavior in clinical neuroscience: A network approach. *Neuroscience & Biobehavioral Reviews, 130*, 81–90. https://doi.org/10.1016/j.neubiorev.2021.07.027

Blanken, T. F., Deserno, M. K., Dalege, J., Borsboom, D., Blanken, P., Kerkhof, G. A., et al. (2018). The role of stabilizing and communicating symptoms given overlapping communities in psychopathology networks. *Scientific Reports, 8*, 5854. https://doi.org/10.1038/s41598-018-24224-2

Blanken, T. F., Van der Zweerde, T., Van Straten, A., Van Someren, E. J. W., Borsboom, D., & Lancee, J. (2019). Introducing network intervention analysis to investigate sequential, symptom-specific treatment effects: A demonstration in co-occurring insomnia and depression. *Psychotherapy and Psychosomatics, 88*(1), 55–57. https://doi.org/10.1159/000495045

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. New York, NY: The Guilford Press.

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

Borsboom, D. (2017). Mental disorders, network models, and dynamical systems. In K. Kendler, & J. Parnas (Eds.), *Philosophical issues in Psychiatry IV: Psychiatric Nosology DSM-5 (International perspectives in philosophy and Psychiatry)* (pp. 80–97). Oxford University Press.

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

Borsboom, D., Cramer, A. O., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 42*, e2. https://doi.org/10.1017/S0140525X17002266

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

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., et al. (2021). Network analysis of multivariate data in psychological science. *Nature Review Methods Primers, 1*, 58. https://doi.org/10.1038/s43586-021-00055-w

Borsboom, D., Robinaugh, D. J., Psychosystems Group, Rhemtulla, M., & Cramer, A. O. (2018). Robustness and replicability of psychopathology networks. *World Psychiatry, 17*(2), 143–144. https://doi.org/10.1002/wps.20515

Bos, F. M., Snippe, E., Bruggeman, R., Wichers, M., & van der Krieke, L. (2019). Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services, 70*(11), 983–991. https://doi.org/10.1176/appi.ps.201900050

Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present and future. *Current Opinion in Psychology, 41*, 59–63. https://doi.org/10.1016/j.copsyc.2021.03.004

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

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review, 125*, 606–615. https://doi.org/10.1037/rev0000108

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. *Multivariate Behavioral Research, 53*(3), 293–314. https://doi.org/10.1080/00273171.2018.1439722

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods, 22*, 409–425. https://doi.org/10.1037/met0000085

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., et al. (2016). Assessing temporal emotion dynamics using networks. *Assessment, 23*(4), 425–435. https://doi.org/10.1177/1073191116645909

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

Brose, A., Lindenberger, U., & Schmiedek, F. (2013). Affective states contribute to trait reports of affective well-being. *Emotion, 13*(5), 940–948. https://doi.org/10.1037/a0032401

Brose, A., Schmiedek, F., Gerstorf, D., & Voelkle, M. C. (2020). The measurement of within-person affect variation. *Emotion, 20*(4), 677–699. https://doi.org/10.1037/emo0000583

Brose, A., Voelkle, M. C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2015). Differences in the between-person and within-person structures of affect are a matter of degree. *European Journal of Personality, 29*(1), 55–71. https://doi.org/10.1002/per.1961

Brouwer, M. E., Molenaar, N. M., Burger, H., Williams, A. D., Albers, C. J., Lambregtse-van den Berg, M. P., et al. (2020). Tapering antidepressants while receiving digital preventive cognitive therapy during pregnancy: An experience sampling methodology trial. *Frontiers in Psychiatry, 11*, 1116. https://doi.org/10.3389/fpsyt.2020.574357

Brouwer, M. E., Williams, A. D., Kennis, M., Fu, Z., Klein, N. S., Cuijpers, P., et al. (2019). Psychological theories of depressive relapse and recurrence: A systematic review and meta-analysis of prospective studies. *Clinical Psychology Review, 74*, 101773. https://doi.org/10.1016/j.cpr.2019.101773

Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (2018). VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. *Psychological Methods, 23*, 740–756. https://doi.org/10.1037/met0000178

