Evaluating General Network Scoring Methods as Alternatives to Traditional Factor Scoring Methods

Jihong Zhang and Xinya Liang

Department of Counseling, Leadership, and Research Methods, University of Arkansas

Author Note

Jihong Zhang http://orcid.org/000-0003-2820-3734

Xinya Liang http://orcid.org/0000-0002-2453-2162

Correspondence concerning this article should be addressed to Jihong Zhang, Department of Counseling, Leadership, and Research Methods, University of Arkansas, 109 Graduate Education Building, Fayetteville, AR 72703, Email: jzhang@uark.edu

Abstract

Network psychometrics have been getting more attention in psychopathology and measurement development. As a novel psychometric method, multiple psychometric properties of psychometric modeling have been compared to factor analysis models. However, compared to scoring in factor analysis and classical test theory, the scoring method in psychometric network modeling is not well examined. Thus, this study aims to propose novel network scoring methods and compare them to factor scores. In this study, we first proposed the general form of network scoring and conducted two simulation studies and one empirical study to examine the performance of different network scores. The results of this study suggest that non-regularized network scores performed better than regularized network scores and factor scores under certain conditions, especially non-regularized network scores with node strength and hybrid centrality. The limitations and future directions are also discussed.

Keywords: Network Score, Psychometrics, Factor Analysis

Evaluating General Network Scoring Methods as Alternatives to Traditional Factor Scoring Methods

Introduction

In psychology and education, measurement scoring plays an essential role in assessment, evaluation, and measurement. Measurement scoring has been considered as an empirical process of "assigning numbers to objects or events according to a rule" and "giving meaning to the theoretical variables" (Avila et al., 2015). In psychometrics, measurement scores rely on measurement modeling that describes the relationships between a *construct* and its *indicators*, which can then quantify the abstract phenomenon of interest (Diamantopoulos et al., 2008). One property of measurement scoring is that its interpretation varies across different psychometric models and psychometric theories. For example, classical test theory assumes each person has an innate *true score*, with the observed score reflecting his/her true score plus random error. Alternatively, common factor theory posits that one person can have one or multiple factor scores corresponding to measured constructs (e.g., a person's liability to mental disorder, severity of symptoms, math ability, etc.). These factor scores aim to capture the entirety of underlying constructs (e.g., He et al., 2022; Uher et al., 2008). In large-scale assessments, such as Programme for International Student Assessment (PISA) or Trends in International Mathematics and Science Study (TIMMS), imputed factor scores (plausible values) are provided as estimates of the student proficiency. Recently, with increasing empirical studies using psychological network models (Borsboom, 2017; e.g., Fried et al., 2017; McElroy et al., 2019; Monteleone et al., 2023; Monteleone & Cascino, 2021; Stella, 2022), a novel measurement score estimation method relying on network theory, named *network score estimation*, receives increasing attention as an alternative to factor scoring.

Recent simulation studies (Christensen, 2018; Ouyang et al., 2023) have shown that network scores obtained using network score estimation were similar to or better than factor scores estimated by the confirmatory factor analysis (CFA). One reason is that network analysis is more flexible than CFA regarding its assumptions of data structures. Network analysis employs unrestricted modeling, often the most natural and flexible approach for calibrating and scoring items, as most items or observable variables are inherently complex (Ferrando & Lorenzo-Seva, 2018). More importantly, network analysis assumes direct connections among observed variables, rather than attributing latent variables as causal factors behind a series of observed behaviors. Thus, it is less affected by reliability issues stemming from model structure misspecifications (cross-loadings and residual covariances), as the analysis does not require strict local independence assumptions (Ouyang et al., 2023). However, previous studies have focused solely on a specific type of network scores obtained using hybrid centrality, rather than general network scores derived from other centrality measures. Currently, no consensus exists on the statistical form of network scoring, making the network scores challenging to interpret. Moreover, no study has investigated how different estimation approaches of network models (e.g., saturated, regularized) and data characteristics affect individual network scoring.

To address these research gaps, this study conducts two simulation studies to compare the performance of various network scores and factor scores. This includes (1) proposing and justifying the general statistical form of network scores; (2) examining the effects of data structures (e.g., test length, sample size, dimensionality, etc.) on the estimation of network scores; (3) examining the effects of centrality measures and regularization approaches on network scoring estimations. In addition, an empirical study with the Big Five personality traits test was conducted to illustrate the utility of network scores in a real-world scenario. Specifically, this study aims to answer the following research questions:

1. What is the general statistical formula of network scores, and can different types of network scores reflecting varied aspects of network structure be developed based on the general formula?
2. Does data structure affect the performance of different types of network scores compared to factor scores? If so, which data characteristic has the strongest influence on network score estimation.
3. Do different network scores vary in stability and accuracy? If so, which network scores perform best under unidimensional and multidimensional data generation model?

The study is organized as follows. First, we provide background on psychological network analysis and regularization. Second, we illustrate the general mathematical form of network scores and their relationships to factor scores. Third, we conducted two Monte Carlo simulation studies to compare the performance of network score variants to factor scores, focusing on their relationships to data-generating factor scores under latent factor models. Finally, we demonstrate the practical application of network scores using the Big Five questionnaire in an empirical study.

Network Analysis

Network analysis offers an alternative framework of factor analysis by representing dependency among observed features as a network structure. Different from the hypothesis of factor analysis that common factors are used to explain the shared variance between observed variables, network analysis assumes that correlations between observed variables result from processes that engage in mutual interactions (Cramer et al., 2010; Van Der Maas et al., 2006). The psychological network modeling with multivariate data aims to identify the network structure of observed variables (also known as *nodes*) and estimate the importance of variables (Borsboom et al., 2021a). By applying psychological network modeling to survey data, psychometric network analysis has gained more attention in areas such as psychopathology, intelligence, and personality. For instance, an empirical network analysis study on internet addiction and depression suggested that two symptoms, "*guilty*" and "*escape*", activate the negative feedback loop and further contribute to the comorbidity between internet addiction and depression (Zhao et al., 2023).

Network analysis has demonstrated strong equivalence with factor analysis, despite their divergent hypothesized causal processes (Christensen & Golino, 2021). The core part of network analysis is the Gaussian graphical model (GGM). GGM is the statistical form of psychometric network analysis for cross-sectional continuous multivariate data. Compared to factor analysis models that estimate model parameters such as factor loadings, factor correlations, and item intercepts in factor analysis, GGM estimates partial correlations between nodes. The GGM reproduces the variance-covariance matrix $\boldsymbol{\Sigma}$ among observed variables using the following equation (Epskamp et al., 2017a; Epskamp, Waldorp, et al., 2018a; Epskamp, 2020):

$$\boldsymbol{\Sigma} = \boldsymbol{\Delta}(\mathbf{I} - \boldsymbol{\Omega})^{-\mathbf{1}}\boldsymbol{\Delta} \qquad (1)$$

where $\boldsymbol{\Delta}$ is a diagonal scaling matrix that controls the variances, and $\boldsymbol{\Omega}$ is a square symmetrical matrix with 0s on the diagonal and partial correlation coefficients on the off diagonal. In psychometric network analysis, off-diagonal elements in the partial correlation matrix, ω_{ij}, represent the partial correlation between node i and node j. A higher value of a partial correlation represents a stronger connection between two nodes controlling for the influences of other nodes.

Regularization

One limitation of estimating a GGM with small sample size is that the number of parameters to estimate grows rapidly as the size of the network increases (Epskamp, Borsboom, et al., 2018). To address this challenge, *regularization* techniques are applied to obtain sparse (or conservative) network structures: only a relatively small number of node edges remain to explain the covariation structure in the data. Regularization can reduce the risk of overfitting and improve model interpretation (Epskamp & Fried, 2018). In this study, the network models that have not been regularized are called *non-regularized* or *saturated networks*, while the network models in which weak node edges has been removed through regularization are called *regularized* or *sparse networks*.

