

Out of bounds? The boundary specification problem for centrality in psychological networks

Zachary P. Neal and Jennifer Watling Neal
Michigan State University

The analysis of psychological networks has become common in multiple subfields including clinical, social, and personality psychology, where the focus is often on identifying highly central nodes that represent symptoms, beliefs, or traits. However, the boundaries of these networks are often ambiguous and relevant nodes are often missing from the network. In this paper, we use a series of simulations to show that even under typical conditions of missingness, the centrality of nodes in an empirical psychological network are poorly correlated or uncorrelated with their centrality in a hypothetical ‘true’ psychological network, and thus are invalid. We illustrate the implications of this lack of validity using an empirical example drawn from a recent study of political belief system networks, demonstrating that the original study would have drawn incorrect conclusions about American’s most central political beliefs. We conclude by recommending that centrality measures should be computed and interpreted only in psychological networks that include (nearly) all the nodes inside a theoretically meaningful boundary.

Introduction

The estimation, analysis, and visualization of psychological networks has become common in multiple subfields. Clinical psychologists have studied networks of symptoms to map the structure of psychopathologies (e.g., Borsboom & Cramer, 2013; Cramer, Waldorp, Van Der Maas, & Borsboom, 2010; Preszler, Marcus, Edens, & McDermott, 2018). Social psychologists have studied networks of political beliefs to define groups’ belief systems (e.g., Brandt, 2020; Brandt, Sibley, & Osborne, 2019; Dalege et al., 2016; Schlicht-Schmälzle, Chykina, & Schmälzle, 2018). Personality psychologists have studied networks of personality traits to understand how they combine to yield higher-order personality dimensions or types (e.g., Christensen, Golino, & Silvia, 2020; Costantini et al., 2019; Dalege, Borsboom, van Harreveld, & van der Maas, 2019; Marcus, Preszler, & Zeigler-Hill, 2018). In many cases, the analysis involves identifying highly central nodes in the network, with the goal of finding the most important or influential symptoms, beliefs, or traits. These methods have been adapted from theoretical and methodological traditions in social network

analysis, but this adaptation has not adequately considered one problem that is well-documented in that literature: the boundary specification problem.

The boundary specification problem refers to the difficulty of deciding which nodes should be included in a network. Although all nodes in a theoretically- or naturalistically-defined population should be included in a network, we argue that this is impossible or impractical for many psychological networks because the boundary of the population is unknown or very wide. For example, because it is unknown how many exhibitable traits or holdable beliefs exist, it is impossible to measure a complete network that includes nodes representing all traits or all beliefs, or indeed even to know how many nodes a network is missing. In practice, such networks are bounded by whatever traits or beliefs have been measured. In this paper, we use a simulation to demonstrate that a psychological network estimated from such incomplete data yields invalid node centralities, then provide an empirical example from the context of political belief systems to illustrate how the boundary specification problem can lead to erroneous conclusions in practice.

Background

Psychological networks and centrality

A network is a collection of objects (called *nodes* or *vertices*) and the relationships (called *edges*) that exist between pairs of these objects. Psychologists have been measuring *social* networks, where the nodes are people and the edges are social relationships (e.g., friendship, advice), for many decades (e.g., Festinger, Schachter, & Back, 1950; Moreno

Correspondence concerning this article should be addressed to Zachary Neal, Department of Psychology, Michigan State University, East Lansing, MI 48824; zpcneal@msu.edu. Earlier versions of this manuscript were posted as pre-prints on PsyArXiv at <https://psyarxiv.com/nz6k3/>.

Zachary P. Neal  <https://orcid.org/0000-0003-3076-4995>

Jennifer Watling Neal  <https://orcid.org/0000-0002-7749-8121>

& Jennings, 1934; Travers & Milgram, 1977).¹ More recently, psychologists have begun examining networks in which the nodes are measured items or scales, and the edges are the statistical associations among them. In this paper, we use the term *psychological network* to refer to this latter type of network.

Psychological networks have grown in popularity and are now studied in many subfields, including clinical (e.g., Borsboom & Cramer, 2013; Cramer et al., 2010; Preszler et al., 2018), social (e.g., Brandt, 2020; Brandt et al., 2019; Dalege et al., 2016; Schlicht-Schmälzle et al., 2018), and political psychology (e.g., Christensen et al., 2020; Dalege et al., 2019; Marcus et al., 2018). Across all of these subfields, the estimation of a psychological network begins by measuring, for example, a set of beliefs held by participants. These raw data take the form of a conventional rectangular case-by-variable dataset. They are then transformed into a psychological network, which takes the form of a square variable-by-variable matrix containing the statistical associations between pairs of variables.² A variety of methods exist for performing this transformation (Epskamp, Boersboom, & Fried, 2018; Forbes, Wright, Markon, & Krueger, 2017a). However, among the most common are Gaussian graphical models (GGM), which estimate the edge weights so they can be interpreted as partial correlation coefficients. Methods for estimating and analyzing psychological networks have been the subject of significant debate recently (Borsboom et al., 2017; Bringmann & Eronen, 2018; Epskamp, Fried, et al., 2018; Forbes et al., 2017a; Forbes, Wright, Markon, & Krueger, 2017b, 2021, 2019; Fried, van Borkulo, & Epskamp, 2020; Hallquist, Wright, & Molenaar, 2019; Jones, Williams, & McNally, 2020; Steinley, Hoffman, Brusco, & Sher, 2017). While this debate has clarified many issues, there remain questions about the meaningfulness of one family of measures that is widely used in psychological networks: centrality.

A node's centrality in a network aims to capture its 'importance' given the network's structure. Centrality is not a single measure, but refers to a family of measures that each capture different conceptualizations of nodal importance (for an overview and comparison of centrality measures, see Borgatti & Everett, 2006). In psychological networks, a node's centrality is sometimes interpreted as evidence of its causal influence over other nodes. For example, in their review of psychopathology networks, Borsboom and Cramer (2013) recommend the use of centrality measures to identify symptoms that may cause other symptoms. In other cases, the focus is not explicitly causal, but instead aims to descriptively identify the most central nodes. For example, in their study of New Zealanders' political belief system, Brandt et al. (2019) found that beliefs about groups were more central than beliefs about issues. Similarly, in their study of dark personality traits, Marcus et al. (2018) found that in-

terpersonal manipulation and callousness were consistently central. Bringmann et al. (2019) articulated several concerns about what, if anything, centrality measures measure in psychological networks. Although they offered several potential solutions, including developing new centrality measures with psychological networks in mind, they mentioned in passing one problem that we contend is so severe that even solving all other problems would still fail to make centrality useful in most psychological networks: the boundary specification problem.