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2018). Improved insight into and prediction of network dynamics by combining VAR and dimension reduction. *Multivariate Behavioral Research, 53*, 853–875. https://doi.org/10.1080/00273171.2018.1516540

Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., et al. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine, 18*, 99. https://doi.org/10.1186/s12916-020-01558-1

Cabrieto, J., Adolf, J., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2018c). Detecting long-lived autodependency changes in a multivariate system via change point detection and regime switching models. *Scientific Reports, 8*, 15637. https://doi.org/10.1038/s41598-018-33819-8

Cabrieto, J., Adolf, J., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2019). An objective, comprehensive and flexible statistical framework for detecting early warning signs of mental health problems. *Psychotherapy and Psychosomatics, 88*(3), 184–186. https://doi.org/10.1159/000494356

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., & Ceulemans, E. (2018a). Testing for the presence of correlation changes in a multivariate time series: A permutation based approach. *Scientific Reports, 8*, 769. https://doi.org/10.1038/s41598-017-19067-2

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Wilhelm, F., Liedlgruber, M., & Ceulemans, E. (2018b). Capturing correlation changes by applying kernel change point detection on the running correlations. *Information Sciences, 447*, 117–139. https://doi.org/10.1016/j.ins.2018.03.010

Chen, M., Chow, S. M., & Hunter, M. D. (2018). Stochastic differential equation models with time-varying parameters. In K. van Montfort, J. H. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 205–238). Cham: Springer.

Cramer, A. O., Van Borkulo, C. D., Giltay, E. J., Van Der Maas, H. L., Kendler, K. S., Scheffer, M., et al. (2016). Major depression as a complex dynamic system. *PloS one, 11*(12), e0167490. https://doi.org/10.1371/journal.pone.0167490

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences, 33*, 137–193. https://doi.org/10.1017/S0140525X09991567

Dablander, F., & Hinne, M. (2019). Node centrality measures are a poor substitute for causal inference. *Scientific Reports, 9*, 6846. https://doi.org/10.1038/s41598-019-43033-9

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. (2018). The Attitudinal Entropy (AE) Framework as a general theory of individual attitudes. *Psychological Inquiry, 29*(4), 175–193. https://doi.org/10.1080/1047840X.2018.1537246

De Haan-Rietdijk, S., Gottman, J. M., Bergeman, C. S., & Hamaker, E. L. (2016). Get over it! A multilevel threshold autoregressive model for state-dependent affect regulation. *Psychometrika, 81*(1), 217–241. https://doi.org/10.1007/s11336-014-9417-x

Deserno, M. K., Borsboom, D., Begeer, S., van Bork, R., Hinne, M., & Geurts, H. M. (2020). Highways to happiness for autistic adults? Perceived causal relations among clinicians. *PLoS One, 15*(12), e0243298. https://doi.org/10.1371/journal.pone.0243298

Doove, L. L., Wilderjans, T. F., Calcagnì, A., & Van Mechelen, I. (2017). Deriving optimal data-analytic regimes from benchmarking studies. *Computational Statistics & Data Analysis, 107*, 81–91. https://doi.org/10.1016/j.csda.2016.10.016

Dormann, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods, 20*(4), 489–505. https://doi.org/10.1037/met0000041

Driver, C. C., & Voelkle, M. C. (2018). Hierarchical Bayesian continuous time dynamic modeling. *Psychological Methods, 23*(4), 774–799. https://doi.org/10.1037/met0000168

Eisele, G., Lafit, G., Vachon, H., Kuppens, P., Houben, M., Myin-Germeys, I., et al. (2021). Affective structure, measurement invariance, and reliability across different experience sampling protocols. *Journal of Research in Personality, 92*, 104094. https://doi.org/10.1016/j.jrp.2021.104094

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., et al. (2020). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*. https://doi.org/10.1177/1073191120957102

Ellis, A. (1956). An operational reformulation of some of the basic principles of psychoanalysis. In H. Feigl, & M. Scriven (Eds.), *The foundations of science and the concepts of psychology and psychoanalysis* (pp. 131–154). Minneapolis: University of Minnesota Press.