Multiple regularization algorithms have been developed. For example, graphical least absolute shrinkage and selection operator (*graphical LASSO;* Friedman et al., 2008) is a well-established and fast algorithm for estimating regularized GGM. Alternatively, Epskamp et al. (2020) proposed a method based on iterative model search and pruning within the SEM or network framework to simply the structure of a network. The SEM-based stepdown model search approach uses partial pruning to compare models with or without the network edges to iteratively select the best-fitting model. The method via the *modelsearch* algorithm can achieve a network structure with the similar level of sparsity. The *modelsearch* algorithm removes edges that are not significant at $\alpha = .01$ in each step, and then re-estimates the network model until no insignificant edges are included (Isvoranu & Epskamp, 2023). The simulation study of Isvoranu and Epskamp (2023) suggested that the stepwise *modelsearch* algorithm performed well in terms of *specificity* and *precision* of network structure estimation. Thus, we used the stepwise *modelsearch* algorithm to estimate the regularized network structures in this study.

Centrality Measures

One objective of the psychometric network analysis is to properly describe the network structure of target constructs. For example, for clinicians and psychopathology researchers, one of the most difficult challenges is how best to conceptualize the co-occurrence of symptoms of psychopathology (Hallquist et al., 2021). To resolve the challenges, recent empirical studies of network analysis focused on node-level metrics (e.g., centrality measures) from graph theory to explore whether particular symptoms (nodes) are central in a network (e.g., Li & Zhang, 2024; Liang et al., 2023; Zhao et al., 2023). Such measures of nodal importance are called *centrality measures*. Centrality measures aggregate information from the overall covariance structure to summarize the properties of one symptom relative to another (Hallquist et al., 2021). It has been argued that centrality measures may help to identify important targets (e.g., symptoms or features) that play a crucial role in precipitating other nodes in the network (Hofmann et al., 2016).

Multiple centrality measures have been developed, including: (1) betweenness centrality (BC); (2) randomized shortest paths betweenness centrality (RSPBC); (3) closeness centrality (LC); (4) node degree (ND); (5) node strength (NS); (6) expected influence (EI); (7) eigenvector centrality (EC; Borkulo et al., 2015); and (8) hybrid centrality (HC; Christensen, 2018). In the current study, betweenness centrality (BC), closeness centrality (LC), node strength (NS), and hybrid centrality (HC) are used to calculate network scores to investigate how different centrality measures affect the estimation of network scores. These measures were chosen for several reasons. Node strength is most frequently used in empirical studies (e.g., Hallquist et al., 2021; Robinaugh et al., 2020). Hybrid centrality has been used to calculate network scores, demonstrating comparable performance with factor scores (Christensen, 2018; Ouyang et al., 2023). Betweenness and closeness centrality are also widely reported in network analysis studies (e.g., Epskamp et al., 2018; Li & Zhang, 2024), which were derived from the concept of distance rather than network strength (Hallquist et al., 2021). Network scores based on distance measures (BC and LC) may reflect different aspects of the network analysis compared to strength-based measures and hybrid centrality. Thus, it is important to compare network scores derived from these different measures.

The node strength (S) centrality measures the sum of the absolute strength of edges (partial correlations in GGM) connected to a single node. The node strength for node i can be formed as:

$$\mathrm{S}_i = \sum_j |W_{ij}|$$

where W is the estimated partial correlation matrix among nodes with diagonal elements as 0.

The hybrid centrality is calculated by averaging various centrality measures' weighted or unweighted ranks, including BC, LC, NS, ND, and EC. The mathematical form of HC (Christensen, 2018; Ouyang et al., 2023) is as:

$$\mathrm{HC} = \frac{R_{BC}^{w} + R_{BC}^{u} + R_{LC}^{w} + R_{LC}^{u} + R_{ND}^{u} + R_{NS}^{w} + R_{EC}^{w} + R_{EC}^{u}}{8 \times (K-1)}$$

where w signifies the weighted measure and u signifies the unweighted measures, R represents the rank of the centrality measure, and K represents the number of nodes. Other centrality metrics, such as BC and LC, were distance-based and thus do not reflect correlation strength but the network structure in psychometric networks.

General Network Score

The mathematical form of network scores is essential within the network analysis framework, as it helps researchers understand how network scores reflect individuals' ranking in the dimensions of measured targets. For example, for a psychometric network of depression, it is not easy to explain why individuals with higher network scores suggest more severe depression without understanding the statistical form of network score. In addition, how network scores can be used to evaluate individuals is still debated. For example, in line with the traditional measurement of "attributes" in psychometrics and measurement of "properties" in metrology, van Bork et al. (2024) demonstrated that the measurement target of network analysis can be conceptualized as a nominal property of each individual rather than continuous or ordered categorical variables (e.g., whether someone is currently in a healthy or depressed stable state) for the diagnosis purpose. Other studies (e.g., Christensen & Golino, 2021) suggested that a network score is more analogous to a formative latent variable (i.e., a continuous weighted composite).

This study employs the latter perspective and considers network scores as continuous quantities. Unlike previous studies that the node importance can only be a certain centrality measure (e.g., hybrid centrality; Ouyang et al., 2023), we consider the node importance as exchangeable across multiple centrality measures, each capturing different aspects of importance within the network. The general form for computing the expected a posteriori (EAP) network scores $\boldsymbol{\eta}_\mathbf{i}$ is expressed as:

$$\mathrm{EAP}(\boldsymbol{\eta}_\mathbf{i}) = \mathbf{W}\mathbf{Y}_\mathbf{i} \qquad (2)$$

where $\mathbf{W}$ is the measure of node (or feature) importance and $\mathbf{Y}_\mathbf{i}$ is a vector of observed responses for person *i*. $\mathbf{W}$ may take any form of network centrality measures.

Factor scores and network scores have a close statistical relationship due to the statistical equivalence between the factor and network analyses (Ouyang et al., 2023). Thurstone's regression factor scores (Thurstone, 1934) consider factor scores as weighted composites, with unidimensional factor structure expressed as:

$$\text{EAP}(\boldsymbol{\theta}_{\mathbf{i}}) = \boldsymbol{\Phi}\boldsymbol{\Lambda}'\left(\widehat{\boldsymbol{\Sigma}}_{\mathbf{F}}\right)^{-\mathbf{1}}\mathbf{Y}_{\mathbf{i}} \qquad (3)$$

where the estimated sample variance-covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\mathbf{F}}$ is a function of factor correlation matrix ($\boldsymbol{\Phi}$), factor loadings ($\boldsymbol{\Lambda}$), and residual variance covariance matrix ($\boldsymbol{\Psi}$):

$$\widehat{\boldsymbol{\Sigma}}_{\mathbf{F}} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \qquad (4)$$

An advantage of general network scoring is its ability to bridge factor score and network score formulations. The mathematical form of network scoring is thus more general and comprehensive, as both Thurstone's regression factor scores and network scores can be viewed as special cases of general network scores. Specifically, when $\mathbf{W} = \boldsymbol{\Phi}\boldsymbol{\Lambda}'(\widehat{\boldsymbol{\Sigma}}_{\mathbf{F}})^{-\mathbf{1}}$ in (2), the network scores $\boldsymbol{\eta}_{\mathbf{i}}$ are equivalent to Thurstone's regression factor scores $\boldsymbol{\theta}_{\mathbf{i}}$.

However, it is noted that centrality measures and factor loadings have different scales and interpretations. For example, *node strength*, one frequently used centrality measure in network modeling, is an estimated partial correlation of item pairs, while factor loadings are estimated slopes of factor scores on item responses. Betweenness and closeness reflect the network's topological structure rather than partial correlations. Hybrid centrality presents the overall importance of each node by averaging rankings of multiple centrality measures. Therefore, using various centrality measures may lead to different performances and interpretations of network scores. In addition, regularization methods affect the performance of network scoring by reducing the total number of parameters in network models. To systematically assess the relative performance of network scores based on various centrality measures, we propose multiple network scores that are computed using different weights **W** based on both the non-regularized and regularized network models.