The boundary specification problem and centrality

The boundary specification problem was first described in the study of social networks, and refers to the *a priori* specification of inclusion criteria for the nodes in a network. Whole network analytic techniques, including measuring centrality, require that the network include the entire population of nodes, not a sample of nodes from a population. Inclusion of the population of nodes requires specifying the boundary of the population. Although some social network analysts study systems where these boundaries are easy to specify (e.g., all students in a classroom, all employees in an organization), others study systems where the boundaries are more ambiguous (e.g., partners in a sexual network). Laumann, Marsden, and Prensky (1992) explain that this ambiguity is problematic because missing nodes lead to inaccurate measurement of the network and invalid network analyses: "...it is obviously of great consequence if a key intervening actor...is omitted due to oversight or use of data that are merely convenient; such an error, because it distorts the overall configuration of actors in a system, may render an entire analysis meaningless" (p. 63).

When nodes are missing, an observed network's topology often differs from the unobserved complete network's topology. Simulation studies varying the amount of node missingness have demonstrated that this is particularly problematic for measuring centrality. In these studies, researchers started with simulated or empirical social networks that they treated as 'true' networks, then dropped some nodes to obtain 'observed' networks, measuring the validity of centrality measures by comparing the centrality scores of nodes appearing in both the true and observed networks. For example, using simulated Erdős-Rényi random graphs to examine the effects of varying levels of node missingness on degree, betweenness, closeness, and eigenvector centrality,

¹Although Moreno was listed as the sole author of this work, "it is impossible to overestimate the importance of [Helen Hall] Jennings" in this work (L. Freeman, 2004, p. 35). Therefore, we cite it as co-authored by Moreno and Jennings.

²This transformation is closely related to what is known in network science as bipartite projection (Breiger, 1974; Latapy, Magnien, & Del Vecchio, 2008; Neal, 2014).

Borgatti, Carley, and Krackhardt (2006) found that “accuracy not only declines with increasing [node missingness], but does so predictably and monotonically” (p. 134). Likewise, additional studies using a variety of different empirical networks and simulation designs have also demonstrated that missing nodes reduce the validity of many centrality measures (Costenbader & Valente, 2003; Galaskiewicz, 1991; Smith & Moody, 2013; Smith, Moody, & Morgan, 2017).

These simulation studies have examined several factors that may affect the extent to which node missingness impacts the validity of centrality measures, including the type of centrality measure, the network’s structure, and the type of missingness. First, although all centrality measures’ validity declines as more nodes are missing (Borgatti et al., 2006), the decline in validity is more dramatic for centrality measures that depend on the whole network’s structure, such as closeness and betweenness (Costenbader & Valente, 2003; Smith & Moody, 2013). Second, there is weak evidence that the missingness-validity relationship is moderated by network size, and mixed evidence that it is moderated by network density (Borgatti et al., 2006; Galaskiewicz, 1991). Finally, there is limited evidence that validity declines more rapidly when more central nodes are systematically missing (Smith et al., 2017). In some cases these findings can help researchers mitigate invalidity (e.g., by choosing a robust centrality measure), while in other cases they have limited practical value (e.g., because a researcher cannot know whether the missing nodes were central or not). Notably, this work has focused on unweighted networks and unweighted centrality measures. However, because the edges in psychological networks are weighted using partial correlations, they offer limited insights in this context.

The problem for psychological networks

Figure 1 illustrates the process of studying centrality in an empirical psychological network, situating the present study among prior related work, and highlighting why studying centrality in psychological networks may be problematic. In this section, we use bold capital letters to refer to specific steps in figure 1.

Measurement of a psychological network begins by empirically measuring variables that capture some of the beliefs, symptoms, or traits held by participants (**D**). The measured variables represent a subset drawn from the unobserved set of *all* the beliefs, symptoms or traits held by those participants (**A**). A psychological network is estimated from the empirically measured data using one of several methods that yield a network containing a node for each measured belief, symptom, or trait (Epskamp, Boorsboom, & Fried, 2018, **E**). This observed psychological network is a subgraph of the unobserved true network, which contains a nodes for *all* the beliefs, symptoms or traits held by the participants (**B**). Finally, the centralities of nodes in the observed psychological

network are computed using graph theoretic methods (Borgatti & Everett, 2006), which yield a score for each measured belief, symptom, or trait (**F**). To be valid, these scores should match the scores that would be obtained from the true network, which are the criterion (**C**).

Prior research has investigated the effect of incomplete data in two ways. First, as we discuss above, a significant body of work in the social network analysis literature has investigated *what happens if some nodes are missing*, finding a monotonic decline in the validity of centrality scores (e.g., Borgatti et al., 2006; Costenbader & Valente, 2003; Smith & Moody, 2013, in figure 1: **B**→**E**→**F**↔**C**). Second, an emerging body of work in the psychological network literature has investigated *what happens if some participants are dropped* (e.g., Beard et al., 2016; Briganti, Kempenaers, Braun, Fried, & Linkowski, 2018; Epskamp, Boorsboom, & Fried, 2018). In this prior work, some participants (i.e., cases) are dropped from the empirical data, a new network is estimated from which centrality scores are computed and compared to those from the full sample (in figure 1: **D**→**G**→**H**→**I**↔**F**). Using bootstrap methods, the goal is to determine whether, in a given sample, centrality scores exhibit sufficient stability, which is typically evaluated using the CS-coefficient. Answering both of these questions in the context of psychological networks is facilitated by the *bootnet* package, which automates both node-drop and case-drop bootstrapping (Epskamp, Boorsboom, & Fried, 2018).

Although prior work has demonstrated that node-missingness and case-dropping can lead to invalidity and instability in centrality scores, respectively, we investigate a different form of missingness in this study. We ask *what happens when some variables are unmeasured?* The path in figure 1 we investigate (**A**→**D**→**E**→**F**↔**C**) may look similar to prior work on node-missingness, but missingness in psychological networks occurs further ‘upstream,’ not among nodes in a network, but among the variables from which the network is estimated. There are two reasons that unmeasured variables may lead to invalid centrality scores. First, when a variable is unmeasured in the empirical data, the corresponding node will be missing in the empirical network, thereby leading to the type of invalidity previously documented in the social network literature. Second, when a variable is unmeasured in the empirical data, all the edge weights among the remaining nodes may change when the network is estimated because psychological network estimation techniques rely on *partial* correlations to define edge weights. Thus, in the psychological network context, an unmeasured variable results not only in a missing node, but potentially also in different edge weights among the nodes that are present. Epskamp, Boorsboom, and Fried (2018) commented on this potential risk, noting that “dropping 50% of the nodes leads to entirely different network structures” (p. 200), but did not investigate its impact on centrality scores.

