



<https://openaccess.leidenuniv.nl>

License: Article 25fa pilot End User Agreement

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication other than authorised under this licence or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: OpenAccess@library.leidenuniv.nl

Article details

Borsboom D., Robinaugh D.J., The Psychosystems Group (including Fried E.I.), Rhemtulla M. & Cramer A.O.J. (2018), Robustness and replicability of psychopathology networks, *World Psychiatry* 17(2): 143-144.
Doi: 10.1002/wps.20515

Robustness and replicability of psychopathology networks

Network approaches to psychopathology hold that mental disorders arise from the interplay between symptoms in a network structure^{1,2}. In the past few years, statistical techniques that estimate networks were developed and applied to many disorders³. As empirical findings start to accumulate, the question arising is which of these findings are robust and replicable. Here we evaluate the state of psychopathological network research based on three methodological criteria: model quality, precision, and replicability.

Model quality. One important quality of statistical modeling techniques is their ability to recover the “true” model that generated the data, a necessary prerequisite for justifying inferences based on network models. This is evaluated through: a) mathematical analysis: prove that a technique will recover the generating network model from data (e.g., by showing that it converges to the true model as sample size increases); and b) simulation studies: evaluate a technique’s performance under various circumstances (e.g., for different network structures, sample sizes, and parameter settings).

Current state-of-the-art network techniques (i.e., pairwise Markov random fields⁴) have been vetted through mathematical proofs and simulation studies^{5,6}: they efficiently recover the “true” model underlying the data. In general, such techniques minimize the false positive rate at the expense of statistical power. As a result, these techniques are more likely to omit “true” network connections, than to include spurious connections^{5,6}. In sum, these techniques are vetted, conservative tools for estimating psychopathology network structures.

Precision and robustness. When a researcher has estimated a network from empirical data using vetted methodology, the question is to what extent the parameter estimates are precise: how robust are the results? For instance, if the relationship between self-worth and suicidal thoughts seems stronger than that between sleep difficulty and suicidal thoughts, it is necessary to investigate if model parameters are estimated with sufficient precision to justify this inference. If not, the result may not replicate in other samples.

Precision of network parameter estimates can vary considerably depending on factors such as sample size, network size, and network structure. Therefore, these factors must be assessed and reported on a case-by-case basis, by evaluating the statistical precision of parameter estimates (e.g., with confidence intervals) and the robustness of the model as a whole (e.g., investigating network structures in subsamples).

Dedicated freeware methodology for doing this recently became available⁴, which allows researchers to report confidence intervals for estimated network parameters as an integral part of their results. This practice was quickly embraced by the majority of the network community, who now publish their work including detailed robustness checks. Naturally, results of such analyses should constrain the researcher’s conclusions proportionately to their content: stronger claims (e.g., “insomnia

is the most central node in the depression network”) require stronger evidence than weaker claims (e.g., “insomnia is connected to the depression network”).

Replicability. When network analysis seems to warrant an empirical conclusion (e.g., a particular symptom is highly central, or one network is more densely connected than another), the next question is whether the relevant phenomenon can be replicated in other samples. Ideally, replication research differs from the original study only in features that are deemed irrelevant to the phenomenon under investigation (e.g., by using a different sample from the same population). However, as is often the case in replication research, it is sometimes unclear whether differences between a study and its purported replication are relevant or not. For instance, if a network is first estimated on a community sample, and a replication is attempted in a patient sample, it may be unrealistic to assume that the same network holds in both populations. In such cases, studies probe not only a finding’s replicability, but also its generalizability. Consequently, if inconsistent findings arise, this may either be because the phenomenon is unstable or illusory (i.e., the finding is not replicable) or because of substantively meaningful differences between studies (i.e., the finding is not generalizable to the context of the new study). In contrast, if an empirical phenomenon is observed consistently across studies, this provides compound evidence for both its replicability and generalizability⁷.

Several recent empirical studies have evaluated the replicability of networks. The general picture which emerges is that network structures replicate and generalize well. For example, networks of major depression and generalized anxiety disorder symptoms are nearly identical in the US and Australia; post-traumatic stress disorder (PTSD) networks are similar across different populations and sources of trauma; and major depression networks are invariant across environmental and genetic risk factors (e.g., age of onset)^{7,8}.

Although network structures appear replicable and generalizable, detailed inferences based on them may be more susceptible to variation across studies. For example, the centrality of nodes seems to vary across PTSD networks, and a reported difference in network density between remitted and persistent major depression cases in adults was not fully replicated in an adolescent sample⁸. Future research should critically interrogate such findings to determine if inconsistency between studies is best characterized as a failure to replicate or a failure to generalize across contexts.

In conclusion, the model quality of network analysis techniques is good, while precision and robustness can now adequately be assessed with freely available methodological tools. Burgeoning replication research suggests that the structure of networks is typically consistent across studies, while stronger inferences based on these structures (e.g., centrality) have occasionally yielded mixed results.

Network analysis is a promising approach that may lead to significant improvements in research on and treatment of psychopathology⁹, but researchers should be careful not to overstate causal conclusions based on network analysis as long as the causal interpretation of models has not been thoroughly investigated. The assessment of network robustness and replicability is an important step in this process and should be an important research focus in the next few years.

Denny Borsboom¹, Donald J. Robinaugh¹, The Psychosystems Group¹, Mijke Rhemtulla², Angélique O.J. Cramer³

¹Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands; ²Department of Psychology, University of California, Davis, CA, USA; ³Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands

The Psychosystems Group includes Sacha Epskamp, Eiko I. Fried, Lourens J. Waldorp, Claudia D. van Borkulo, and Han L.J. van der Maas.

1. Borsboom D. *World Psychiatry* 2017;16:5-13.
2. Cramer AOJ, van Borkulo CD, Giltaij EJ et al. *PLoS One* 2016;11:e0167490.
3. Fried EI, van Borkulo CD, Cramer AOJ et al. *Soc Psychiatry Psychiatr Epidemiol* 2017;52:1-10.
4. Epskamp S, Borsboom D, Fried EI. *Behav Res Methods* 2018;50:195-212.
5. Bühlmann P, van de Geer S. *Statistics for high-dimensional data: methods, theory and applications*. Berlin: Springer, 2011.
6. Van Borkulo CD, Borsboom D, Epskamp S et al. *Sci Rep* 2014;4:5918.
7. Borsboom D, Fried EI, Epskamp S et al. *J Abnorm Psychol* 2017;126:989-99.
8. Fried EI. <http://psych-networks.com/7-new-papers-network-replicability/>.
9. McNally RJ. *Behav Res Ther* 2016;86:95-104.

DOI:10.1002/wps.20515