In the present study, to investigate the performance of network scores under various conditions, nine scores were used for comparison, including (1) factor scores estimated with confirmatory factor analysis (*FS*); (2) network scores based on hybrid centrality and node strength (*NS-H* and *NS-S*) with non-regularized network analysis; (3) network scores based on betweenness and closeness centrality (*NS-B* and *NS-C*) with non-regularized network analysis; (4) network scores based on hybrid centrality and node strength with regularized network analysis (*RegNS-H* and *RegNS-S*); (5) network scores based on betweenness and closeness centrality with regularized network analysis (*RegNS-B* and *RegNS-C*). Table 1 provides an overview of the eight network scores.

Table 1

Network Scores computed using different centrality measures for non-regularized and regularized models

Index	Scores	Regularization	Centrality Measures
1	NS-H	Non-regularized	Hybrid Centrality
2	NS-S	Non-regularized	Node Strength
3	NS-B	Non-regularized	Betweenness Centrality
4	NS-C	Non-regularized	Closeness Centrality
5	RegNS-H	Regularized	Hybrid Centrality
6	RegNS-S	Regularized	Node Strength
7	RegNS-RS	Regularized	Betweenness Centrality

Index	Scores	Regularization	Centrality Measures
8	RegNS-RH	Regularized	Closeness Centrality

The Present Study

Previous research has established the equivalency between network analysis and factor analysis, as well as similarity between network loadings and factor loadings (Christensen & Golino, 2021; Golino & Epskamp, 2017). Additional research also suggests that network scores based on the hybrid centrality are similar to CFA factor scores under correctly specified model (Christensen, 2018; Ouyang et al., 2023).

However, as Ouyang et al. (2023) suggested, the performance of various network scores formulated by different centrality measures remains unclear and needs further examination. In addition, how different estimation approaches of network analysis and what data characteristics affect individual network scoring has yet to be well examined. To fill the research gaps, this study develops different network scores that reflect different aspects of network structure and individual characteristics. The network scores are then compared with factor scores in terms of their relations to actual scores, providing guidelines on the most effective network score under various conditions. This study is organized into the following sections. First, simulation study 1 investigates the performance of eight network scores and factor scores under varied conditions of unidimensional factor models. Next, simulation study 2 extends to a multidimensional context. Finally, an empirical study using Big Five dataset demonstrates the application of these network scores in real-world settings.

Simulation Study 1

Simulation Study 1 investigated how various settings of generated data structures and network estimation methods influenced the relative performance of network scoring in the framework of the unidimensional factor structure.

Method

Study design

Since general network scores are influenced by observed data and estimated centrality measures (see Eq. 2), it is essential to examine the factors that affect data structures and network estimation. Drawing on the designs of previous simulation studies of network analysis (Curran et al., 2016; Hallquist et al., 2021; Ouyang et al., 2023), we manipulated four design factors in data generation (see Table 2): (1) sample size (N = 100, 200, and 500); (2) test length ($J = 6$, 12 or 24), reflecting a range of potential applications from short to long tests; (3) measurement quality, indicated by factor loading magnitudes ($\lambda \in \{.4, .7\}$); (4) residual correlation between the first and second indicators ($\psi_{12} \in \{0, .3\}$), to compare simple structure versus complex structure. In total, there were $3 \times 3 \times 2 \times 2 = 36$ conditions in Study 1. For each condition, 1,000 data sets were generated based on the above design factors in R version 4.2.1.

Table 2

Design factors for Study 1

Design factor	One-factor Model
Sample size	100, 200, 500
Number of Items	6, 12, 24

Design factor	One-factor Model
Factor loadings	0.5, 0.7
Residual correlation	0, .3

Data analysis

Both non-regularized and regularized network models are used to calculate network scores. Specifically, non-regularized network structures were estimated using the *psychonetrics* package (Epskamp, 2023). To obtain a sparse network structure, we applied the *modelsearch* algorithm (Epskamp et al., 2020) using the *prune* function of *psychonetrics* package. The *pruning* used a stepwise model search and continued until no more significant edges can be removed. Factor analysis was performed using the *lavaan* package (Rosseel, 2012). For centrality measures, node strength, hybrid centrality, betweenness, and closeness were computed using the *psychonetrics* package and the *NetworkToolbox* package (Christensen, 2018), respectively. All data and R scripts can be found on the Open Science Framework (OSF, https://osf.io/rsp7h/).

For each condition, three analysis models were estimated: (1) Model 1: a simple-structure one-factor model, (2) Model 2: a non-regularized network model, and (3) Model 3: a regularized network model. Note that the one-factor analytic model was a simple one-factor structure without residual covariance, which was correctly specified for data generated with $\psi_{12} = 0$, but misspecified when $\psi_{12} = .3$. In cases of mild model misspecification, the estimated factor scores may deviate from the true factor scores. After models were estimated, eight network scores were calculated using centrality measures, with four network scores from non-regularized network Model 2 (NS-S, NS-H, NS-S, NS-C) and four from regularized network Model 3 (RegNS-S, RegNS-H, RegNS-S, RegNS-C). Although multiple factor scores are available, we only reported factor scores calculated using Thurstone's regression method because other alternatives such as the Bartlett's method has shown similar accuracy (Ouyang et al., 2023).

Performance criterion

As for the performance criterion, correlations and root mean square errors (RMSE) between estimated network/factor scores and "true" factor scores were used (Ouyang et al., 2023). Throughout the study, we used $\bar{r}$ to represent the average correlation across replications and $\dot{r}$ to represent the average correlation across replications and conditions.

Results

According to the criteria for good model fitting (RMSEA < .05 and CFI/TLI > .90), across all factor models analyzed, 23 out of 36 conditions (63.89%) had an acceptable rate higher than .5, indicating more than half of replications having acceptable model fitting. Only 14 out of 36 conditions (38.8%) had more than 90% replications with the acceptable model fitting. It turned out that both model misspecification and small sample sizes could lead to low acceptable rates. For non-regularized network models, the acceptable rates were always 100% because the non-regularized network models perfectly represent the observed data sets. In contrast, regularized network models had lower acceptable rates on average (nearly zero) based on traditional factor-analysis fitting criteria because regularization led to varying degrees of model misspecification. In previous empirical studies, goodness-of-fit criteria were typically neither reported nor used for model rejection. Instead, the *Extended Bayesian Information Criteria* (EBIC) was used to select the best-fitting model (Epskamp, Waldorp, et al., 2018b, 2018a). How to use absolute model fit indices for psychological network models is out of the scope of the current study and is still debated. For interested readers, please refer to Epskamp, Waldorp, et al. (2018a, 2018b) for more details.