Epskamp, S. (2019). Reproducibility and replicability in a fast-paced methodological world. *Advances in Methods and Practices in Psychological Science, 2*(2), 145–155. https://doi.org/10.1177/2515245919847421

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika, 85*, 206–231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods, 23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Isvoranu, A., & Cheung, M. W. (2021). Meta-analytic gaussian network aggregation. *Psychometrika*. https://doi.org/10.1007/s11336-021-09764-3.
Epskamp, S., Maris, G., Waldorp, L., & Borsboom, D. (2018). Network psychometrics. In P. Irwing, D. Hughes, & T. Booth (Eds.), *Handbook of psychometrics*. New York, NY, USA: Wiley.
Epskamp, S., Rhemtulla, M. T., & Borsboom, D. (2017). Generalized network psychometrics: combining network and latent variable models. *Psychometrika, 82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x
Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823
Ernst, A. F., Albers, C. J., Jeronimus, B. F., & Timmerman, M. E. (2020). Inter-individual differences in multivariate time-series. *European Journal of Psychological Assessment, 36*(3), 482–491. https://doi.org/10.1027/1015-5759/a000578
Eronen, M. I. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology, 59*, 100785. https://doi.org/10.1016/j.newideapsych.2020.100785
Eronen, M. I. (2021). The levels problem in psychopathology. *Psychological Medicine, 51*(6), 927–933. https://doi.org/10.1017/S0033291719002514
Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 8*, 1044–1056. https://doi.org/10.1037/abn0000311
Forbes, M. K., Wright, A. G., Markon, K. E., & Krueger, R. F. (2019). Quantifying the reliability and replicability of psychopathology network characteristics. *Multivariate Behavioral Research, 56*, 224–242. https://doi.org/10.1080/00273171.2019.1616526
Fried, E. I., Bockting, C., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, A. O., et al. (2015). From loss to loneliness: The relationship between bereavement and depressive symptoms. *Journal of Abnormal Psychology, 124*(2), 256–265. https://doi.org/10.1037/abn0000028
Fried, E. I., & Cramer, A. O. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science, 12*(6), 999–1020. https://doi.org/10.1177/1745691617705892
Fried, E. I., Eidhof, M. B., Palic, S., Costantini, G., Huisman-van Dijk, H. M., Bockting, C. L., et al. (2018). Replicability and generalizability of posttraumatic stress disorder (PTSD) networks: A cross-cultural multisite study of PTSD symptoms in four trauma patient samples. *Clinical Psychological Science, 6*(3), 335–351. https://doi.org/10.1177/2167702617745092
Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z
Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage, 63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026
Hallquist, M. N., Wright, A. G., & Molenaar, P. C. (2021). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research, 56*(2), 199–223. https://doi.org/10.1080/00273171.2019.1640103
Hamaker, E. (2009). Using information criteria to determine the number of regimes in threshold autoregressive models. *Journal of Mathematical Psychology, 53*(6), 518–529. https://doi.org/10.1016/j.jmp.2009.07.006
Hamaker, E. L. (2012). Why researchers should think "within-person" a paradigmatic rationale. In M. R. Mehl, & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York, NY: Guilford Publications.
Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research, 53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819
Hamaker, E. L., & Dolan, C. V. (2009). Idiographic data analysis: Quantitative methods—from simple to advanced. In *Dynamic process methodology in the social and developmental sciences* (pp. 191–216). New York, NY: Springer.
Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. https://doi.org/10.1037/a0038889
Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton university press.
Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences, 18*, 83–90. https://doi.org/10.1016/j.cobeha.2017.07.018
Haslbeck, J. M. B., Bringmann, L., & Waldorp, L. J. (2021). A tutorial on estimating time-varying vector autoregressive models. *Multivariate Behavioral Research, 56*(1), 120–149. https://doi.org/10.1080/00273171.2020.1743630
Haslbeck, J. M. B., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (2021). Modeling psychopathology: From data models to formal theories. *Psychological Methods*. https://doi.org/10.1037/met0000303
Haslbeck, J. M. B., & Waldorp, L. J. (2020). mgm: Structure estimation for time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software, 93*(8), 1–46. https://doi.org/10.18637/jss.v093.i08
Helmich, M. A., Olthof, M., Oldehinkel, A. J., Wichers, M., Bringmann, L. F., & Smit, A. C. (2021). Early warning signals and critical transitions in psychopathology: Challenges and recommendations. *Current Opinion in Psychology, 41*, 51–58. https://doi.org/10.1016/j.copsyc.2021.02.008