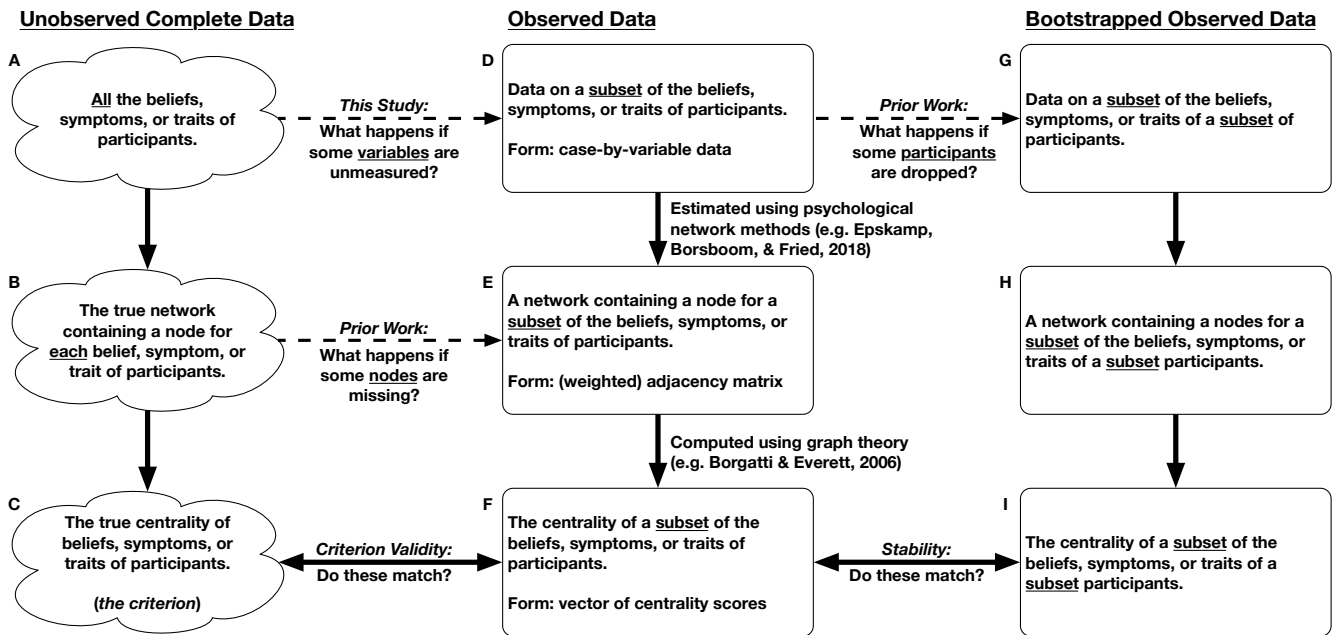


Figure 1

Boundary specification, psychological network construction, and assessments of stability and validity.

This observation by Epskamp, Boorsboom, and Fried (2018) highlights a second reason that psychological networks may be different from social networks in a way that could exacerbate the invalidity of centrality scores. When psychological networks are missing nodes, the percentage of missing nodes may be quite large. The size of the universe of beliefs, symptoms, or traits is often unknown, but is likely to be large because it consists of, for example, all holdable beliefs or all exhibitable traits. Even when the scope of the universe is theoretically bounded (e.g., political beliefs), it is likely still very large (e.g., all holdable political beliefs). Because the size of these universes is often unknown, it is difficult to know how many nodes are missing in psychological networks, but some examples from the literature suggest it is substantial.

In some cases, individual items serve as the nodes in psychological networks. For example, Brandt (2020) computed centrality scores from a network of 19 political belief items. It is impossible to know how many holdable political beliefs exist, but Curry, Alfano, Brandt, and Pelican (2020) hypothesize that morality is a combinatorial system and speculate there could be more than 8 million distinct moral values. However, even if we very conservatively estimate there are only 200 holdable political beliefs, Brandt's network would still be missing more than 90% of the possible nodes. In a similar example, Preszler et al. (2018) computed centrality scores from a network of either 12 or 20 personality trait items. Again, it is impossible to know how many exhibitable indicators of personality traits exist, but there are 3,320 items

in the International Personality Item Pool (IPIP; Goldberg et al., 2006). This suggests that Preszler et al.'s network could be missing more than 99% of the possible nodes. Because individual beliefs or indicators of personality traits are so numerous, psychological networks that treat items-as-nodes risk very high levels of missingness.

In other cases, entire domains or dimensions measured from multi-item scales serve as the nodes in psychological networks. For example, Marcus et al. (2018) computed centrality scores from a network of 9 dimensions of dark personality. However, they note as a limitation that "the networks did not include other dark personality traits, most notably sadism...and did not include broadband personality traits such as agreeableness" (p. 62). Although it is difficult to know how many dimensions of personality exist, considering only those that the authors specifically identify as missing, their network could be missing 50% of the possible nodes. Because scales are, by definition, less numerous than items, both true and observed scales-as-nodes psychological networks will be smaller in size. However, because multi-item scales are still costly to collect, such networks may still exhibit high levels of missingness.

Overview of Analyses

To investigate the impact of the boundary specification problem on centrality measures in psychological networks, we conduct a simulation, then provide a concrete empirical example. In the simulation, we generate synthetic 'true' psy-

chological networks of varying size and structure, computing centrality scores from these networks, which we treat as a criterion. We then compare these criterion centrality scores to centrality scores computed from an empirical psychological network that has been estimated from data with unmeasured variables and thus is missing some nodes. In the empirical example, we illustrate the practical impact of boundary specification using data from a recent political belief network study (Brandt, 2020), comparing the original results to those obtained when the network is estimated from data that contains more variables. The replication code for both analyses is available at <http://osf.io/x4nbj>.

Simulation

Methods

Following the lead of other network studies (Borgatti et al., 2006; Costenbader & Valente, 2003; Hoffman, Steinley, Gates, Prinstein, & Brusco, 2018), we use simulation to understand the impact of variable missingness on the validity of centrality scores computed from an empirical psychological network. Specifically, we simulate a ‘true’ psychological network from which we compute ‘true’ centrality scores, and an empirical psychological network from which we compute empirical centrality scores, then compare these two sets of centrality scores. There are two ways to perform such simulations: numeric and analytic.

