Critiques of network analysis of multivariate data in psychological science

Zachary P. Neal*[1], Miriam K. Forbes[2], Jennifer Watling Neal[1], Michael J. Brusco[3], Robert Krueger[4], Kristian Markon[5], Douglas Steinley[6], Stanley Wasserman[7], and Aidan G. C. Wright[8]

[1]Department of Psychology, Michigan State University, USA
[2]School of Psychological Sciences, Macquarie University, AUSTRALIA
[3]College of Business, Florida State University, USA
[4]Psychology Department, University of Minnesota, USA
[5]Department of Psychological and Brain Sciences, University of Iowa, USA
[6]Department of Psychological Sciences, University of Missouri, USA
[7]Department of Psychological and Brain Sciences, Indiana University, USA
[8]Department of Psychology, University of Pittsburgh, USA

Abstract

Using network methods to analyze multivariate psychological data on symptoms or beliefs has become popular, and numerous tutorials and primers provide detailed guides. However, these guides make such methods appear artificially robust because they often fail to adequately review serious critiques of psychometric network analysis. We briefly review critiques concerning model selection, study design, estimation reliability, and interpretation of measures, and provide a comprehensive annotated bibliography of published critiques.

This is a preprint. The final published paper is available at and should be cited as:
Neal, Z. P., Forbes, M. K. Neal, J. W., Brusco, M. J., Krueger, R., Markon, K., Steinley, D., Wasserman, S., and Wright, A. G. C. (2022). Critiques of network analysis of multivariate data in psychological science. *Nature Reviews Method Primers.*

A recent primer on the network analysis of multivariate data in psychological science provided an overview of 'psychometric network analysis,' including graphical models, estimation methods for those models, and descriptive tools[1]. These techniques are used for obtaining and examining the statistical associations among psychological variables as a network. Unfortunately, the primer systematically omitted or glossed over core critiques of these models and methods, making them appear artificially robust and widely-accepted. In this comment, we highlight four categories of critique: (1) selecting network models when better-suited multivariate methods already exist, (2) adopting study designs that are mismatched to research questions, (3) estimating networks using methods that yield unreliable estimates, and (4) interpreting network metrics that are invalid when applied to networks of statistical associations. We point readers to selected critiques that were available when the original primer was written, and provide a complete annotated bibliography of critiques of psychometric network analysis at https://doi.org/10.31234/osf.io/ke6qn.

1 Model selection

The goal of psychometric network analysis is to understand the relationships among measured psychological variables, including symptoms, beliefs, or traits. However, many multivariate methods already exist for this purpose. For example, latent variable models such as factor analysis and structural equation methods already allow researchers to examine the relationships among a set of measured variables, and offer robust ways to evaluate model fit. A series of debates in both clinical and personality psychology have challenged whether psychometric network analysis is the most appropriate method for research questions focused on understanding patterns of variable associations. Some have observed that network models may be identical in many cases to latent variable models; however, because of their structure and number of parameters network models may overfit and be lacking in parsimony[2]. Similarly, others note that network models do not provide value beyond a graphical depiction of the correlations or partial correlations among the variables[3]. Following these observations, *we recommend that researchers conduct psychometric network analysis only when it is better matched to their research questions than other existing multivariate analytic methods.*

2 Study design

The majority of psychometric network analyses are conducted on between-person and cross-sectional data. This design ensures that psychometric networks cannot "offer clues to causal dynamics" among symptoms or beliefs in individuals. Despite the primer's claims that "network models...explicitly represent pairwise interactions between components in a system," they only explicitly represent pairwise statistical associations, and therefore do not "form a natural bridge from data analysis to theory formation"[1]. First, it is not possible to use between-person statistical associations among symptoms or beliefs to draw conclusions about how those symptoms or beliefs are related for a given individual, even a member of the sample. To do so is an ecological fallacy, and "the visually attractive network graphs enhance this tendency" to commit such fallacies[4]. Second, it is not possible to use cross-sectional data to make causal inferences because correlations among variables – even regularized partial correlations estimated using sophisticated techniques – do not indicate causality. Indeed, comparisons to inferences drawn from longitudinal data reveal that cross-sectional psychometric networks "do not reflect how symptoms trigger each other over time"[5]. Therefore, *we recommend that researchers refrain from making inferences about causality or within-person processes when using between-person cross-sectional data.*

3 Estimation reliability

The strength of connections between symptoms or beliefs in psychometric networks are operationalized using measures of statistical association estimated from empirical data on a sample of subjects. Identifying which of these connections are robust and reliable is critical because they define the structure of the network. However, the estimation of psychometric networks from multivariate data typical in psychological science often yields unreliable estimates. For example, the fitting of multiple types of conditional independence models, when applied to multiple epidemiological samples, yielded unreliable estimates of individual connections and nodes' centralities[6]. Indeed, compared to Monte Carlo fixed-marginal models common in network science, psychometric networks models yielded "results...indistinguishable from what would be expected by chance"[7]. More broadly, although psychometric networks may appear similar when "relying on global summary statistics," the detailed characteristics of the networks interpreted in the literature typically vary substantially with even small differences in study samples and methods[8]. Because the robustness of estimated psychometric networks remains questionable, *we recommend that researchers exercise caution in drawing inferences for clinical research or practice before results have been rigorously replicated.*

