Comparison of Latent Variable and Psychological Network Models in PROMIS Data: Output

Metrics and Factor Structure

Joshua Starr and Carl F. Falk

Department of Psychology

McGill University

Acknowledgements: The authors would like to thank Mijke Rhemtulla for her helpful comments on a previous version of this manuscript and Felix Fischer for discussions regarding PROMIS.

Correspondence: Carl F. Falk, Department of Psychology, 2001 McGill College Ave., 7th Floor, Montreal, QC H3A 1G1, Canada. Email: carl.falk@mcgill.ca.

Abstract

Purpose: Much research is still needed to compare traditional latent variable models such as confirmatory factor analysis (CFA) to emerging psychometric models such as the Gaussian graphical model (GGM). Previous comparisons of GGM centrality indices with factor loadings from CFA have discovered redundancies, and investigations into how well a GGM-based alternative to exploratory factor analysis (i.e., exploratory graph analysis, or EGA) is able to recover the hypothesized factor structure show mixed results. Importantly, such comparisons have not typically been examined in real mental and physical health symptom data, despite such data being an excellent candidate for the GGM. Our goal was to extend previous work by comparing the GGM and CFA using data from Wave 1 of the Patient Reported Outcomes Measurement Information System (PROMIS). *Methods:* Models were fit to PROMIS data based on 16 test forms designed to measure 9 mental and physical health domains. Our analyses borrowed a two-stage approach for handling missing data from the structural equation modeling literature. *Results:* We found weaker correspondence between centrality indices and factor loadings than found by previous research, but in a similar pattern of correspondence. EGA recommended a factor structure discrepant with PROMIS domains in most cases yet may be taken to provide substantive insight into the dimensionality of PROMIS domains. *Conclusion:* In real mental and physical health data, GGM and EGA may provide complementary information to traditional CFA metrics.

Keywords: PROMIS, network analysis, centrality indices, exploratory graph analysis, mental and physical health, measurement

Most mental health researchers hold that psychological disorders must be measured indirectly through observable symptoms. Yet, competing frameworks exist for explaining how such symptoms represent the mental disorders they purport to measure. Latent variable models have been considered the “gold-standard” in psychopathology measurement [1]. In this framework, a mental disorder is an underlying common cause of one or more manifest symptoms. In contrast, recently developed network models conceptualize mental illness as a system of symptoms, linked by their direct and indirect associations [2]. From this perspective, symptoms are not mere reflections of mental disorders but comprise mental disorders at their most basic level [3].

While great emphasis has been placed on core conceptual differences between network and latent variable models, some researchers have illuminated statistical similarities and model equivalencies [4, 5]. One may then ask, what novel psychometric information does network modeling offer that latent variable modeling does not? What complementary insight is offered by each framework? In this manuscript, we compare network and factor analytic models by addressing two issues: 1) the redundancy of factor loadings in confirmatory factor analysis with centrality indices in network models [6]; and 2) the utility of conclusions from a network modeling alternative to exploratory factor analysis called exploratory graph analysis [7]. We consider both issues jointly as the core shortcoming in prior research is the lack of consideration of real symptom data. As context, we hope to gain further insight into data from Wave 1 of the Patient Reported Outcomes Measurement Information System (PROMIS) [8].

Statistical Similarities and Output

Some network models are statistically equivalent to latent variable models [4, 5, 9, 10]. Both network modeling and latent variable models can propose a model for the covariance

among variables, or symptoms, but do so in different ways. Confirmatory factor analysis proposes that the covariance among symptoms is due to the causal effect of latent variables, and estimated parameters consist of factor loadings, error variances, and correlations among latent variables. While there are many variants of psychological network models, we focus on the Gaussian graphical model (GGM) as it is one of the most widely implemented for cross-sectional data [11]. The GGM similarly operates on the covariance matrix among symptoms, but does so by proposing a model for its inverse, the precision matrix. The elements of the precision matrix can be standardized and signs reversed to produce a matrix of partial correlations between all pairs of variables. In the GGM, edges are partial correlations that represent undirected and weighted associations between two nodes (i.e., symptoms) after conditioning on all other nodes in the network [9].

In CFA models, factor loadings and factor correlations are the usual focus of model interpretation. With the GGM, partial correlations are the primary output, but additional summary indices are often computed from partial correlations or additional algorithms can help simplify interpretation. For example, *centrality indices* can help make conclusions about which symptoms appear to be most influential in the network. The concept of centrality is built on this perspective, such that more central nodes are judged as having more potential for influence on other nodes, and therefore have more promise as targets for clinical intervention [12, 13].