Helmich, M. A., Snippe, E., Kunkels, Y. K., Riese, H., Smit, A. C., & Wichers, M. (2021). *Transitions in Depression (TRANS-ID) Recovery: Study protocol for a repeated intensive longitudinal n = 1 study design to search for personalized early warning signals of critical transitions towards improvement in depression*. Retrieved from. https://doi.org/10.31234/osf.io/fertq.
Henry, T. R., Robinaugh, D. J., & Fried, E. I. (2021). On the control of psychological networks. *Psychometrika*, 1–26. https://doi.org/10.1007/s11336-021-09796-9
Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., et al. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry, 5*(3), 237–286. https://doi.org/10.1016/S2215-0366(17)30513-8
Insel, T. R. (2017). Digital phenotyping: Technology for a new science of behavior. *Journal of the American Medical Association, 318*(13), 1215–1216. https://doi.org/10.1001/jama.2017.11295
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*, 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379
Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science, 348*(6234), 499–500. https://doi.org/10.1126/science.aab2358
Isvoranu, A. M., Epskamp, S., & Cheung, M. W. L. (2021). Network models of posttraumatic stress disorder: A meta-analysis. *Journal of Abnormal Psychology, 130*(8), 841–861. https://doi.org/10.1037/abn0000704
Isvoranu, A. M., van Borkulo, C. D., Boyette, L. L., Wigman, J. T., Vinkers, C. H., Borsboom, D., et al. (2016). A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms. *Schizophrenia Bulletin, 43*(1), 187–196. https://doi.org/10.1093/schbul/sbw055
Jacobson, N. C., Lekkas, D., Huang, R., & Thomas, N. (2021). Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17–18 years. *Journal of Affective Disorders, 282*, 104–111. https://doi.org/10.1016/j.jad.2020.12.086
Jones, P. J., Heeren, A., & McNally, R. J. (2017). Commentary: A network theory of mental disorders. *Frontiers in Psychology, 8*, 1305. https://doi.org/10.3389/fpsyg.2017.01305
Jones, P. J., Williams, D. R., & McNally, R. J. (2021). Sampling variability is not nonreplication: A Bayesian reanalysis of Forbes, Wright, Markon, and Krueger. *Multivariate Behavioral Research, 56*(2), 249–255. https://doi.org/10.1080/00273171.2020.1797460
Kalis, A., & Borsboom, D. (2020). Folk psychology as a causal language. *Theory & Psychology, 30*(5), 723–728. https://doi.org/10.1177/0959354320933940
Kim, C. J., & Nelson, C. R. (1999). *State-Space models with regime switching: Classical and gibbs-sampling approaches with applications*. Cambridge: MIT Press.
Kirk, D. E. (2004). *Optimal control theory: an introduction*. Mineola, New York: Dover Publications, Inc.
Kroeze, R., Van Veen, D., Servaas, M. N., Bastiaansen, J. A., Oude Voshaar, R., Borsboom, D., et al. (2016). Personalized feedback on symptom dynamics of psychopathology: A proof-of-principle study. *Journal for Person-Oriented Research, 3*(1), 1–10. https://doi.org/10.17505/jpor.2017.01
Lafit, G., Adolf, J., Dejonckheere, E., Germeys, I., Viechtbauer, W., & Ceulemans, E. (2020). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial to perform power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science, 4*(1), Article 2515245920978738. https://doi.org/10.1177/2515245920978738
Lafit, G., Meers, K., & Ceulemans, E. (2021). A systematic study into the factors that affect the predictive accuracy of multilevel VAR(1) models. *Psychometrika*. https://doi.org/10.1007/s11336-021-09803-z
Liu, S. (2017). Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels. *British Journal of Mathematical and Statistical Psychology, 70*(3), 480–498. https://doi.org/10.1111/bmsp.12096
Lunansky, G., van Borkulo, C., & Borsboom, D. (2020). Personality, resilience, and psychopathology: A model for the interaction between slow and fast network processes in the context of mental health. *European Journal of Personality, 34*(6), 969–987. https://doi.org/10.1002/per.2263
Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer.
Lutz, W., Schwartz, B., Hofmann, S. G., Fisher, A. J., Husen, K., & Rubel, J. A. (2018). Network analysis predicts treatment dropout in patients with mood and anxiety disorders. *Scientific Reports, 8*, 7819. https://doi.org/10.1038/s41598-018-25953-0
Mansell, W., & Marken, R. S. (2015). The origins and future of control theory in psychology. *Review of General Psychology, 19*(4), 425–430. https://doi.org/10.1037/gpr0000057
McNally, R. J. (1990). Psychological approaches to panic disorder: A review. *Psychological Bulletin, 108*(3), 403–419. https://doi.org/10.1037/0033-2909.108.3.403
McNally, R. J. (1994). *Panic disorder: A critical analysis*. New York: Guilford Press.
McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy, 86*, 95–104. https://doi.org/10.2016/j.brat.2016.06.006
McNally, R. J. (2021). Network analysis of psychopathology: Controversies and challenges. *Annual Review of Clinical Psychology, 17*, 31–53. https://doi.org/10.1146/annurev-clinpsy-081219-092850