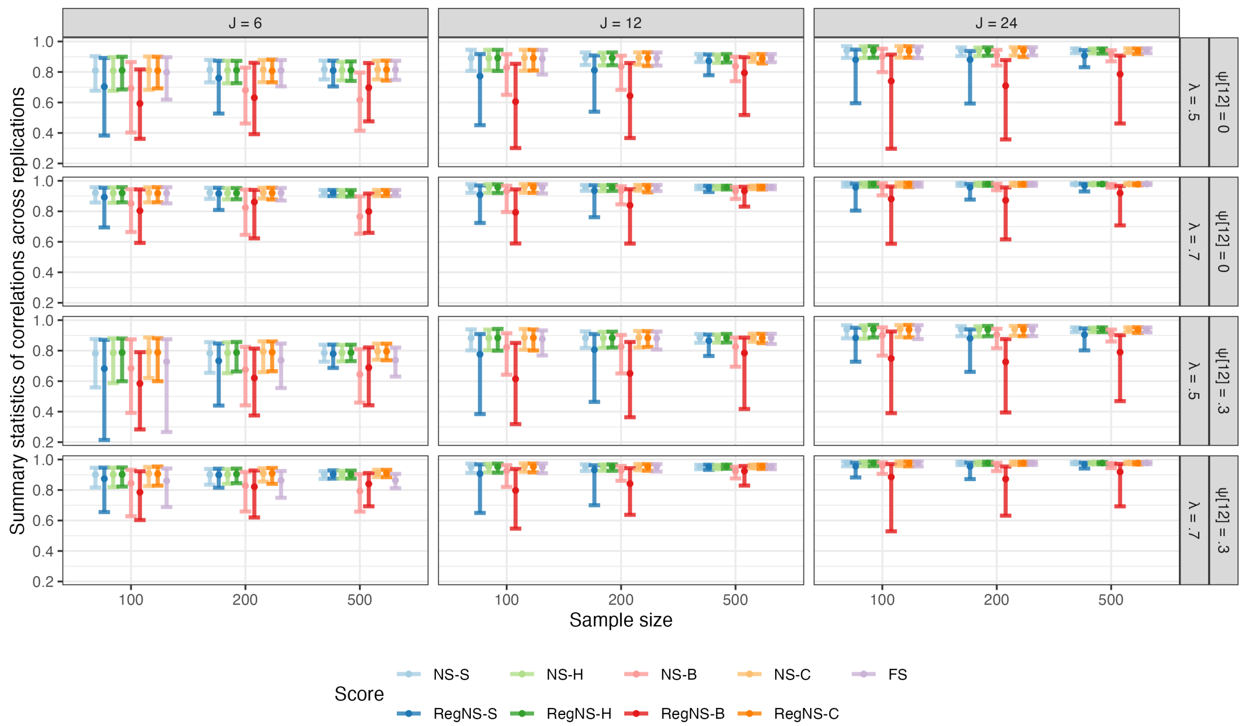
Figure 1 and Figure 2 present the correlations and RMSEs between the estimated scores with the true factor scores from data generation across 1,000 replications by various design factors. In addition, Table S1 and Table S2 in the supplementary materials present the estimated regression coefficients with correlation values as outcomes and design factors as predictors, which indicate the estimated effects on correlations on average.

The results suggested that without residual covariances ($\psi = 0$), except for network scores using betweenness (NS-B and RegNS-B) and regularized network scores using node strength (RegNS-S), network scores have higher correlations ($\bar{r} \in [.807, .979]$) than factor scores ($\bar{r} \in [.799, .979]$) across conditions. Among all 8 network scores, NS-S presented highest correlations ($\bar{r} \in [.809, .979]$) than other network scores. Non-regularized network scores ($\bar{r} \in [.617, .979]$) generally had a higher correlation than regularized network scores ($\bar{r} \in [.593, .979]$) across conditions. The results of RMSE (see Figure 2) suggested that all network scores except for betweenness-based network scores had similar RMSEs. FS yielded the lowest RMSEs followed by RegNS-S.

When the residual covariance was present ($\psi = .3$), the average correlations of network scores across all replications and conditions were close to those without the residual covariance ($\Delta\bar{r} \in [.0004, .0114]$). In addition, non-regularized network scores based on closeness, hybrid centrality, and node strength (NS-H, NS-S, NS-C) were shown to have higher correlations ($\bar{r} \in [.782, .978]$) than FS ($\bar{r} \in [.728, .978]$). As for the effects of sample sizes and item qualities, the results showed that larger sample sizes, longer test lengths, and higher factor loadings slightly improved accuracy and stability. Like conditions without model misspecification, network scores based on betweenness centrality (NS-B and RegNS-B) presented the lowest correlation ($\bar{r} \in [.584, .967]$) among all network scores.

Figure 1

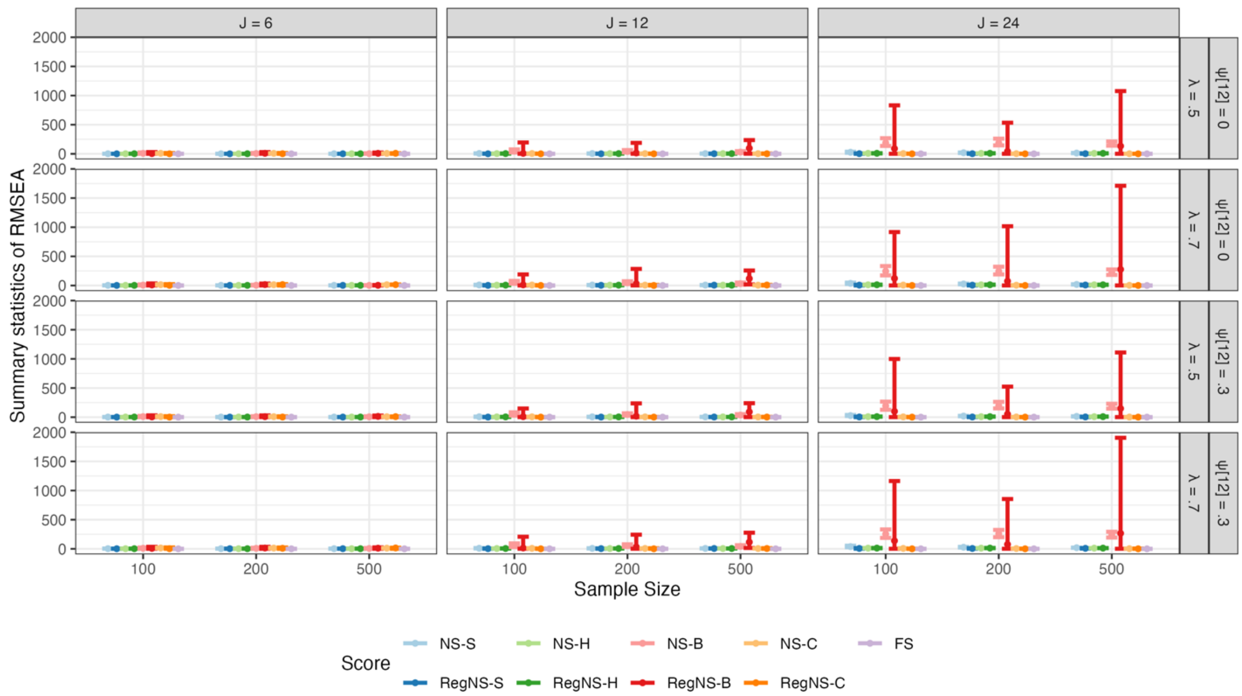
Correlations between estimated scores and data-generating factor scores in study 1



Note. The upper bound and the lower bound of the bar represent maximum and minimum values of correlations between each network score with factor scores across 1000 replications. The point denotes the average value of correlations.

Figure 2

RMSEs between estimated scores and data-generating factor scores in study 1



Note. The upper bound and the lower bound of the bar represent maximum and minimum values of correlations between each network score with factor scores across 1000 replications. The point denotes the average value of correlations.

Simulation Study 2

Simulation Study 2 examined how different settings of data generation and network estimation methods affected the relative performance of network scoring in the framework of the multidimensional factor structure.

Method

Study design

As shown in Table 3, we expanded upon the previous study (Ouyang et al., 2023) by incorporating additional design factors regarding factor correlations and model misspecifications. In total, we considered five design factors within the framework of three-factor models: (1) sample size (N = 100, 500); (2) number of items per dimension ($J_k = 3$ or 6, where k is the k^{th} dimension); (3) factor loadings ($\lambda_j \in \{.5, .7\}$) (4) factor correlations ($\phi_{kk'} \in \{.3, .7\}$); (5) residual covariances. For all population models, the factor loadings of the first item per factor was fixed to 1 (anchor items). Three residual covariance matrices were examined: RC1, with no residual covariance; RC2, with residual covariance between two items loading on the same factor; and RC3, with residual covariance between two items loading on different factors. Specifically, let $\psi_{ij,k}$ denote the residual covariance between items *i* and *j* loading on the factor *k,* and $\psi_{ij,kl}$ as the residual covariance between item *i* loading on factor *k* and item *j* loading on Factor l. For the population model with RC1, we generated data with a three-factor structure without any residual correlations ($\psi_{ij,k} = \psi_{ij,kl} = 0$). For models with 3 indicators per factor ($J_1 = J_2 = J_3 = 3$) or 6 indicators per factor ($J_1 = J_2 = J_3 = 6$), the residual covariance matrix included $\psi_{23,1} = \psi_{56,2} = \psi_{89,3} = .3$ for RC2, and $\psi_{26,12} = \psi_{59,23} = \psi_{83,31} = .3$ for RC3. In total, there were $2 \times 2 \times 2 \times 2 \times 3 = 48$ conditions.