Traditional centrality indices for undirected weighted networks include strength, expected influence, closeness, and betweenness [14, 15]. These measures are calculated for each node in the network so that every symptom can be assessed relative to one another. Strength is defined as the sum of the absolute values of all edge weights (i.e., partial correlations) connected to the target node. Expected influence is nearly identical to strength, only it accounts for negative edge

signs and may be appropriate when both positive and negative relationships exist [15]. Closeness and betweenness depend on edges beyond those directly adjacent to the target node. Closeness is defined as the inverse sum of all shortest paths between a node and all other nodes, and betweenness is defined as the number of shortest paths that pass through the target node [15].

In comparing latent variable and network models, Hallquist and colleagues [6] investigated how centrality indices relate to factor loadings when latent variables are known *a priori*. In one simulation study, the authors generated data from CFA models with 10 indicators per factor (loadings from .4 to .95), one to three standardized factors (either orthogonal or correlated at $r = .40$), and equal residual item variances. The authors found that strength from the GGM correlated highly with factor loadings from a CFA (r 's $\geq .97$), and that this correspondence was sustained as the number of factors increased. Closeness and betweenness were less strongly related to factor loadings, and correspondence dipped substantially with multiple factors. Centrality indices exhibited stronger correspondence with loadings when the factors were correlated.

Assuming representativeness, results from Hallquist et al. [6] may have implications for network theory and model interpretation. For example, symptoms high on strength may be little more than good indicators of an underlying latent variable. However, the data generating conditions may be unlikely to hold in real symptom data as a single dataset might include more than three factors, an unbalanced number of items per factor, some variability in the strength of loadings across factors and in factor correlations, and some model misspecification.

Exploratory Graph Analysis

Another way to interpret results from network models involves forming clusters of items with many substantial edges, often known as *communities*. Golino and Epskamp [7] argue that

each community may represent symptoms that interact with one another and exemplify some psychological disorder, or may represent a latent variable or underlying common cause. Once communities of items are identified, they can be examined for their compatibility with theory. Both within and across communities, researchers may then focus on the density or number of connections [16] or examine relationships of individual nodes.

Exploratory graph analysis (EGA) was invented as an alternative to exploratory factor analysis and can empirically define communities from networks [7]. EGA operates by estimating a GGM, and then using the Walktrap algorithm [17] to characterize node clusters based on the partial correlation estimates. This algorithm identifies the number of communities and assigns each item to a community [7, 18]. This procedure is analogous to exploratory factor analysis in that both approaches can be used to choose the number of latent dimensions and with which factor (or community) each item is most strongly associated.

Golino and Epskamp [7] found in simulations that EGA could better detect the number of factors than traditional methods. EGA was initially only tested with simulated data generated from a model with 2 or 4 factors, with an equal number of items per factor (5 or 10), and with equal loadings for each item. Follow-up simulations were conducted under more realistic conditions [19]: loadings were different for each factor (.4, .55, .7, and .85 with a range of $\pm .1$ around these values) and small crossloadings were possible. The number of items per factor was the same for each factor (3, 4, 8, or 12) and factor inter-correlations were the same among all factors (0, .3, .5, or .7). The simulated data came from either a multivariate normal distribution or dichotomous items with skewness added to the underlying variables.

The above simulations only evaluated EGA's ability to detect the correct number of latent variables. Results for how EGA assigns items to communities are based on empirical

applications, but initial applications have come from datasets that test cognitive abilities, such as inductive reasoning [7] and intelligence [20], or social desirability [19]. In general, such applications sometimes resulted in the exact theoretical structure being replicated by EGA [7], but sometimes suggested a different plausible and improved solution [e.g., 19].

PROMIS and Study Purpose

In light of recent methodological and statistical comparisons between these frameworks [21], there is a need for more empirical comparisons with real data relevant to the measurement of mental and physical health related symptoms. Data from PROMIS [8] represents a good candidate. PROMIS item banks contain hundreds of items measuring symptoms from at least 9 different mental and physical health domains – Anxiety, Depression, Fatigue, Alcohol Use, Anger, Pain Interference, Satisfaction with participation in discretionary activities, Satisfaction with participation in social roles, and Physical Function. PROMIS is one of the most ambitious and rigorous recent non-cognitive test development endeavors in that item development and refinement involved qualitative and multiple quantitative stages, and used state-of-the-art techniques [8, 22, 23]. Given the widespread use of PROMIS, practitioners may find the results of symptom networks to be informative and EGA analyses may shed light on the validity of symptom domains.