McNally, R. J., Mair, P., Mugno, B. L., & Riemann, B. C. (2017). Co-morbid obsessive-compulsive disorder and depression: a Bayesian network approach. *Psychological Medicine, 47*(7), 1204. https://doi.org/10.1017/S0033291716003287
McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods, 25*(5), 610–635. https://doi.org/10.1037/met0000250
Mehl, M. R., & Conner, T. S. (Eds.). (2012). *Handbook of research methods for studying daily life*. New York: Guilford Press.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2), 246–268. https://doi.org/10.1037/0033-295X.102.2.246
Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1
Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & Van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine, 39*(9), 1533–1547. https://doi.org/10.1017/S0033291708004947
Newman, M. (2018). *Networks*. Oxford: Oxford University Press.
Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica, 49*(6), 1417–1426. https://doi.org/10.2307/1911408
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (Third ed.). New York: McGraw-Hill, Inc.
Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2016). Bayesian data analysis with the bivariate hierarchical Ornstein-Uhlenbeck process model. *Multivariate Behavioral Research, 51*(1), 106–119. https://doi.org/10.1080/00273171.2015.1110512
Oude Maatman, F. (2020). Folk psychology and network theory: Fact or gamble? A reply to Kalis and Borsboom. *Theory & Psychology, 30*(5), 729–734. https://doi.org/10.1177/0959354320952863
Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*(2), 199–215. https://doi.org/10.1007/BF02294374
Oud, J. H. L., Voelkle, M. C., & Driver, C. C. (2018). SEM based CARMA time series modeling for arbitrary N. *Multivariate Behavioral Research, 53*(1), 36–56. https://doi.org/10.1080/00273171.2017.1383224
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*(2), 292–300. https://doi.org/10.1177/2167702614540645
R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.
Richardson, T., & Spirtes, P. (2002). Ancestral graph Markov models. *Annals of Statistics, 30*(4), 962–1030. https://doi.org/10.1214/aos/1031689015
Robinaugh, D. J., Haslbeck, J. M. B., Waldorp, L. J., Kossakowski, J. J., Fried, E. I., Millner, A. J., et al. (2019). *Advancing the network theory of mental disorders: A computational model of Panic Disorder*. PsyArXiv. https://doi.org/10.31234/osf.io/km37w
Riese, H., & Wichers, M. (2021). Comment on: Eronen MI (2019). The levels problem in psychopathology. *Psychological Medicine, 51*(3), 525–526. https://doi.org/10.1017/S0033291719003623
Robinaugh, D., Hoekstra, R., Toner, E., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine, 50*(3), 353–366. https://doi.org/10.1017/S0033291719003404
Rodebaugh, T. L., Tonge, N. A., Piccirillo, M. L., Fried, E., Horenstein, A., Morrison, A. S., et al. (2018). Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? *Journal of Consulting and Clinical Psychology, 86*(10), 831–844. https://doi.org/10.1037/ccp0000336
Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science, 1*(1), 27–42. https://doi.org/10.1177/2515245917745629
Rubel, J. A., Fisher, A. J., Husen, K., & Lutz, W. (2018). Translating person-specific network models into personalized treatments. Development and demonstration of the dynamic assessment treatment algorithm for idiographic networks. *Psychotherapy and Psychosomatics, 87*, 249–251. https://doi.org/10.1159/000487769
Ruzzano, L., Borsboom, D., & Geurts, H. M. (2015). Repetitive behaviors in autism and obsessive compulsive disorder: New perspectives from a network analysis. *Journal of Autism and Developmental Disorders, 45*, 192–202. https://doi.org/10.1007/s10803-014-2204-9
Ryan, O. (2020). *Dynamic Systems and Causal Structures in Psychology: Connecting Data and Theory [Doctoral dissertation]*. Utrecht University.
Ryan, O., & Hamaker, E. L. (2021). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika, 1–39*. https://doi.org/10.1007/s11336-021-09767-0
Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous time approach to intensive longitudinal data: What, Why and How? In K.v. Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 27–54). New York: Springer.
Schmiedek, F., Lövdén, M., von Oertzen, T., & Lindenberger, U. (2020). Within-person structures of daily cognitive performance differ from between-person structures of cognitive abilities. *PeerJ, 8*, e9290. https://doi.org/10.7717/peerj.9290
Schreuder, M. J., Groen, R. N., Wigman, J. T., Hartman, C. A., & Wichers, M. (2020). Measuring psychopathology as it unfolds in daily life: Addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry, 20*, 351. https://doi.org/10.1186/s12888-020-02674-1
Shrout, P. E., Stadler, G., Lane, S. P., McClure, M. J., Jackson, G. L., Clavél, F. D., et al. (2018). Initial elevation bias in subjective reports. *Proceedings of the National Academy of Sciences, 115*(1), E15–E23. https://doi.org/10.1073/pnas.1712277115
Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass, 2*(1), 414–433. https://doi.org/10.1111/j.1751-9004.2007.00044.x
Slofstra, C., Nauta, M. H., Bringmann, L. F., Klein, N. S., Albers, C. J., Batalas, N., et al. (2018). Individual negative affective trajectories can be detected during different depressive relapse prevention strategies. *Psychotherapy and Psychosomatics, 87*(4), 243–245. https://doi.org/10.1159/000489044
Snippe, E., Viechtbauer, W., Geschwind, N., Klippel, A., de Jonge, P., & Wichers, M. (2017). The impact of treatments for depression on the dynamic network structure of mental states: Two randomized controlled trials. *Scientific Reports, 7*, 46523. https://doi.org/10.1038/srep46523
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B, 58*(1), 267–288.
Tong, H., & Lim, K. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society: Series B (Methodological), 42*(3), 245–268.
Trull, T. J., Lane, S. P., Koval, P., & Ebner-Priemer, U. W. (2015). Affective dynamics in psychopathology. *Emotion Review, 7*(4), 355–361. https://doi.org/10.1177/1754073915590617
Vachon, H., Rintala, A., Viechtbauer, W., & Myin-Germeys, I. (2018). Data quality and feasibility of the experience sampling method across the spectrum of severe psychiatric disorders: A protocol for a systematic review and meta-analysis. *Systematic Reviews, 7*, 7. https://doi.org/10.1186/s13643-018-0673-1
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. https://doi.org/10.1038/srep05918
van Borkulo, C., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R., Borsboom, D., & Waldorp, L. (2017). Comparing network structures on three aspects: A permutation test. Retrieved from http://doi.org/10.13140/RG.2.2.29455.38569.
van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., et al. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences, 111*(1), 87–92. https://doi.org/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. https://doi.org/10.1037/0033-295X.113.4.842
van der Zweerde, T., Van Straten, A., Effting, M., Kyle, S. D., & Lancee, J. (2019). Does online insomnia treatment reduce depressive symptoms? A randomized controlled trial in individuals with both insomnia and depressive symptoms. *Psychological Medicine, 49*(3), 501–509. https://doi.org/10.1017/S0033291718001149
van Montfort, K., Oud, J. H. L., & Voelkle, M. C. (Eds.). (2018). *Continuous time modeling in the behavioral and related sciences*. New York, NY: Springer.
Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research, 49*, 193–213. https://doi.org/10.1080/00273171.2014.889593
Voelkle, M. C., Gische, C., Driver, C. C., & Lindenberger, U. (2018). The role of time in the quest for understanding psychological mechanisms. *Multivariate Behavioral Research, 53*(6), 782–805. https://doi.org/10.1080/00273171.2018.1496813
Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods, 17*, 176–192. https://doi.org/10.1037/a0027543
von Klipstein, L., Riese, H., Servaas, M. N., & Schoevers, R. A. (2020). Using person-specific networks in psychotherapy: Challenges, limitations, and how we could use them anyway. *BMC Medicine, 18*(1), 345. https://doi.org/10.1186/s12916-020-01818-0
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440–442. https://doi.org/10.1038/30918
Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine, 44*(7), 1349–1360. https://doi.org/10.1017/S0033291713001979
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. https://doi.org/10.1159/000441458
Wichers, M., Riese, H., Hodges, T. M., Snippe, E., & Bos, F. M. (2021). A narrative review of network studies in depression: What different methodological approaches tell us about depression. *Frontiers in Psychiatry, 12*, 719490. https://doi.org/10.3389/fpsyt.2021.719490
Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine, 49*(3), 380–387. https://doi.org/10.1017/S0033291718002064
Wichers, M., Simons, C. J. P., Kramer, I. M. A., Hartmann, J. A., Lothmann, C., Myin-Germeys, I., et al. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica, 124*(4), 262–272. https://doi.org/10.1111/j.1600-0447.2011.01749.x
Wichers, M., Smit, A. C., & Snippe, E. (2020). Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: A confirmatory single-subject time-series study. *Journal for Person-Oriented Research, 6*(1), 1–15. https://doi.org/10.17505/jpor.2020.22042
Wichers, M., Wigman, J. T., Bringmann, L. F., & de Jonge, P. (2017). Mental disorders as networks: Some cautionary reflections on a promising approach. *Social Psychiatry and Psychiatric Epidemiology, 52*(2), 143–145. https://doi.org/10.1007/s00127-016-1335-z
Wigman, J., van Os, J., Borsboom, D., Wardenaar, K., Epskamp, S., Klippel, A., … Wichers, M. (2015). Exploring the underlying structure of mental disorders: Cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine, 45*(11), 2375–2387. https://doi.org/10.1017/S0033291715000331