Table 3

Design factors for Study 2

Design factor	Three-factor Model
Sample size (N)	100, 500

Design factor	Three-factor Model
Number of Items Per Dimension	3, 6
Factor loadings (λ_j)	.5, .7
Factor correlation ($\phi_{kk'}$)	.3, .7
Residual covariance matrix	RC1: one without residual correlations RC2: one with the within-factor residual correlations as .3 RC3: one with the between-factor residual correlations as .3

Data analysis

Like simulation study 1, we generated 1,000 data sets for each condition. For each data set, we fit three analysis models: (1) a simple structure three-factor model, (2) a non-regularized network model, and (3) a regularized network model. Nine scores were calculated: one factor score, four non-regularized network scores, and four regularized network scores. The relative performance of these scores was compared in terms of their correlations and RMSEs to true factor scores. The global model fitting for the three-factor models were examined.

Results

According to average values of RMSEA, 16 out of 48 conditions (33.33%) had acceptable model fit. As for CTI and TLI, 20 out of 48 conditions (41.67%) had acceptable model fit.

Figure 3 and Figure S1 (in the supplemental materials) present the average correlations between estimated scores and true factor scores across conditions for J = 6 and J = 3, respectively. Overall, network scores showed similar average correlations among different dimensions (factors). Network scores had slightly lower correlation ($\bar{r} \in [.912, 0.971]$) on average than FS ($\bar{r} \in [.976,1]$) in the conditions of moderate sample sizes (N = 500) and long test length (J = 6 per dimension). In contrast, lower factor loadings (λ = .5), smaller sample sizes (N = 100), and short test length (J = 3 per dimension) led to decreased correlations for both network scores ($\bar{r} \in [.747, .998]$) and FS ($\bar{r} \in [.771, .989]$). In addition, as factor correlations ϕ_{ij} increased, the correlations between network scores and true factor scores also increased, ranging from $\bar{r} \in [.747,1]$ to $\bar{r} \in [.799,1]$. For factor scores, it was expected that covariance matrix without residual covariances (RC1; $\dot{r} = .995, \bar{r} \in [.976, 1]$) achieved the highest correlations than within-factor residual covariances (RC2; $\dot{r} = .917, \bar{r} \in [.771, 1]$) and between-factor residual covariances (RC3; $\dot{r} = .995, \bar{r} \in [.972, 1]$). Similar to factor scores, all network scores in the conditions of within-factor residual covariances (RC2) showed lower average correlations and higher variation of correlations across conditions (e.g., for NS-S, $\dot{r} = .947, \bar{r} \in [.820, 890]$) than between-factor residual covariances (RC3; e.g., for NS-S, $\dot{r} = .966, \bar{r} \in [.890, 8934]$). However, both RC2 and RC3 performed worse than RC1 (e.g., for NS-S, $\dot{r} = .967, \bar{r} \in [.895, 938]$).

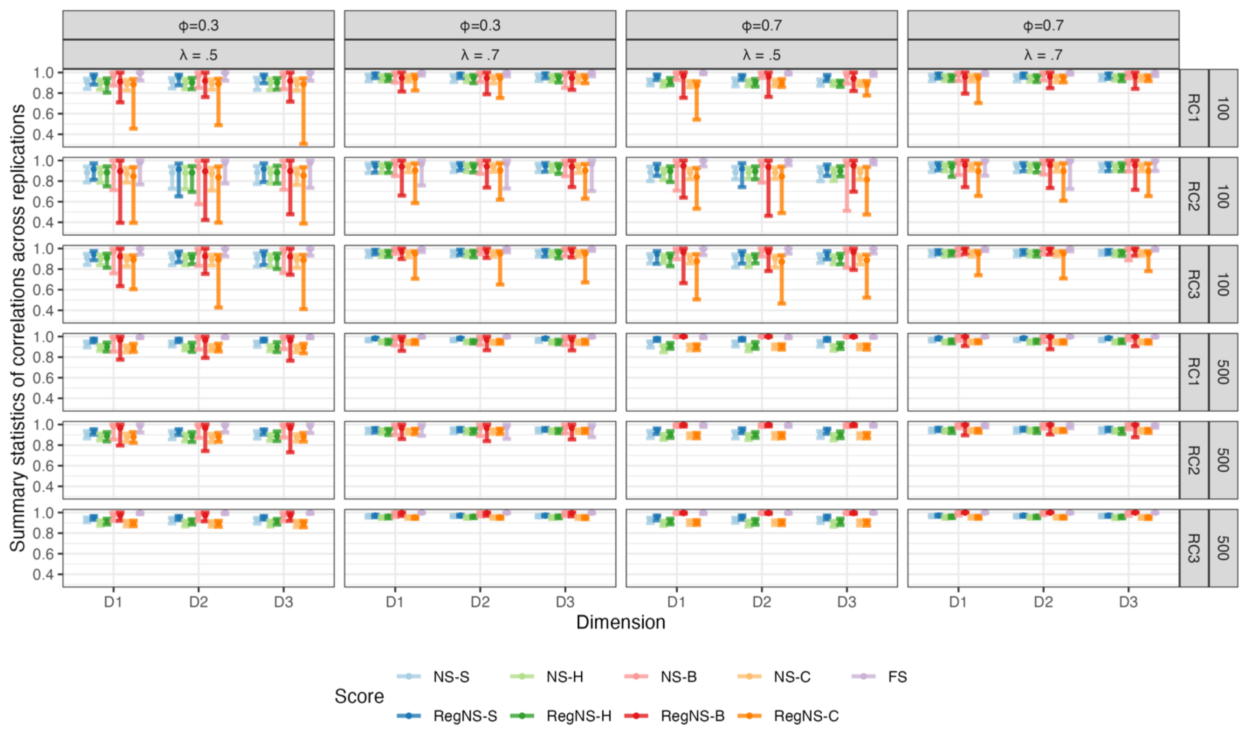
When factor correlations and factor loadings were high, network scores based on betweeness (NS-B: $\bar{r} \in [.915,1]$ and RegNS-B: $\bar{r} \in [.939,1]$) produced the highest correlation than other network scores ($\bar{r} \in [.882, .984]$) and FS ($\bar{r} \in [.891, .954]$) (see the 8th column in Figure 3). However, betweenness-based network scores had decreased correlations when sample size, factor correlations and factor loadings decreased. In addition, network scores based on strength (NS-S and RegNS-S; $\bar{r} \in [.820, .967]$) and non-regularized network scores based on hybrid measures (NS-H; $\bar{r} \in [.824, .955]$) showed highest average correlations among network scores ($\bar{r} \in [.807, .965]$) except when the population models contained high factor correlations and factor loadings (e.g., $\phi = .7$ and $\lambda = .7$). In addition, regularized network scores ($\bar{r} \in$

$[.747,1]$) exhibited greater variation in average correlations across conditions than their non-regularized NS ($\bar{r} \in [.795,1]$).

Figure 4 and Figure S2 (in the supplemental materials) present the RMSEs for the estimated factor scores and network scores relative to the population factor scores. Consistent with the results for correlations, network scores based on betweeness (NS-B and RegNS-B) performed best when factor correlations and factor loadings were relatively high, but they were highly unstable in other conditions. Overall, network scores based on strength (NS-S and RegNS-S) showed the lowest bias in the remaining conditions.