The PROMIS Wave 1 data we use was designed for test construction and therefore assumes that latent variables are a reasonable way to model the data. However, the self-report items represent a variety of mental and physical health symptoms, which are exactly the kinds of variables that network researchers like to examine for causal relationships among symptoms. With this dataset, we seek to conceptually replicate the first simulation study by Hallquist and colleagues [6] with many items, many factors, an unbalanced number of items per factor, and

more variability in the strength of loadings. Our second goal was to investigate network dimensionality and community membership with EGA.

Methods

Data

PROMIS Wave 1 data were collected from general and clinical U.S. populations between July 2006 and March 2007, and are publicly available at <https://dataverse.harvard.edu> [24, 25]. To ensure reasonable CFA models, we selected only items that were identifiable from final item bank version 1.0 of Anxiety, Depression, Fatigue, Satisfaction with participation in discretionary activities, Satisfaction with participation in social roles, Alcohol Use, version 1.1 of Anger and Pain Interference, and version 1.2 of Physical Function [25]. Candidate items not included in a final item bank were not included in our analysis¹.

Data collection followed a planned missing data design with 23 overlapping test forms. If some items are never observed together at the same time (i.e., do not have pairwise coverage), the GGM may not be identified². Forms A through G were designed with pairwise coverage within any given domain, but had at most only 2 domains represented at a time. Since our goal was to examine high-dimensional network and CFA models, we focused on 16 test forms H through W and conducted separate analyses on items from these test forms. Each form included complete data on between 7 and 21 candidate items from each of the 9 domains, ensuring pairwise coverage among the selected items. We excluded domains from analysis if fewer than 3 items from the final bank were available and combined satisfaction with participation in discretionary activities and satisfaction with participation in social roles into one (Satisfaction) to

¹ For details on data collection and sample selection criteria, see Cella et al. [8].

² We tried a model based on full depression and anxiety banks with short form items from other domains, but it was not identified. These details shall be reported in a cautionary note on estimation of the GGM under missing data.

boost item retention. Across all forms, the number of domains ranged from 4 to 8, the number of items per domain ranged from 3 to 19, and the total number of items ranged from 36 to 51. After selecting items based on these 16 test forms, all available data from that test form, and from participants who completed at least some of the corresponding items in forms A through G, were included in our analysis. For brevity, we refer to each analysis being from each “test form.” The final sample sizes across the 16 datasets ranged from N 's = 6,261 to 9,022. On average, each test form had 67% missing data.

All items used a 5-point Likert scale, and most items referenced “the past 7 days” for context; temporal context varied for items in the Physical Function domain. For items in the Anxiety, Depression, Fatigue, Alcohol Use, Anger, and Pain Interference domains, a higher score indicated a more negatively valenced endorsement. For the remaining domains, a higher score meant more Satisfaction or better Physical Function.

Analysis

All analyses were executed in R version 4.0.3 [26]. On data from each test form, we conducted a CFA with freely estimated factor correlations and fit a GGM with EBICGLASSO regularization [27, 28, 29]. In total, we estimated 16 CFA models and 16 GGMs. The CFA models were estimated with *lavaan* (version 0.6-8) [30], using full information maximum likelihood (FIML) to handle missing data with robust (Hubert-White) standard errors. The GGMs were fitted with the *bootnet* package (version 1.4.7) [31]. *bootnet*'s estimation routine resembles a two-stage (TS) estimation approach from structural equation modeling [32, 33]. The saturated correlation matrix is first estimated using FIML, and that matrix is then used to estimate a regularized GGM using the *glasso* package [34] as if the correlation matrix were based on complete data. To choose the tuning parameter for regularization, the likelihood used in

EBIC is based on assuming complete data from this second stage, with N chosen as the pairwise average sample size³. We used default minimum tuning parameter ($\lambda = .01$) and hyperparameter ($\gamma = .5$) values from the *qgraph* package (version 1.7.1) [31, 35]. Community detection with the Walktrap algorithm was done on the already estimated GGMs using *igraph* (version 1.2.6) [36]⁴.