A *numeric* simulation begins by simulating a traditional case-by-variable dataset intended to represent data on all the beliefs, symptoms, or traits of participants (Figure 1A). From these simulated data, the true psychological network can be estimated (**B**) and the true centrality scores can be computed (**C**). Additionally, by dropping some of the variables to produce a smaller dataset intended to represent the observed data (**D**), an empirical psychological network can be estimated (**E**) and the empirical centrality scores can be computed (**F**). This strategy has the advantage of mirroring the process used in practice for studying centrality in psychological networks; however, it also has several disadvantages. First, there are many ways to estimate a psychological network from data (i.e., **A**→**B**, or **D**→**E**), and the choice of estimation model could impact the simulation results. Second, these estimation models require large numbers of observations to ensure stable estimates, and the number of simulated observations could impact the simulation results. Third, most of these estimation models are computationally intensive, limiting the number of simulations that can be performed. Finally, simulating data from which a network is estimated, rather than simulating the network itself, makes it challenging to experimentally manipulate characteristics of the network to explore how they impact the validity of centrality scores.

An *analytic* simulation overcomes these challenges by exploiting the relationship between zero-order and partial cor-

relations. It begins by simulating a partial correlation matrix intended to represent a true psychological network (**B**), from which centrality scores can be computed (**C**). To obtain an empirical psychological network (i.e., a partial correlation matrix) that would be observed had it been estimated from data containing only a subset of variables (**E**), (1) the partial correlation matrix is transformed into a correlation matrix, then (2) the rows and columns corresponding to the missing variables are dropped, and finally (3) this reduced correlation matrix is transformed back to a partial correlation matrix. Empirical centrality scores (**F**) can then be computed from this new partial correlation matrix representing an empirical psychological network. In the simulations below, we use this analytic approach because it is faster, provides greater control over the characteristics of the simulated networks, and eliminates the influence of a particular network estimation model or the number of observations.³

We simulate the true psychological networks varying their size and structure at two levels to achieve a 2×2 simulation design. The *small* networks containing 20 nodes (i.e., estimated from 20 variables) represent a scales-as-nodes context like that studied by Marcus et al. (2018), while the *large* networks containing 200 nodes (i.e., estimated from 200 variables) represent an items-as-nodes context like that studied by Brandt (2020) and Preszler et al. (2018). Similarly, the *random* networks represent a context in which there is no higher-order structure in the variables’ associations, while the *clustered* networks represent a context in which the variables’ associations are clustered in four domains.

To simulate networks of the desired size, we used the `ggm.simulate.pcor` function in the GeneNet package to generate a random partial correlation matrix, and the `corpcor` package to convert between partial correlation and zero-order correlation matrices (Schaefer, Opgen-Rhein, & Strimmer, 2020; Schafer et al., 2017). To obtain clustered partial correlation matrices for the clustered simulation condition, we assigned 25% of the variables to each of four ‘domains,’ then shuffled the values in this matrix until we achieved a modularity of $0.35 < Q < 0.45$ (Newman & Girvan, 2004). To ensure that all simulated networks resembled psychological networks, we verified that each partial correlation matrix (1) was positive semi-definite, (2) contained more positive than negative values, and (3) did not contain any values equal to exactly zero. Within each of the 4 simulation conditions, we simulated 1000 ‘true’ psychological networks.

From each simulated true psychological network, we used the `centrality_auto` function in the `qgraph` package to compute four centrality measures commonly used in psychological network analysis: strength, closeness, betweenness, and expected influence (Borgatti & Everett, 2006;

³In an earlier version of this paper, available on *PsyArXiv*, we used the numeric approach and obtained similar results.

Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012; L. C. Freeman, 1978; Robinaugh, Millner, & McNally, 2016). We then used the analytic method described above to obtain the empirical psychological network that would be observed if it had been estimated from the same data, but with only a subset of variables and thus containing only a subset of nodes. We consider empirical psychological networks estimated from data in which up to 75% of the variables are missing from the complete data, and thus in which up to 75% of the nodes in the true network are missing. We again used `qgraph` to compute the same four centrality measures from the empirical psychological network. Finally, to measure the validity of these empirical centrality scores, we computed their Spearman correlation with the centrality scores computed from the true network.⁴

Results

Figure 2 shows the simulation results for four types of psychological network: (A) small and random, (B) small and clustered, (C) large and random, and (D) large and clustered. Within each panel, an example simulated true psychological network is shown, with edge color indicating the partial correlation sign (red = negative, blue = positive) and edge thickness indicating the partial correlation magnitude. For simulated true psychological networks that are random, the nodes are all white, while for networks that are clustered, the nodes are colored by their domain membership. Each line in the plots displays, for the respective centrality measure, the mean and 95% confidence interval of the Spearman correlation between centrality scores computed (a) from a true network and (b) from an empirical network estimated from data missing some variables, and thus also missing some nodes.

All four simulation conditions display the same pattern of monotonic decrease in Spearman correlations. This indicates that as more variables are missing from the data used to estimate a empirical psychological network, resulting in an empirical network with more missing nodes, the remaining nodes' empirical centrality scores fail to match their true centrality scores. That is, the empirical centrality scores are invalid because they fail to measure what they are intended to measure. None of the centrality measures perform well, although expected influence (Robinaugh et al., 2016) appears to perform somewhat better than classical graph-theoretic measures (Borgatti & Everett, 2006; L. C. Freeman, 1978).

Given these findings, it is worth asking: *Are these correlations too low?* That is, are centrality scores computed from an incomplete psychological network too invalid to be useful? Of course, one's tolerance for invalidity is necessarily context-dependent. However, consider the proportion of nodes that are likely missing from a typical psychological network. As we show above, it is not uncommon for an empirical psychological network to be missing 50%-99% of the nodes that would be contained in the unobserved but com-

plete network (Brandt, 2020; Marcus et al., 2018; Preszler et al., 2018). When 50% of relevant variables are measures, even under the best circumstances – the true network is large and random, and expected influence is used to measure node centrality – the empirical centrality scores capture less than half the variation in the true centrality scores ($\rho = 0.695$, $\rho^2 = 0.483$). When more variables are unmeasured, or when other centrality measures are used, or when the true network has a different structure, the correspondence between true and empirical centrality scores is worse. Thus, under a broad set of circumstances, we find that when computed from psychological networks, *empirical centrality scores often provide little information* about which beliefs, symptoms, or traits are actually more or less central.

Empirical Example

Methods

The simulations reported in the previous section demonstrate that when variables are missing from the data from which a psychological network is estimated, and thus when the network is missing nodes, the computed centrality scores of the remaining nodes are invalid. However, because the simulations are performed using abstract, synthetic data, they fail to show what this invalidity might look like in practice. For this reason, in this section we provide a concrete empirical example of the practical impact of boundary specification in psychological networks, and specifically how it can lead to erroneous conclusions about node centralities.

Providing an empirical example is challenging for two reasons. First, as we argue above, psychologists often do not have empirical data on all relevant beliefs, symptoms, or traits, and therefore often do not have a complete psychological network. Second, even if empirical data was available that included variables measuring all relevant beliefs, symptoms, or traits, in many cases estimating the network would require estimating so many parameters that the sample size would almost certainly be insufficient (Epskamp, Boorsboom, & Fried, 2018).