Figure 3

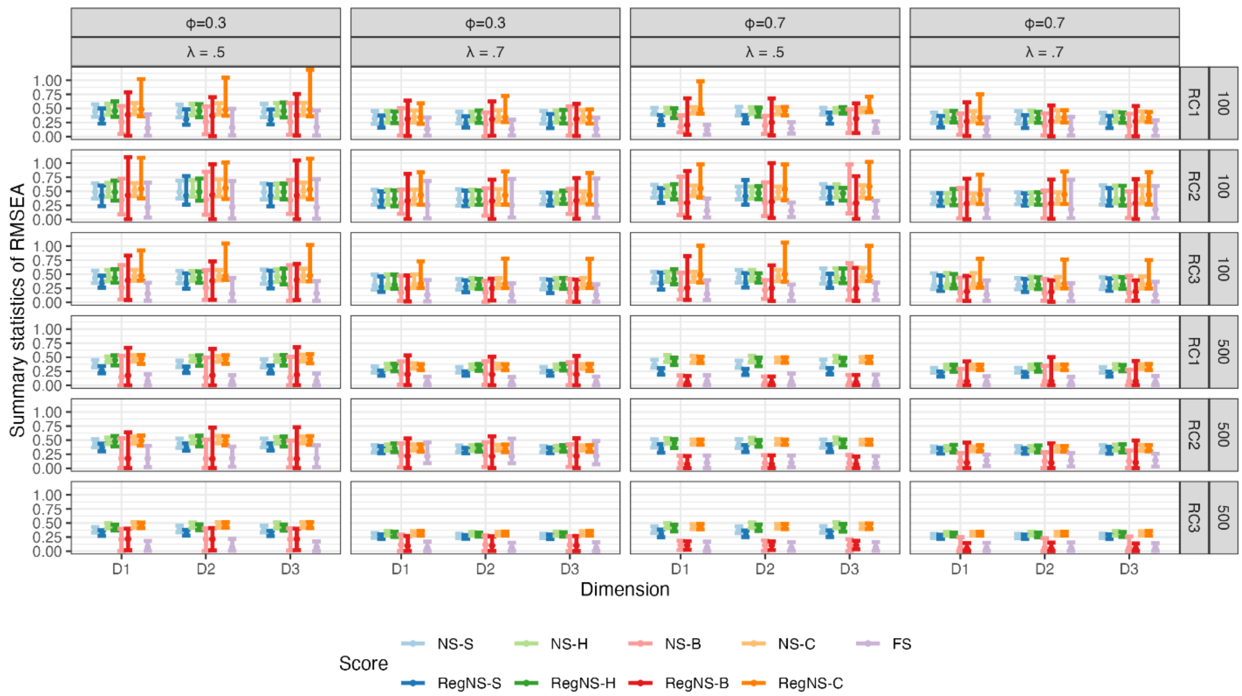
Correlations between estimated scores and population factor scores for J = 6 in Study 2



Note. ϕ denotes factor correlations of population models; λ denotes factor loadings of population models; RC1 = no residual covariances is included in population models; RC2 = within-factor residual covariances are included in population models; RC3 = between-factor residual covariances are included in population models

Figure 4

RMSEs between estimated scores and population factor scores for J = 6 in Study 2



Note. ϕ denotes factor correlations of population models; λ denotes factor loadings of population models; RC1 = no residual covariances is included in population models; RC2 = within-factor residual covariances are included in population models; RC3 = between-factor residual covariances are included in population models

Empirical Study

Method

For empirical illustration, the Big Five Personality Test (Goldberg, 1992) was utilized. Theoretically, the Big Five framework comprises five latent constructs: extraversion (E), neuroticism (N), agreeableness (A), conscientiousness (C), and openness to experience (O). The assessment instrument consists of 50 items rated on a five-point Likert scale, where 1 denotes "Disagree", 3 denotes "Neutral", and 5 denotes "Agree", with each construct represented by 10 items. Eighteen negatively worded items were reverse coded, and a random sample of 500 participants was selected for analysis. Prior to analysis, all items were standardized.

An exploratory analysis was initially conducted to determine dimensionality using both Exploratory Factor Analysis (EFA, Mulaik, 1987) and Exploratory Graphical Analysis (EGA, Epskamp, Cramer, et al., 2018). For factor analysis, parallel analysis (Horn, 1965) identified the number of factors to retain, followed by EFA to investigate the relationships between observed and latent variables. In contrast, EGA utilized a network-based approach to identify communities and clarify inter-variable relationships by estimating partial correlation networks, with communities representing potential latent dimensions. Bootstrapping methods were applied to assess the stability of the identified communities. The *psych* (Revelle, 2023) package was used for EFA, and the *EFAnet* (Epskamp, 2022) package for EGA.

Based on the identified dimensions and item-dimension relationships, we estimated three models: a simple structure CFA model (CFA_SS; an item only loads on one factor), a non-regularized network model, and a regularized network model. The nine scoring methods used in the simulation study (CFA_SS, NS-S, NS-H, NS-S, NS-C, RegNS-S, RegNS-H, RegNS-S, RegNS-C) were examined and compared with the EFA scoring method. Correlations among the ten types of scores for each dimension were analyzed, and RMSEs were calculated using the EFA scores as reference. Additionally, model fit was examined to understand how model specification affects the performance of various scoring methods.

Results

Across multiple samples we drew, parallel analysis consistently overextracted factors, while EGA reliably identified a five-community structure with minor item misalignments from their theoretical dimensions. For demonstration purposes, we selected a sample where EGA indicated a controlled level of community misspecification, with item A3 misaligning with conscientiousness and item A10 with extraversion (see Figure 5). Stability analysis showed that items A10 and A3 had only a 69% and 64% probability of being classified into conscientiousness and extraversion, respectively, in contrast to nearly all other items, which had over 99% confidence of being classified into their respective communities. With the same sample, parallel analysis suggested a seven-factor solution, and EFA indicated that neuroticism might divide into three subscales. Based on the EGA-derived communities, we fitted CFA and network models to evaluate various scoring methods, enabling to assess these methods within the workflow of the network analysis.