Results

Model Fit Summary

As a preliminary check, model fit for CFA models was generally good (CFI's $\geq .87$; SRMR's $\leq .07$; RMSEA's $\leq .05$) with full results in Supplementary Materials (Tables S1 through S33). All standardized loadings were above .35 and in the expected direction based on theory (only Physical Function had any loadings below .5), mean standard errors rarely exceeded .01, and factor correlations were congruent with content.

For each GGM, nonparametric bootstrapping indicated that edge weights were accurate (Figures S1-S3). Case dropping to investigate centrality stability yielded correlation stability coefficients [11] that were high for expected influence (.75 across all test forms), strength (median = .75), and closeness (median = .52), but lower for betweenness (median = .28; Table S34). Finally, within-domain network densities (median = .76, range = [.65, .83]) were substantially higher than between-domain densities (median = .24, range = [.19, .28]; Table S35).

Comparing Centrality Indices and Factor Loadings

Correspondence of centrality indices with standardized loadings was assessed with Pearson correlations separately for each test form but across all factors within a given form. Results showed similar patterning compared to Hallquist et al. [6] yet overall weaker associations (see Figure 1). Average correlations per centrality index were computed by applying

³ Computation of this sample size required forking *bootnet*: <https://github.com/falkcarl/bootnet>.

⁴ Code for all analyses is available at: <https://osf.io/79ys4/>

Fisher's z -transformation to all correlations in a given row of Figure 1, taking the unweighted average, and then backtransforming to r . Recall that in simulated data, Hallquist et al. [6] found that strength was highly correlated with factor loadings ($r \geq .97$), and that closeness and betweenness were less so ($r \leq .55$ when models had multiple factors). In our case, correspondence with loadings was strongest on average for strength ($r = .70$), less strong for expected influence ($r = .49$), weaker still for closeness ($r = .24$), and near zero for betweenness ($r = .08$). If, alternatively, we compute correlations separately for each test form *and* each factor, correspondence becomes stronger for strength ($r = .91$), expected influence ($r = .79$), and betweenness ($r = .27$), yet remains the same for closeness. Thus, within a given factor, items with stronger loadings are typically also high on strength, which suggests more consistency with results from Hallquist et al. [6].

Within individual test forms, the pattern of correspondence across centrality indices was similar with the following exceptions. In forms K, N, and P, correspondence with loadings was stronger for closeness than for expected influence. For form O, expected influence correlated more strongly with loadings than strength, while closeness exhibited a weaker (albeit negative) correlation than betweenness. In general, correspondence within test forms ranged from $r = .35$ to $.83$ for strength, $r = .14$ to $.74$ for expected influence, $r = -.16$ to $.68$ for closeness, and $r = -.24$ to $.43$ for betweenness.

EGA Community Detection

EGA results were evaluated relative to PROMIS domains within a given test form. The Walktrap algorithm determines community membership by optimizing the modularity index, a measure of cluster discreteness [37]. Modularity can range from 0 to 1, with 0 suggesting that within-cluster edges appear no more often than if generated by random assignment, and values

above .70 are unlikely in practice [37]. In our analysis, the median modularity index was .65 and 14 of 16 models had indices above .50 (excepting forms P and R; see Table S36).

EGA returned some discrepancies with PROMIS domains in the number of communities in a given network and in community composition (see Figures 2 and 3, and S4 through S17). More communities than PROMIS domains were recommended in 7 of 16 models, and fewer communities were suggested in 3. For the remaining 6 models, while the algorithm recommended an equal number of communities as PROMIS domains, EGA clustering never matched PROMIS domains exactly. Across most test forms, the Physical Function items split into two or more distinct communities, and Anger, Anxiety, and Depression items clustered together into one large community (e.g., form H, in Figure 2). On occasion, items from the Fatigue domain joined together with all Anger, Anxiety, and Depression items (e.g., form T, in Figure 3). The Walktrap also sometimes recommended communities comprised of only a single item, with Satisfaction items representing the most single-item communities of any domain (the others being Physical Function, Pain Interference, and Fatigue items). We note that the least modular models (i.e., weakest modularity indices) were also those with the most single-item communities (forms P and R).

Weak standardized loadings for Physical Function items from our CFA models further supports the possibility that multiple clusters exist. One explanation could be that different types of physical functioning are captured. In form K, for example, item content suggests that the Walktrap may be picking up on a meaningful distinction between strenuous physical activities (e.g., doing two hours of physical labor) and tasks requiring minimal whole-body mobility (e.g., changing the bulb in a table lamp).