To overcome the first issue, we use the example of a political belief network studied by Brandt (2020). Although the original study used 19 variables in the World Values Survey (WVS; Inglehart et al., 2014) to measure Americans' political belief network between 2005 and 2008, the WVS actually contains 69 variables measuring political beliefs. These

⁴We compute the Spearman correlation, rather than the Pearson correlation because, in practice, the substantive aim is usually to identify the relative ordering of traits' or beliefs' centrality, but not their actual centrality scores. Forbes et al. (2019) recommended also measuring the "proportion of [nodes] with the same rank-order for...centrality between networks (p. 7). In all cases, this alternative measure is lower than the Spearman correlation, pointing toward the same conclusion that centrality scores computed from incomplete data are invalid.

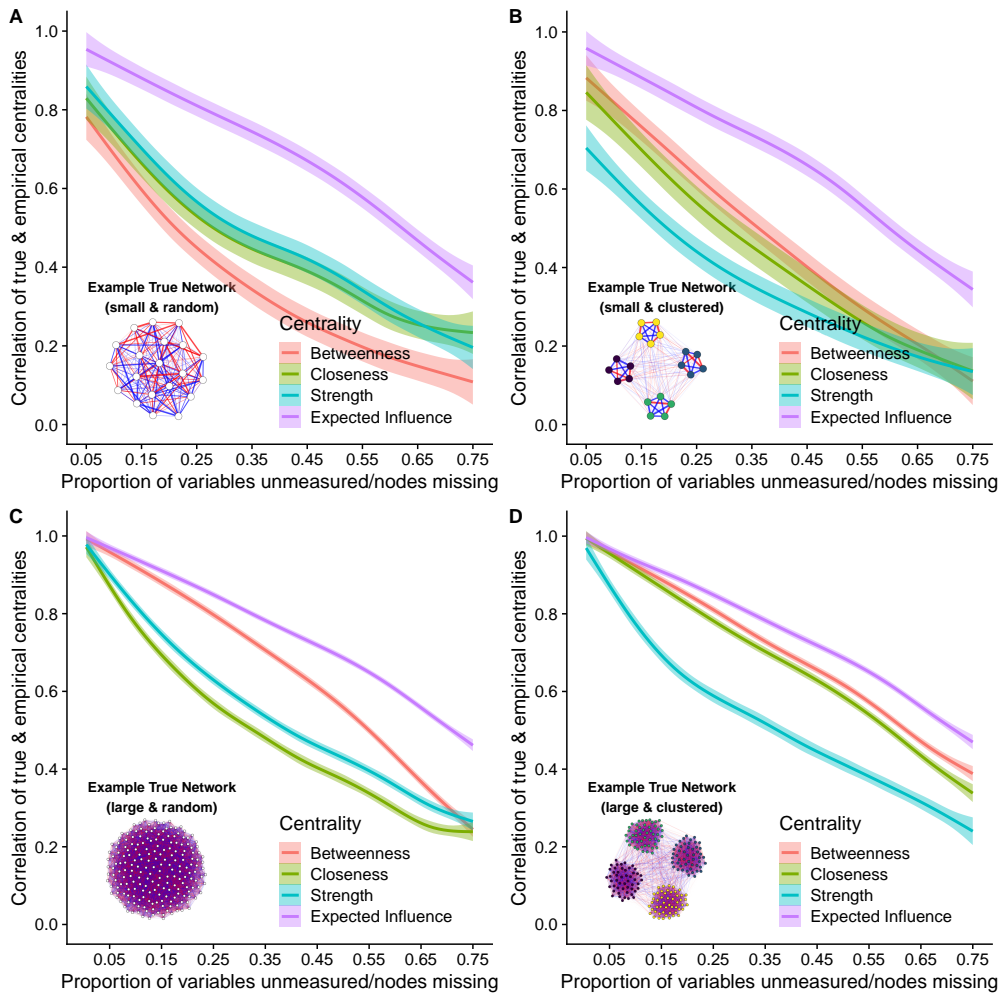


Figure 2

Validity of centrality scores in (A) small and random, (B) small and clustered, (C) large and random, and (D) large and clustered psychological networks. Lines show the mean and 95% CI of the Spearman correlation (y-axis) between centrality scores computed from a true network and an empirical network that is estimated from a subset of variables and thus missing a given proportion of nodes (x-axis). Example true networks show the sign (red = negative, blue = positive) and magnitude of partial correlations among variables.

data therefore permit an illustrative thought experiment: If those 69 variables represented all holdable political beliefs and thus allowed the measurement of a complete political belief system network, how valid are the centralities of beliefs computed in the original study from a network measured using only 19 variables and thus containing only 19 nodes? Of course, there are almost certainly more than 69 holdable political beliefs, but even in this reduced thought experiment it is clear that the original study's data omitted at least 72% of relevant variables and examined networks missing at least 72% of nodes.

To overcome the second issue, which is significant because these data contain only 931 observations, we use the

analytic approach we employed in the simulations above. First, we compute the zero-order correlations among the 69 (in the 'complete' network) or 19 (in the 'observed' network) variables, which is straightforward because $N = 931$ is sufficient for computing a zero-order correlation. Second, in both cases we treat these as correlations as coming from an arbitrarily large sample, and transform them into partial correlations, which we analyze as a psychological network. As in the simulations, this approach does not require the large number of observations necessary for most network estimation approaches, and does not depend on any particular network estimation approach.⁵

⁵In his original analysis, Brandt (2020) estimated the network

To investigate the practical impact of variable and node missingness on centrality scores, we follow the same approach as the simulation. First, we computed four common centrality measures in both the complete 69-node network and the observed 19-node network.⁶ Second, for each measure, we computed the Spearman correlation between the centrality scores of the 19 nodes present in both networks.

Results

Table 1 reports the strength centrality for each node in the 19-node network studied by Brandt (2020), as well as the strength centrality for each of these nodes and three additional example nodes in the larger 69-node network. Computing strength centrality from the 19-node observed network would lead a researcher to conclude, among other things, that beliefs about homosexuality are most central to Americans’ belief system, and that beliefs about euthanasia are the second most central.

However, computing strength centrality from a complete 69-node network reveals that such conclusions about Americans’ most central political beliefs are incorrect. First, beliefs about homosexuality are *not* the most central among these 19 beliefs. Instead, the most central of these 19 beliefs is the belief that experts should rule the country. Second, beliefs about euthanasia are *not* the second most central, and indeed are not particularly central at all, ranking 10th of 19. Third, none of these beliefs are as central to Americans’ belief systems as some other beliefs that were omitted from the observed network entirely, including beliefs about equal rights for women and about the Women’s movement. Taken together, whereas the observed network would suggest Americans’ belief system is dominated by moral beliefs about such issues as homosexuality, euthanasia, and abortion, the complete network suggests that Americans’ belief system is actually dominated by concern for women and the environment.