Figure 5

Exploratory Graph Analysis on Big Five Data

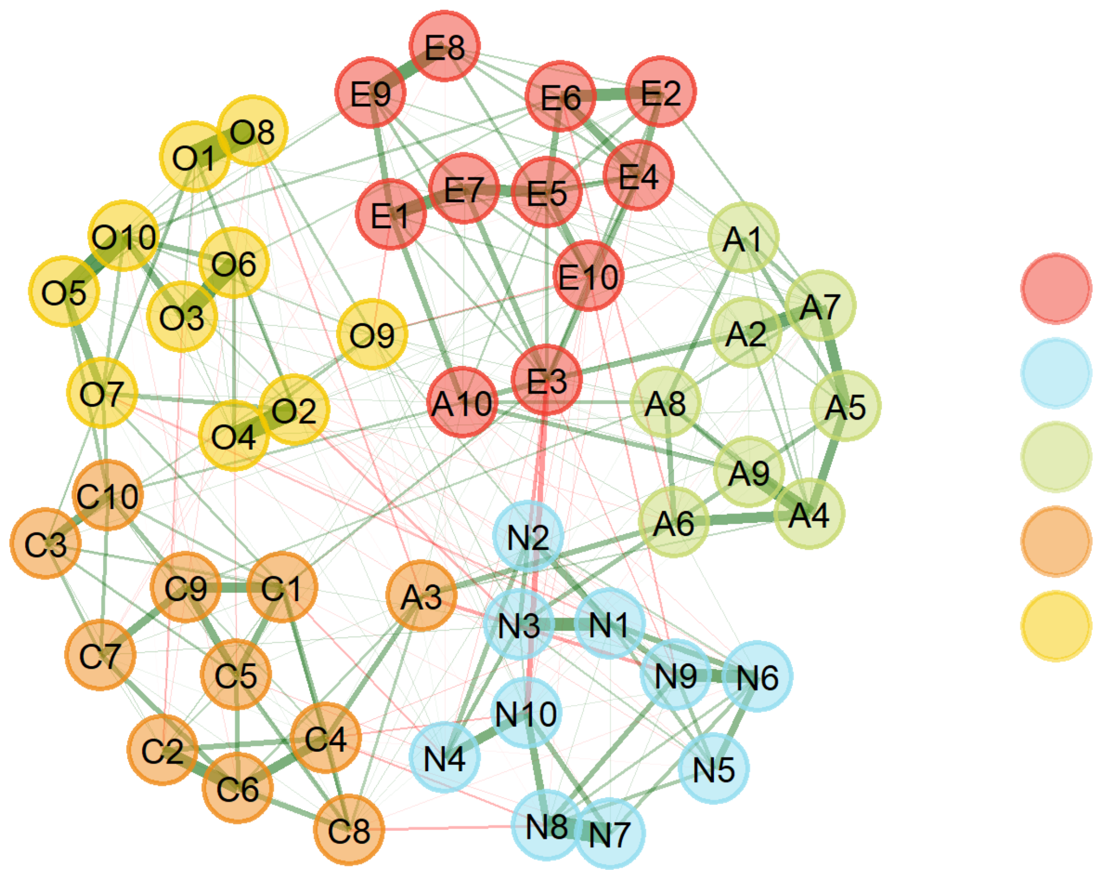


Table 4 presents the correlation ranges among ten scores across five dimensions. Overall, factor scores and network scores were highly correlated, with a minimum correlation above .94 for all scores except NS-B and RegNS-B. When betweenness centrality was used to compute network scores, the minimum correlations could drop to .91 in certain communities, such as the correlation between RegNS-B and EFA within neuroticism. In general, non-regularized network scores showed slightly stronger correlations with factor scores than regularized network scores. Network scores computed using node strength and closeness demonstrated slightly higher correlations with factor scores compared to those computed using hybrid centrality.

Table 4

Correlation Range Matrix among Scores

	EFA	CFA_SS	NS_S	NS_H	NS_C	NS_B	RegNS_S	RegNS_H	RegNS_C	RegNS_B
EFA	1									
CFA_SS	[0.97, 0.99]	1								
NS_S	[0.96, 0.99]	[0.98, 0.99]	1							
NS_H	[0.96, 0.99]	[0.98, 0.99]	[1, 1]	1						
NS_C	[0.96, 0.99]	[0.98, 0.99]	[1, 1]	[1, 1]	1					
NS_B	[0.93, 0.97]	[0.93, 0.99]	[0.95, 0.99]	[0.95, 0.99]	[0.95, 0.99]	1				
RegNS_S	[0.97, 0.99]	[0.98, 0.99]	[0.99, 1]	[1, 1]	[0.99, 1]	[0.96, 0.99]	1			
RegNS_H	[0.94, 0.98]	[0.96, 0.99]	[0.97, 0.99]	[0.97, 0.99]	[0.97, 0.99]	[0.97, 0.99]	[0.99, 1]	1		
RegNS_C	[0.95, 0.99]	[0.97, 0.99]	[1, 1]	[1, 1]	[1, 1]	[0.95, 0.99]	[0.99, 1]	[0.97, 0.99]	1	
RegNS_B	[0.91, 0.97]	[0.92, 0.97]	[0.94, 0.96]	[0.94, 0.97]	[0.93, 0.96]	[0.91, 0.98]	[0.94, 0.97]	[0.94, 0.99]	[0.94, 0.97]	1

Table 5 presents the RMSEs for nine scores relative to the EFA scores. RMSEs for network scores based on node strength, hybrid, and closeness centrality were comparable. The highest RMSEs were observed for CFA scores, followed by NS-B and RegNS-B. Notably, EFA and network scores exhibited similar score distributions, while the CFA scores displayed a more condensed distribution, which may contribute to the larger RMSEs when compared with EFA.

Model fit did not appear to noticeably impact scoring performance. The non-regularized network model showed perfect fit, while EFA, CFA, and regularized network exhibited varying degrees of misspecification. Consistently across fit measures: CFI, TLI, RMSEA, AIC, and BIC, the regularized network model fit best, followed by EFA and then CFA. Although RMSEA indicated good fit for all models (.05 - .07), CFI (.70 - .84) and TLI (.69 - .83) suggested overall poor fit. Despite these differences, scores from all models remained highly correlated.

Table 5

Root Mean Square Errors

Scoring	F1	F2	F3	F4	F5
CFA_SS	0.35	0.37	0.56	0.61	0.44
NS_S	0.17	0.28	0.24	0.18	0.20
NS_H	0.16	0.29	0.23	0.18	0.20
NS_C	0.17	0.29	0.25	0.18	0.20
NS_B	0.25	0.36	0.23	0.32	0.26
RegNS_S	0.15	0.26	0.19	0.16	0.20
RegNS_H	0.19	0.34	0.19	0.19	0.28
RegNS_C	0.17	0.31	0.24	0.18	0.20
RegNS_B	0.26	0.43	0.28	0.39	0.33

Discussion

The current study developed a general formula for network scoring, and compared the performance of various network scores derived from different centrality measures to factor scores. This study also examined the potential sources of instability arising from violations of the local independence assumption, small sample sizes, short test lengths, and regularization methods employed in network analysis.

In simulation studies, we investigated the relative performance of network scores and factor scores with the unidimensional and multidimensional factor structure. The results showed that except for betweenness-based network scores, other network scores exhibited comparable performance to factor scores regarding RMSEs and correlations, regardless of whether the data-generation models contained residual covariances. In addition, some network scores (NS-C, NS-S, and NS-H) performed better than factor scores for unidimensional models with residual covariances, which were not found in multidimensional models. In the empirical study, we further demonstrated the utility of network scores in empirical settings with the Big Five Personality Test. Consistent with results from the simulation studies, non-regularized network scores showed slightly stronger correlations with factor scores than regularized network scores. Network scores computed using node strength and closeness demonstrated the best performance in terms of correlations with factor scores. Findings imply that all network scores except those based on betweenness centrality can serve as a viable alternative scoring method to factor scores in nearly all practical conditions.

Based on our findings, we recommend using non-regularized network scores over regularized network scores. Network scores based on node strength (NS-S) are most suitable for models with moderate factors loadings, followed by network scores based on hybrid centrality (NS-H). Network scores based on betweenness (NS-B and RegNS-B) are generally not recommended, except when both factor loadings and factor correlations are extremely high. Under these conditions, the observed data likely resemble a fully connected network with high edge strengths. Therefore, betweenness centrality measures for all items might approximate factor loadings, resulting in greater accuracy for network scores. In addition, we found that all network scores had high stability and accuracy even when the population covariance matrix contained between-factor residual covariances (RC2), compared to those containing within-factor residual covariances (RC3). This finding suggests that network scores may be less impacted by residual covariances occurring between different dimensions (or communities) than factor scores, which is consistent with the previous research indicating that network scores are robust to model misspecification (Ouyang et al., 2023).

Various network scores have been proven useful in simulation and empirical studies, especially non-regularized network scores with node strength and hybrid centrality. As more and more researchers consider the construct as a dynamic system comprising of observable variables rather than measures driven by latent variables (Borsboom, 2008; Borsboom et al., 2021b), network scores may be crucial in capturing each individual's overall status within a group-level psychological network and in developing personalized interventions. Unlike CFA, network scores are highly flexible and do not require a strict structural specification. However, further investigation is needed on how to properly estimate and interpret network scores under different data structures, network structures, estimation methods, and domain contexts.

Limitations and Future Direction

The limitations of this study provide potential directions for future research. First, we investigated only four types of centrality indicators. Other types of network scores that could perform better under certain conditions warrant further investigation. Second, more empirical evidence is needed to evaluate the effectiveness and validity of network scores. For example, examining whether an individual's network score for a mental disorder is linked to his or her clinical outcome. In addition, how network scores can be applied to the higher-level network is worth investigating. Epskamp et al. (2017b) proposed a *latent network model* in which relationships among latent variables are modeled as a network. However, these methods require confirmatory factor analysis to identify latent variables. The development of network scores enables more advanced network modeling. For example, the dependency among nodes can be modeled as a network to calculate person-level measurement scores, which are then used as outcomes or predictors of other variables.