When multiple domains clustered together as one community, it could reflect the presence of cross-domain processes among symptoms (i.e., comorbidity). For instance, the relatively strong association between EDANX46 and EDDEP44 suggests that the co-occurrence of feeling nervous and feeling disproportionately upset may represent a symptom-level marker of co-occurring anxiety and depression states. Similarly, the strong association between EDANG18 and EDDEP21 might suggest that feelings of guilt or blameworthiness could be elevated in individuals who are both angry and depressed. Otherwise, a strong between-domain association could represent content redundancy, a misassigned item, or a poor item.

When the Walktrap recommended a single-item cluster, the division sometimes made sense based on content. For example, on form H (see Figure 2), satisfaction related to time spent doing leisurely activities (SRPSAT05) was separate from satisfaction related to engagement in family (SRPSAT06) or work (SRPSAT07) responsibilities.

Discussion

To compare the latent variable and network modeling frameworks, we set out with two main objectives. First, drawing inspiration from Hallquist et al. [6] we evaluated the correspondence between factor loadings and centrality indices with real symptom data from PROMIS with many variables, an imbalance of variables across domains, and all hypothesized to covary due to several correlated latent factors. Second, we used the Walktrap algorithm from EGA to evaluate groupings of items defined by PROMIS and borne from the latent variable modeling tradition.

Across 16 correlated-factors CFA models and GGMs, the pattern of correspondence between factor loadings and centrality indices matched the pattern found by Hallquist and colleagues [6]; strength had the strongest correspondence, followed by closeness then

betweenness. However, these correlations all had magnitudes smaller than their counterparts in Hallquist et al. [6]. This theme held true even when considering each domain separately. Furthermore, expected influence – which was not previously evaluated – usually correlated less strongly with factor loadings than strength, and sometimes even correlated less strongly than closeness. Based on these results, strength centrality may provide slightly different information than the loadings, such that symptoms with strong factor loadings may not always be those purported to have high local influence on the network.

Closeness and betweenness appear to be different entities altogether compared to factor loadings. This makes sense as these indices are believed to reflect more networkwide than local connectivity [38] and our CFA models did not allow cross-loadings. Closeness and betweenness showed more instability compared to strength, and substantial instability in some GGMs (see Supplementary Materials). This is consistent with later simulation results in Hallquist et al. [6] that suggest vulnerability under small increases in sampling variability and some misspecification due to unmodeled small factors or unaccounted for local dependence. Other researchers [e.g., 39] argue that closeness and betweenness are conceptually problematic as the *distance* between nodes may not have clear meaning when nodes are symptoms.

To address our second objective, EGA applied to each of the 16 GGMs found that across models, 1) Physical Function items often separated into two or more communities, 2) Anger, Anxiety, Depression, and sometimes Fatigue items typically merged together as one large community, and 3) many single-item communities emerged from different domains. We therefore agree with Golino and Demetriou [20] that EGA and CFA can be complementary for test construction. In general, despite good CFA model fit on the whole, our EGA results could be taken as evidence that some PROMIS domains may benefit from revision.

Our EGA results support the idea that Physical Function is multidimensional. Regarding the fusion of Anxiety and Depression, in a highly regarded prevalence and comorbidity study using actual diagnosis [40], generalized anxiety disorder and major depression had a tetrachoric correlation of .62, implying substantial co-occurrence. With Fatigue, the DSM-5 includes fatigue as a criterion for both generalized anxiety disorder and major depressive disorder, and researchers have implicated fatigue as a potential bridge between anxiety and depression presentations [2]. For Anger, the connection with anxiety, depression, and fatigue is less straightforward, though links via irritability have been suggested in the literature. In clinical populations, irritability and anger have actually been shown to share a single dimension [e.g., 41]. In the DSM-5, irritability is a criterion for numerous anxiety disorders, and despite only appearing as a criterion for major depression in children and adolescents, studies have also found high rates of “overt irritability/anger” [42] and “subjective anger” [43] in depressed adults.

We make note of several limitations. First, differences in several small analytical decisions could be further explored. For example, we chose to focus on traditional centrality indices (strength, closeness, betweenness) and added a new index: expected influence. We made this decision so as to maintain close correspondence with Hallquist et al. [6]. Exploration of other indices – like so-called “network-loadings” [21] or bridge centrality indices [44] – may shed additional light on similarities between the GGM and CFA approaches.