4 Interpretation of measures

The symptoms or beliefs in psychometric networks are often characterized by their centrality in the network. However, most network centrality metrics were developed for networks that are both bounded and composed of social actors, and lose their interpretability when applied to unbounded networks of variables. The primer briefly noted that "if important nodes are missing, this can affect the structure of the network" and that "there are no clear guidelines" for interpreting centrality in psychometric networks[1], but this glosses over the severity of the problem. First, it is often impossible for psychometric networks to include all important nodes because the universe of relevant symptoms or beliefs is unknown and uncountably large. Exacerbating this issue, the primer recommends that researchers decide which nodes to include in the network *after* data collection, rather than *a priori* to inform data collection. Simulations have revealed that these practices render centrality measures invalid in psychometric networks[9]. Second, most centrality metrics are designed to measure the flow of information or influence among distinct and exchangeable nodes. In contrast, psychometric networks measure statistical associations among correlated variables representing symptoms and beliefs that vary in severity or intensity. Therefore, hypotheses that link node centrality to symptom or belief importance are 'ill-defined'[10]. Because psychometric networks typically have unknown levels of node missingness, and in the absence of an explicit conceptual rationale for their applicability, *we recommend that researchers consider more appropriate alternatives to centrality metrics when analyzing psychometric networks.*

5 Conclusion

Network methods have a long history and many potential applications in psychology. However, these critiques highlight that there are important but under-recognized limitations in psychometric network analyses of variables measuring symptoms, beliefs, or traits.

Contributor Statement. ZPN conceptualized the paper. ZPN, MKF, and JWN drafted the proposal. All authors drafted and reviewed the final paper.

References

[1] Denny Borsboom, Marie K Deserno, Mijke Rhemtulla, Sacha Epskamp, Eiko I Fried, Richard J McNally, Donald J Robinaugh, Marco Perugini, Jonas Dalege, Giulio Costantini, et al. Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, 1:58, 2021. https://doi.org/10.1038/s43586-021-00055-w.

[2] Robert F. Krueger, Colin G. DeYoung, and Kristian E. Markon. Toward scientifically useful quantitative models of psychopathology: The importance of a comparative approach. *Behavioral and Brain Sciences*, 33(2-3):163–164, 2010. https://doi.org/10.1017/S0140525X10000658.

[3] Jens B Asendorpf. What do the items and their associations refer to in a network approach to personality? *European Journal of Personality*, 26(4):432–433, 2012. https://doi.org/10.1002/per.1867.

[4] Elisabeth H Bos and Rob BK Wanders. Group-level symptom networks in depression. *JAMA psychiatry*, 73(4):411–411, 2016. https://doi.org/10.1001/jamapsychiatry.2015.3103.

[5] Fionneke M Bos, Evelien Snippe, Stijn de Vos, Jessica A Hartmann, Claudia JP Simons, Lian van der Krieke, Peter de Jonge, and Marieke Wichers. Can we jump from cross-sectional to dynamic interpretations of networks implications for the network perspective in psychiatry. *Psychotherapy and psychosomatics*, 86(3):175–177, 2017. https://doi.org/10.1159/000453583.

[6] Miriam K Forbes, Aidan GC Wright, Kristian E Markon, and Robert F Krueger. Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology*, 126(7):969—-988, 2017. https://doi.org/10.1037/abn0000276.

[7] Douglas Steinley, Michaela Hoffman, Michael J Brusco, and Kenneth J Sher. A method for making inferences in network analysis: Comment on Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, 126:1000–1010, 2017. https://doi.org/10.1037/abn0000308.

[8] Miriam K Forbes, Aidan GC Wright, Kristian E Markon, and Robert F Krueger. Quantifying the reliability and replicability of psychopathology network characteristics. *Multivariate Behavioral Research*, 56:224–242, 2021. https://doi.org/10.1080/00273171.2019.1616526.

[9] Zachary P Neal and Jennifer Watling Neal. Out of bounds? the boundary specification problem for centrality in psychological networks. *Psychological Methods*, 2021. https://doi.org/10.1037/met0000426.

[10] Tobias R Spiller, Ofir Levi, Yuval Neria, Benjamin Suarez-Jimenez, Yair Bar-Haim, and Amit Lazarov. On the validity of the centrality hypothesis in cross-sectional between-subject networks of psychopathology. *BMC Medicine*, 18 (1):297, 2020. https://doi.org/10.1186/s12916-020-01740-5.