In this example, measuring a belief system network from incomplete data leads to incorrect conclusions about which beliefs are most central. Taking a wider perspective, the order of strength centrality scores computed from an incomplete network are correlated with centrality scores computed from a complete network at only $\rho = 0.4649$. That is, the apparent ordering of beliefs by strength centrality in the observed network captures only about 23% of the variation in their actual ordering, indicating that the strength centralities derived from incomplete data are invalid. We observe similarly low validity in the other three centrality measures: betweenness ($\rho = 0.1265$, $\rho^2 = 0.02$), closeness ($\rho = 0.5596$, $\rho^2 = 0.31$), and expected influence ($\rho = 0.8824$, $\rho^2 = 0.78$).

Clearly there is variation in these correlations, ranging from very low for betweenness to higher (although still low for establishing criterion validity) for expected influence. However, it is important to note that we regard the 69-node network as ‘complete’ in this example only for the sake of

Table 1

Strength centrality of beliefs in the observed network ($N = 19$) and a hypothetically complete network ($N = 69$).

| Variables/Nodes | Observed | Complete |
|---------------------------------------|----------|----------|
| Homosexuality ^a | 1.528 | 3.131 |
| Euthanasia | 1.403 | 2.620 |
| Strong leader should rule country | 1.397 | 2.403 |
| Abortion | 1.395 | 3.045 |
| Political ideology | 1.312 | 2.816 |
| Army should rule country | 1.250 | 2.693 |
| Private vs. public ownership | 1.237 | 2.421 |
| Experts should rule country | 1.179 | 3.135 |
| Value of competition | 1.111 | 3.009 |
| Income inequality | 1.101 | 2.456 |
| Men are better political leaders | 1.093 | 2.636 |
| Welfare vs. personal responsibility | 1.078 | 2.918 |
| Environmental taxes | 1.057 | 2.579 |
| Economic migration | 1.023 | 2.465 |
| Prostitution | 1.015 | 2.962 |
| Political system should be democratic | 1.003 | 2.441 |
| Men should have priority for jobs | 0.973 | 2.288 |
| Natives should have priority for jobs | 0.806 | 2.142 |
| Environment vs. economy | 0.785 | 2.453 |
| Equal rights for women | — | 3.275 |
| Women’s movement | — | 3.145 |
| Environmental protection movement | — | 3.102 |
| 47 additional variables/nodes | — | NR |
| Spearman ρ | | 0.4649 |

^a The WVS asked “whether you think [“homosexuality”] can always be justified, never be justified, or something in between.” We retain the original language here, but note that the Gay and Lesbian Alliance Against Defamation (GLAAD) now recommends using using alternative language (e.g., “gay and lesbian lives”).

illustration. We do not actually believe that the centrality scores computed from it are valid either, because even this larger network is almost certainly missing many holdable political beliefs. If we were able to measure a complete network of all holdable political beliefs, our simulation results suggest that the centrality scores computed from our 69-node network would be a poor approximation, and the centrality scores computed from the 19-node network would be even worse.

using a Bayesian Gaussian graphical model (BGGM). In an earlier version of this paper, available on *PsyArXiv*, we reported this empirical example using BGGM and obtained similar results.

⁶For comparability with the simulation results, we report the same four centrality measures: betweenness, closeness, strength, and expected influence. In his original study, Brandt (2020) did not compute expected influence, but did compute eigenvector, 1-step expected influence, and 2-step expected influence. We observed the same patterns in these three additional centrality measures.

Discussion

Psychological networks – where nodes are symptoms, traits, or beliefs and where edges capture their statistical associations – are now commonly studied in many subfields of psychology, with a particular emphasis on identifying the most central (i.e., important) nodes. However, the boundaries of these networks, which define the universe of relevant variables from which these networks should be estimated, and thus also the nodes that should be included in the network, are often unknown or uncountably large. This leads many empirical psychological networks to be poorly bounded, estimated from data in which key variables are unmeasured and missing some and often a very large number of nodes. By simulating ‘true’ psychological networks, we have shown that centrality scores computed from networks estimated from incomplete data are invalid. The magnitude of invalidity depends primarily on the proportion of variables that are unmeasured and nodes that are missing. We illustrate the practical impact of invalid centrality scores in the context of Americans’ political belief network, demonstrating that conclusions drawn from incomplete data about Americans’ central political beliefs are incorrect.

There are three related causes for invalid centrality scores in psychological networks. First, as has previously been demonstrated in social networks (e.g., Borgatti et al., 2006; Costenbader & Valente, 2003), the omission of a node changes the topology of a network, thereby changing other nodes’ centrality scores. Second, psychological networks are not measured directly, but are estimated from non-network data using statistical models (e.g., Gaussian graphical models) in which the weight of the edge connecting two nodes is their statistical association (e.g., their partial correlation). This estimation process means that when a variable (e.g., a belief item) is unmeasured in the raw data, not only is the corresponding node missing from the estimated network, but in addition the edge weights among the remaining nodes may also change. Third, the boundaries of psychological networks are often ambiguous or unknown, leading the proportion of variables that are unmeasured and nodes that are missing to sometimes be very large. For example, a complete political belief network would contain a node representing every holdable political belief, the number of which is unknown but surely very large. However, in practice, an empirical network contains only nodes representing political beliefs that happen to be measured and appear in the data.

In response to these challenges to the validity of centrality scores in psychological networks, researchers might be tempted to frame their conclusions more tentatively as conditional on their empirical network’s boundary. For example, in the empirical example we present in Table 1, one might be tempted to say “In a belief network of these 19 items, beliefs about homosexuality were the most central.” Strictly speaking, this is not incorrect and might even repli-

cate in subsequent studies using different samples (Brandt, 2020; Williams, Rast, Pericchi, & Mulder, 2020), but it is also not particularly informative. As we show, although beliefs about homosexuality appear to be the most central belief in Americans’ belief system, this is an artifact of the empirical network’s boundary (i.e., which beliefs happened to be measured) and does not reflect the true nature of Americans’ belief system.

A second possible response may be to ask: *what proportion of variables can be unmeasured, and what proportion of nodes can be missing, while still achieving reasonably valid centrality scores?* Answers to this question are of limited practical value in dealing with unmeasured variables and missing nodes for two reasons. First, in practice, researchers rarely know how large the universe of beliefs or traits is, and therefore rarely know what proportion of nodes are missing from their empirical network. Second, when the universe of beliefs or traits is large (e.g., greater than 200, as is likely the case for belief or trait items), our simulation results illustrate that centrality indices are invalid even when the empirical network includes 100 nodes. Therefore, even if the necessary proportion of nodes were known, our simulations suggest that when the universe is large, the required number of nodes would be impractically large for even the most ambitious attempts at empirical data collection and estimating a psychological network from such data would require an impractically large sample.