Next, the choice of psychological scale (or particular items) may have influence over the structure of the estimated network [45] and could impact symptom rankings on centrality indices, their relationship with factor loadings, or the formation of communities. While we fit models to 16 different subsets of the PROMIS data, these models should not be taken as completely independent; they were still fitted to subsets of a single dataset and based on the same core item

banks. Future studies could evaluate correspondence between centrality indices and factor loadings or EGA results with similar domains but developed independently of PROMIS.

In our study, correspondence between factor loadings from CFA and centrality indices from network modeling partially replicated Hallquist et al. [6] with real data. Despite the statistical similarities between latent variable and network models, strength centrality is highly related to factor loadings but may provide unique information in real mental and physical health data. In the same context, community detection with EGA did not recover the hypothesized factor structure based on PROMIS domains. Instead, results imply that EGA may offer new insight into the composition of the investigated constructs and associated PROMIS test forms, as well as the relationships among them.

References

1. Borsboom D, Mellenbergh GJ, van Heerden J. The theoretical status of latent variables. *Psychol Rev.* 2003;110(2):203–19.
2. Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol.* 2013;9(1):91–121.
3. McNally RJ. Can network analysis transform psychopathology? *Behav Res Ther.* 2016;86:95–104.
4. Marsman M, Borsboom D, Kruis J, Epskamp S, van Bork R, Waldorp LJ, et al. An introduction to network psychometrics: Relating Ising network models to Item Response Theory models. *Multivariate Behav Res.* 2018;53(1):15–35.
5. van Bork R, Rhemtulla M, Waldorp LJ, Kruis J, Rezvanifar S, Borsboom D. Latent variable models and networks: Statistical equivalence and testability. *Multivariate Behav Res.* 2019;1–24.
6. Hallquist MN, Wright AGC, Molenaar PCM. Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behav Res.* 2019;1–25.
7. Golino HF, Epskamp S. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One.* 2017;12(6):e0174035.
8. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol.* 2010;63(11):1179–94.
9. Epskamp S, Waldorp LJ, Möttus R, Borsboom D. The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behav Res.* 2018;53(4):453–80.
10. Molenaar PC. Latent variable models are network models. *Behav Brain Sci.* 2010;33(2–3):166.
11. Epskamp S, Rhemtulla M, Borsboom D. Generalized network psychometrics: Combining network and latent variable models. *Psychometrika.* 2017;82(4):904–27.
12. McNally RJ, Robinaugh DJ, Wu GW, Wang L, Deserno MK, Borsboom D. Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clin Psychol Sci.* 2015;3(6):836–49.

13. Ruzzano L, Borsboom D, Geurts HM. Repetitive behaviors in autism and obsessive-compulsive disorder: new perspectives from a network analysis. *J Autism Dev Disord*. 2015;45(1):192–202.
14. Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths. *Soc Networks*. 2010;32(3):245–51.
15. Robinaugh DJ, Millner AJ, McNally RJ. Identifying highly influential nodes in the complicated grief network. *J Abnorm Psychol*. 2016;125(6):747–57.
16. Burger J, Isvoranu A-M, Lunansky G, Haslbeck J, Epskamp S, Hoekstra RH, et al. Reporting standards for psychological network analyses in cross-sectional data. *PsyArXiv*. 2020. Available from: <http://dx.doi.org/10.31234/osf.io/4y9nz>
17. Pons P, Latapy M. Computing communities in large networks using random walks. In: *Computer and Information Sciences - ISCIS 2005*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005. p. 284–93.
18. Christensen AP, Garrido LE, Golino H. Comparing community detection algorithms in psychological data: A Monte Carlo simulation. *PsyArXiv*. 2020. Available from: <http://dx.doi.org/10.31234/osf.io/hz89e>
19. Golino H, Shi D, Christensen AP, Garrido LE, Nieto MD, Sadana R, et al. Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychol Methods*. 2020;25(3):292–320.
20. Golino HF, Demetriou A. Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence*. 2017;62:54–70.
21. Christensen AP, Golino H. On the equivalency of factor and network loadings. *Behav Res Methods*. 2021; Available from: <http://dx.doi.org/10.3758/s13428-020-01500-6>
22. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22-31.
23. DeWalt DA, Rothrock N, Yount S, Stone AA, PROMIS Cooperative Group. Evaluation of item candidates: the PROMIS qualitative item review: The PROMIS qualitative item review. *Med Care*. 2007;45(5 Suppl 1):S12-21.
24. Liu HH, Cella D, Gershon R, Shen J, Morales LS, Riley W, et al. Representativeness of the PROMIS Internet panel. *J Clin Epidemiol*. 2010;63(11):1169-78.
25. Cella D. PROMIS 1 Wave 1. Harvard Dataverse; 2015. Available from: <https://dx.doi.org/10.7910/DVN/ONGAKG>

26. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: <https://www.R-project.org/>
27. Epskamp S, Fried EI. A tutorial on regularized partial correlation networks. *Psychol Methods*. 2018;23(4):617–34.
28. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*. 2008;9(3):432–41.
29. Chen J, Chen Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*. 2008;95(3):759–71.
30. Rosseel Y. lavaan: An R package for structural equation modeling. *J Stat Softw*. 2012;48(2):1–36. Available from: <http://www.jstatsoft.org/v48/i02/>
31. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. *Behav Res Methods*. 2018;50(1):195–212.
32. Yuan K-H, Bentler PM. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociol Methodol*. 2000;30(1):165–200.
33. Allison PD. Missing data techniques for structural equation modeling. *J Abnorm Psychol*. 2003;112(4):545–57.
34. Friedman J, Hastie T, Tibshirani R. glasso: Graphical Lasso: Estimation of Gaussian Graphical Models. R-project.org. 2011. Available from: <http://CRAN.R-project.org/package=glasso>
35. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network visualizations of relationships in psychometric data. *J Stat Softw*. 2012;48(4):1–18. Available from: <http://dx.doi.org/10.18637/jss.v048.i04>
36. Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal Complex Systems*. 2006;1695(5):1–9.
37. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69(2 Pt 2):026113.
38. Borsboom D. A network theory of mental disorders. *World Psychiatry*. 2017;16(1):5–13.
39. Bringmann LF, Elmer T, Epskamp S, Krause RW, Schoch D, Wichers M, Wigman JT, Snippe E. What do centrality measures measure in psychological networks? *J Abnorm Psychol*. 2019;128(8):892–903.

40. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617–27.
41. Pasquini M, Picardi A, Biondi M, Gaetano P, Morosini P. Relevance of anger and irritability in outpatients with major depressive disorder. *Psychopathology*. 2004;37(4):155–60.
42. Judd LL, Schettler PJ, Coryell W, Akiskal HS, Fiedorowicz JG. Overt irritability/anger in unipolar major depressive episodes: Past and current characteristics and implications for long-term course. *JAMA Psychiatry*. 2013;70(11):1171–80.
43. Genovese T, Dalrymple K, Chelminski I, Zimmerman M. Subjective anger and overt aggression in psychiatric outpatients. *Compr Psychiatry*. 2017;73:23–30.
44. Jones PJ, Ma R, McNally RJ. Bridge centrality: A network approach to understanding comorbidity. *Multivariate Behav Res*. 2019;1–15.
45. Fried EI, Cramer AOJ. Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspect Psychol Sci*. 2017;12(6):999–1020.

Statements and Declarations

Funding: The authors acknowledge the support of the Fonds de la recherche en Santé du Québec (FRQS), doctoral training award reference number BF2-2022-302883; the Natural Science and Engineering Research Council of Canada (NSERC), funding reference number DGEER-2018-00083; and the Fonds de recherche du Québec – Nature et technologies (FRQNT), funding reference number 2019-NC-255344.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Author Contributions: Both authors contributed to study conception and design, co-wrote code for conducting analyses and generating results. Joshua Starr wrote the first draft of the manuscript, and both authors contributed to editing. All authors read and approved the final manuscript.

Ethics Approval: As the manuscript involves re-analysis of publicly available data, it did not require ethics approval.

Figure 1. Plots of Correspondence Between Centrality Measures and Factor Loadings

Note. Exp. Infl. = Expected Influence

Figure 2. EGA Results for Test Form H

PROMIS Domains

Communities by Walktrap

Note. acl = Alcohol Use; ang = Anger; anx = Anxiety; dep = Depression; fat = Fatigue; pain = Pain Interference; ph = Physical Function; sp = Satisfaction

Figure 3. EGA Results for Test Form T

PROMIS Domains

Communities by Walktrap

Note. anx = Anxiety; dep = Depression; fat = Fatigue; pain = Pain Interference; ph = Physical Function; sp = Satisfaction