Reference

Avila, M. L., Stinson, J., Kiss, A., Brandão, L. R., Uleryk, E., & Feldman, B. M. (2015). A critical review of scoring options for clinical measurement tools. *BMC Research Notes*, *8*(1), 612. https://doi.org/10.1186/s13104-015-1561-6

Borkulo, C. van, Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of symptom network structure with the course of depression. *JAMA Psychiatry*, *72*(12), 1219–1226. https://doi.org/10.1001/jamapsychiatry.2015.2079

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*(9), 1089–1108. https://doi.org/10.1002/jclp.20503

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., Borkulo, C. D. van, Bork, R. van, & Waldorp, L. J. (2021a). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., Borkulo, C. D. van, Bork, R. van, & Waldorp, L. J. (2021b). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Christensen, A. P. (2018). NetworkToolbox: Methods and measures for brain, cognitive, and psychometric network analysis in r. *R J.*, *10*(2), 422. https://files.osf.io/v1/resources/6kmav/providers/osfstorage/5aec6ad0b80ad4001054a820?action=download&version=1&displayName=NetworkToolbox_Preprint_psyarxiv-2018-05-04T14:14:40.992Z.pdf&direct

Christensen, A. P., & Golino, H. (2021). On the equivalency of factor and network loadings. *Behavior Research Methods*, *53*(4), 1563–1580. https://doi.org/10.3758/s13428-020-01500-6

Christensen, A. P., Kenett, Y. N., Aste, T., Silvia, P. J., & Kwapil, T. R. (2018). Network structure of the Wisconsin Schizotypy ScalesShort Forms: Examining psychometric network filtering approaches. *Behavior Research Methods*, *50*(6), 2531–2550. https://doi.org/10.3758/s13428-018-1032-9

Cramer, A. O. J., Waldorp, L. J., Maas, H. L. J. van der, & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–150. https://doi.org/10.1017/S0140525X09991567

Curran, P. J., Cole, V., Bauer, D. J., Hussong, A. M., & Gottfredson, N. (2016). Improving factor score estimation through the use of observed background characteristics. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(6), 827–844. https://doi.org/10.1080/10705511.2016.1220839

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, *61*(12), 1203–1218. https://doi.org/10.1016/j.jbusres.2008.01.009

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S. (2022). EGAnet: An R package for exploratory graph analysis. *Journal of Statistical Software*, *101*(1), 1–33. https://doi.org/10.18637/jss.v101.i01

Epskamp, S. (2023). *Psychonetrics: Structural equation modeling and confirmatory network analysis*. https://cran.r-project.org/web/packages/psychonetrics/index.html

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. (2018). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological data. *PLOS ONE*, *13*(2), e0193169. https://doi.org/10.1371/journal.pone.0193169

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

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017b). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, *82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017a). Generalized Network Psychometrics: Combining Network and Latent Variable Models. *Psychometrika*, *82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018a). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018b). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Epskamp, S., Isvoranu, A.-M., & Cheung, M. (2020). Meta-analytic Gaussian network aggregation. Retrieved from https://psyarxiv.com/236w8/

Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, *78*(5), 762–780. https://doi.org/10.1177/0013164417719308

Fried, E. I., Borkulo, C. D. van, Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*(1), 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Hallquist, M. N., Wright, A. G. C., & Molenaar, P. C. M. (2021). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research*, *56*(2), 199–223. https://doi.org/10.1080/00273171.2019.1640103

He, J., Kang, J., Sun, S., Cooper, M., Zickgraf, H. F., & Zhai, Y. (2022). The landscape of eating disorders research: A 40-year bibliometric analysis. *European Eating Disorders Review*, *30*(2), 96–109. https://doi.org/10.1002/erv.2884

Hofmann, S. G., Curtiss, J., & McNally, R. J. (2016). A Complex Network Perspective on Clinical Science. *Perspectives on Psychological Science*, *11*(5), 597–605. https://doi.org/10.1177/1745691616639283

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289095

Isvoranu, A.-M., & Epskamp, S. (2023). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods*, *28*(4), 925–946. https://doi.org/10.1037/met0000439

(Li & Zhang, 2024)Levinson, C. A., Vanzhula, I., & Brosof, L. C. (2018). Longitudinal and personalized networks of eating disorder cognitions and behaviors: Targets for precision intervention a proof of concept study. *International Journal of Eating Disorders*, *51*(11), 1233–1243. https://doi.org/10.1002/eat.22952

Li, J., & Zhang, J. (2024). Personality traits and depressive symptoms among chinese older people: A network approach. *Journal of Affective Disorders*, *351*, 74–81. https://doi.org/10.1016/j.jad.2024.01.215

Liang, G., Cheng, Y., Barnhart, W. R., Song, J., Lu, T., & He, J. (2023). A network analysis of disordered eating symptoms, big-five personality traits, and psychological distress in Chinese adults. *International Journal of Eating Disorders*, *n/a*(n/a), 1–12. https://doi.org/10.1002/eat.24012

McElroy, E., McIntyre, J. C., Bentall, R. P., Wilson, T., Holt, K., Kullu, C., Nathan, R., Kerr, A., Panagaki, K., McKeown, M., Saini, P., Gabbay, M., & Corcoran, R. (2019). Mental health, deprivation, and the neighborhood social environment: A network analysis. *Clinical Psychological Science*, *7*(4), 719–734. https://doi.org/10.1177/2167702619830640

Monteleone, A. M., & Cascino, G. (2021). A systematic review of network analysis studies in eating disorders: Is time to broaden the core psychopathology to non specific symptoms. *European Eating Disorders Review*, *29*(4), 531–547. https://doi.org/10.1002/erv.2834

Monteleone, A. M., Cascino, G., Salerno, L., Albano, G., Barone, E., Cardi, V., & Lo Coco, G. (2023). The interplay between emotion regulation, interpersonal problems and eating symptoms in individuals with obesity: A network analysis study. *Journal of Affective Disorders*, *324*, 61–67. https://doi.org/10.1016/j.jad.2022.12.056

Mulaik, S. A. (1987). A brief history of the philosophical foundations of exploratory factor analysis. *Multivariate Behavioral Research*, *22*, 267–305. https://doi.org/10.1207/s15327906mbr2202_4

Ouyang, J., Jiang, Z., DiStefano, C., Pan, J., Han, Y., Xu, L., Shi, D., & Cai, F. (2023). Comparing methods for factor score estimation in structural equation modeling: The role of network analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *0*(0), 119. https://doi.org/10.1080/10705511.2023.2253496

Revelle, W. (2023). *Psych: Procedures for psychological, psychometric, and personality research*. https://CRAN.R-project.org/package=psych

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Stella, M. (2022). Network psychometrics and cognitive network science open new ways for understanding math anxiety as a complex system. *Journal of Complex Networks*, *10*(3), cnac012. https://doi.org/10.1093/comnet/cnac012

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, *41*(1), 1–32. https://doi.org/10.1037/h0075959

Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., Mors, O., Elkin, A., Williamson, R. J., Schmael, C., & al., et. (2008). Measuring depression: Comparison and integration of three scales in the GENDEP study. *Psychological Medicine*, *38*(2), 289–300. https://doi.org/10.1017/S0033291707001730

van Bork, R., Lunansky, G., & Borsboom, D. (2024). Measurement targets for network constructs in psychopathology. *Measurement*, *232*, 114643. https://doi.org/10.1016/j.measurement.2024.114643

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861. https://doi.org/10.1037/0033-295X.113.4.842

Zhao, Y., Qu, D., Chen, S., & Chi, X. (2023). Network analysis of internet addiction and depression among chinese college students during the COVID-19 pandemic: A longitudinal study. *Computers in Human Behavior*, *138*, 107424. https://doi.org/10.1016/j.chb.2022.107424