To summarize, when centrality is computed in a psychological network that does not include (nearly) all the nodes within a theoretically meaningful boundary, the centrality scores computed from the network are invalid. That is, they do not measure the nodes’ true centrality. Therefore, in addition to the list of recommendations Bringmann et al. (2019) offered, we offer an additional and overarching recommendation: *Centrality should be computed and interpreted only in psychological networks that measure (nearly) all the variables, and thus include (nearly) all the nodes, inside a theoretically meaningful boundary.* In this analysis, we have focused narrowly on the implications of boundary specification for centrality because it is among the most commonly used network analytic measures. However, in view of the social network literature on problems associated with boundary specification and missingness (e.g., Butts, 2009; Kossinets, 2006; Laumann et al., 1992; Smith & Moody, 2013; Smith et al., 2017), future research should consider whether a similar recommendation should apply to other forms of analysis on psychological networks such as the detection of communities or small worldness (e.g., Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Mullen & Jones, 2021).

We conclude by returning to some open questions in the literature. Brandt et al. (2019) asked “What is central in political belief system networks?” Our simulations and empirical example offer an answer: Unless a belief system net-

work includes a node for every holdable belief, *we simply cannot know which beliefs are actually central*. Relatedly, [Bringmann et al. \(2019\)](#) asked “What do centrality measures measure in psychological networks?” Again, our simulations and empirical example offer an answer: When a psychological network is estimated from data that did not measure all the variables in a theoretically-defined universe, *the centrality scores often do not measure anything meaningful* about the true centrality of the remaining variables in that universe.

Given these answers, do networks have a role in psychology? We believe they do. First, psychological networks like those we discuss here may still be informative in contexts where a theoretically-bounded complete network possible, for example, when the universe of diagnostic criteria for the universe of psychopathologies is known and measurable ([Tio, Epskamp, Noordhof, & Borsboom, 2016](#)). Second, because there is “no clear boundary where network models end and latent variable (or common cause) models begin” ([Bringmann & Eronen, 2018](#), p. 607), developments from one modeling approach may inform the other (e.g., [Chandrasekaran, Parrilo, & Willsky, 2010](#); [Christensen & Golino, 2021](#); [Epskamp, Rhemtulla, & Borsboom, 2017](#)). Third, inferential network models may offer opportunities to expand the scope of existing methods of dyadic data analysis (e.g., [Kenny, Kashy, & Cook, 2006](#)) to accommodate triadic and higher-order groups of interdependent actors. Finally, as we noted at the outset, psychologists can stake some claim as being among the earliest social network analysts, and psychology remains critical for understanding social networks in which the nodes are people and edges are social or affective relations.

References

- Beard, C., Millner, A. J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M., ... Björngvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, *46*(16), 3359–3369. <https://doi.org/10.1017/S0033291716002300>
- Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, *28*(2), 124–136. <https://doi.org/10.1016/j.socnet.2005.05.001>
- Borgatti, S. P., & Everett, M. G. (2006). A graph-theoretic perspective on centrality. *Social Networks*, *28*(4), 466–484. <https://doi.org/10.1016/j.socnet.2005.11.005>
- 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., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS One*, *6*(11), e27407. <https://doi.org/10.1371/journal.pone.0027407>
- Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L., & Cramer, A. O. (2017). False alarm? A comprehensive reanalysis of “evidence that psychopathology symptom networks have limited replicability” by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, *126*(7), 989–999. <https://doi.org/10.1037/abn0000306>
- Brandt, M. J. (2020). Estimating and examining the replicability of belief system networks. *Collabra: Psychology*, *6*(1), 24. <https://doi.org/10.1525/collabra.312>
- Brandt, M. J., Sibley, C. G., & Osborne, D. (2019). What is central to political belief system networks? *Personality and Social Psychology Bulletin*, *45*(9), 1352–1364. <https://doi.org/10.1177/0146167218824354>
- Breiger, R. L. (1974). The duality of persons and groups. *Social Forces*, *53*(2), 181–190. <https://doi.org/10.1093/sf/53.2.181>
- Briganti, G., Kempenaers, C., Braun, S., Fried, E. I., & Linkowski, P. (2018). Network analysis of empathy items from the interpersonal reactivity index in 1973 young adults. *Psychiatry Research*, *265*, 87–92. <https://doi.org/10.1016/j.psychres.2018.03.082>
- Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., ... Snippe, E. (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*(4), 606. <https://doi.org/10.1037/rev0000108>
- Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science*, *325*(5939), 414–416. <https://doi.org/10.1126/science.1171022>
- Chandrasekaran, V., Parrilo, P. A., & Willsky, A. S. (2010). Latent variable graphical model selection via convex optimization. In *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (pp. 1610–1613).
- Christensen, A. P., & Golino, H. (2021). On the equivalency of factor and network loadings. *Behavior Research Methods*, 1–18. <https://doi.org/10.3758/s13428-020-01500-6>
- Christensen, A. P., Golino, H., & Silvia, P. J. (2020). A psychometric network perspective on the validity and validation of personality trait questionnaires. *European Journal of Personality*. <https://doi.org/10.1002/per.2265>
- Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, *136*, 68–78. <https://doi.org/10.1016/j.paid.2017.06.011>
- Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks*, *25*(4), 283–307. [https://doi.org/10.1016/S0378-8733\(03\)00012-1](https://doi.org/10.1016/S0378-8733(03)00012-1)
- Cramer, A. O., Waldorp, L. J., Van Der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–150. <https://doi.org/10.1017/S0140525X09991567>
- Curry, O. S., Alfano, M., Brandt, M., & Pelican, C. (2020). Moral molecules: Morality as a combinatorial system. *OSF*

Preprints.

- Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. (2016). Toward a formalized account of attitudes: The causal attitude network (can) model. *Psychological Review*, *123*(1), 2. <https://doi.org/10.1037/a0039802>
- Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. (2019). A network perspective on attitude strength: Testing the connectivity hypothesis. *Social Psychological and Personality Science*, *10*(6), 746–756. <https://doi.org/10.1177/1948550618781062>
- 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., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18. Retrieved from <http://www.jstatsoft.org/v48/i04/>
- Epskamp, S., Fried, E. I., van Borkulo, C. D., Robinaugh, D. J., Marsman, M., Dalege, J., ... Cramer, A. O. (2018). Investigating the utility of fixed-margin sampling in network psychometrics. *Multivariate Behavioral Research*, 1–15. <https://doi.org/10.1080/00273171.2018.1489771>
- Epskamp, S., Rhemtulla, M., & 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>
- Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups; A study of human factors in housing*. Harper.
- Forbes, M. K., Wright, A. G., Markon, K. E., & Krueger, R. F. (2017a). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology*, *126*(7), 969. <https://doi.org/10.1037/abn0000276>
- Forbes, M. K., Wright, A. G., Markon, K. E., & Krueger, R. F. (2017b). Further evidence that psychopathology networks have limited replicability and utility: Response to Borsboom et al. (2017) and Steinley et al. (2017). *Journal of Abnormal Psychology*, *126*(7), 1011–1016. <https://doi.org/10.1037/abn0000313>
- Forbes, M. K., Wright, A. G., Markon, K. E., & Krueger, R. F. (2021). On unreplicable inferences in psychopathology symptom networks and the importance of unreliable parameter estimates. *Multivariate Behavioral Research*, 1–9.
- Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2019). Quantifying the reliability and replicability of psychopathology network characteristics. *Multivariate Behavioral Research*, 1–19. <https://doi.org/10.1080/00273171.2019.1616526>
- Freeman, L. (2004). *The development of social network analysis: A study in the sociology of science*. Vancouver, BC: Empirical Press.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Fried, E. I., van Borkulo, C. D., & Epskamp, S. (2020). On the importance of estimating parameter uncertainty in network psychometrics: A response to Forbes et al. (2019). *Multivariate Behavioral Research*, 1–6. <https://doi.org/10.1080/00273171.2020.1746903>
- Galaskiewicz, J. (1991). Estimating point centrality using different network sampling techniques. *Social Networks*, *13*(4), 347–386. [https://doi.org/10.1016/0378-8733\(91\)90002-B](https://doi.org/10.1016/0378-8733(91)90002-B)
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. <https://doi.org/10.1016/j.jrp.2005.08.007>
- Hallquist, M. N., Wright, A. G., & Molenaar, P. C. (2019). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research*, 1–25. <https://doi.org/10.1080/00273171.2019.1640103>
- Hoffman, M., Steinley, D., Gates, K. M., Prinstein, M. J., & Brusco, M. J. (2018). Detecting clusters/communities in social networks. *Multivariate Behavioral Research*, *53*(1), 57–73. <https://doi.org/10.1080/00273171.2017.1391682>
- Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., ... et al. (2014). World values survey: All rounds - country-pooled datafile 1981–2014. *JD Systems Institute*. Retrieved from <http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp>
- Jones, P. J., Williams, D. R., & McNally, R. J. (2020). Sampling variability is not nonreplication: A bayesian reanalysis of Forbes, Wright, Markon, and Krueger. *Multivariate Behavioral Research*, 1–7. <https://doi.org/10.1080/00273171.2020.1797460>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Guilford press.
- Kossinets, G. (2006). Effects of missing data in social networks. *Social Networks*, *28*(3), 247–268. <https://doi.org/10.1016/j.socnet.2005.07.002>
- Latapy, M., Magnien, C., & Del Vecchio, N. (2008). Basic notions for the analysis of large two-mode networks. *Social Networks*, *30*(1), 31–48. <https://doi.org/10.1016/j.socnet.2007.04.006>
- Laumann, E. O., Marsden, P. V., & Prensky, D. (1992). The boundary specification problem in network analysis. In L. C. Freeman, D. R. White, & A. K. Romney (Eds.), *Research methods in social network analysis* (pp. 61–87).
- Marcus, D. K., Preszler, J., & Zeigler-Hill, V. (2018). A network of dark personality traits: What lies at the heart of darkness? *Journal of Research in Personality*, *73*, 56–62. <https://doi.org/10.1016/j.jrp.2017.11.003>
- Moreno, J. L., & Jennings, H. H. (1934). *Who shall survive?: A new approach to the problem of human interrelations*. Nervous and Mental Disease Publishing Co.
- Mullen, R., & Jones, E. S. (2021). Network analysis of competitive state anxiety. *Frontiers in Psychology*, *11*, 3726. <https://doi.org/10.3389/fpsyg.2020.586976>
- Neal, Z. (2014). The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. *Social Networks*, *39*, 84–97. <https://doi.org/10.1016/j.socnet.2014.06.001>

- Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113. <https://doi.org/10.1103/PhysRevE.69.026113>
- Preszler, J., Marcus, D. K., Edens, J. F., & McDermott, B. E. (2018). Network analysis of psychopathy in forensic patients. *Journal of Abnormal Psychology*, *127*(2), 171. <https://doi.org/10.1037/abn0000325>
- Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of abnormal psychology*, *125*(6), 747. <https://doi.org/10.1037/abn0000181>
- Schaefer, J., Opgen-Rhein, R., & Strimmer, K. (2020). GeneNet: Modeling and inferring gene networks [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=GeneNet> (R package version 1.2.15)
- Schafer, J., Opgen-Rhein, R., Zuber, V., Ahdesmaki, M., Silva, A. P. D., & Strimmer, K. (2017). *corpcor: Efficient estimation of covariance and (partial) correlation*. CRAN.
- Schlicht-Schmälzle, R., Chykina, V., & Schmälzle, R. (2018, 12). An attitude network analysis of post-national citizenship identities. *PloS One*, *13*(12), 1-19. <https://doi.org/10.1371/journal.pone.0208241>
- Smith, J. A., & Moody, J. (2013). Structural effects of network sampling coverage i: Nodes missing at random. *Social Networks*, *35*(4), 652–668. <https://doi.org/10.1016/j.socnet.2013.09.003>
- Smith, J. A., Moody, J., & Morgan, J. H. (2017). Network sampling coverage ii: The effect of non-random missing data on network measurement. *Social Networks*, *48*, 78–99. <https://doi.org/10.1016/j.socnet.2016.04.005>
- Steinley, D., Hoffman, M., Brusco, M. J., & Sher, K. J. (2017). A method for making inferences in network analysis: Comment on Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, *126*(7), 1000-1010. <https://doi.org/10.1037/abn0000308>
- Tio, P., Epskamp, S., Noordhof, A., & Borsboom, D. (2016). Mapping the manuals of madness: Comparing the icd-10 and dsm-iv-tr using a network approach. *International Journal of Methods in Psychiatric Research*, *25*(4), 267–276. <https://doi.org/10.1002/mpr.1503>
- Travers, J., & Milgram, S. (1977). An experimental study of the small world problem. In *Social networks* (pp. 179–197). Elsevier. <https://doi.org/10.1016/B978-0-12-442450-0.50018-3>
- 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*. <https://doi.org/10.1037/met0000254>