Williams, D. R., Rast, P., Pericchi, L. R., & Mulder, J. (2020). Comparing Gaussian graphical models with the posterior predictive distribution and Bayesian model selection. *Psychological Methods, 25*(5), 653–672. https://doi.org/10.1037/met0000254

Wittenborn, A. K., Rahmandad, H., Rick, J., & Hosseinichimeh, N. (2016). Depression as a systemic syndrome: Mapping the feedback loops of major depressive disorder. *Psychological Medicine, 46*(3), 551–562. https://doi.org/10.1017/S0033291715002044

Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research, 91*(2), 303–347. https://doi.org/10.1111/phpr.12095

Wright, A. G., & Woods, W. C. (2020). Personalized models of psychopathology. *Annual Review of Clinical Psychology, 16*, 49–74. https://doi.org/10.1146/annurev-clinpsy-102419-125032

Wright, A. G., & Zimmermann, J. (2019). Applied ambulatory assessment: Integrating idiographic and nomothetic principles of measurement. *Psychological Assessment, 31*(12), 1467–1480. https://doi.org/10.1037/pas0000685

Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., et al. (2020). From data to causes I: Building A general cross-lagged panel model (GCLM). *Organizational Research Methods, 23*(4), 651–687. https://doi.org/10.1177/1094428119847278

Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., et al. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods, 23*(4), 688–716. https://doi.org/10.1177/1094428119847280