Multivariate Behavioral Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

Unique Variable Analysis: A Network Psychometrics Method to Detect Local Dependence

Alexander P. Christensen, Luis Eduardo Garrido & Hudson Golino

To cite this article: Alexander P. Christensen, Luis Eduardo Garrido & Hudson Golino (2023) Unique Variable Analysis: A Network Psychometrics Method to Detect Local Dependence, Multivariate Behavioral Research, 58:6, 1165-1182, DOI: 10.1080/00273171.2023.2194606

To link to this article: https://doi.org/10.1080/00273171.2023.2194606

MULTIVARIATE BEHAVIORAL RESEARCH
2023, VOL. 58, NO. 6, 1165–1182
https://doi.org/10.1080/00273171.2023.2194606

Unique Variable Analysis: A Network Psychometrics Method to Detect Local Dependence

Alexander P. Christensen[a], Luis Eduardo Garrido[b], and Hudson Golino[c]

[a]Department of Psychology and Human Development, Vanderbilt University; [b]Escuela de Psicología, Pontificia Universidad Católica Madre y Maestra; [c]Psychology, University of Virginia

ABSTRACT

The local independence assumption states that variables are unrelated after conditioning on a latent variable. Common problems that arise from violations of this assumption include model misspecification, biased model parameters, and inaccurate estimates of internal structure. These problems are not limited to latent variable models but also apply to network psychometrics. This paper proposes a novel network psychometric approach to detect locally dependent pairs of variables using network modeling and a graph theory measure called weighted topological overlap (wTO). Using simulation, this approach is compared to contemporary local dependence detection methods such as exploratory structural equation modeling with standardized expected parameter change and a recently developed approach using partial correlations and a resampling procedure. Different approaches to determine local dependence using statistical significance and cutoff values are also compared. Continuous, polytomous (5-point Likert scale), and dichotomous (binary) data were generated with skew across a variety of conditions. Our results indicate that cutoff values work better than significance approaches. Overall, the network psychometrics approaches using wTO with graphical least absolute shrinkage and selector operator with extended Bayesian information criterion and wTO with Bayesian Gaussian graphical model were the best performing local dependence detection methods overall.

KEYWORDS
Local dependence; correlated residuals; minor factors

The local independence assumption states that variables are unrelated after conditioning on a latent variable. Violations of local independence can arise in different ways (Leising et al., 2020). Shared semantic references, such as similar item phrasing or item content, can create similar response patterns that lead to conditional dependence between items (Reise et al., 2018; Rosenbusch et al., 2020). Shared substantive causes, such as a common underlying attribute (e.g., social desirability), can lead to undesired dimensions (e.g., minor factors; Leising et al., 2020). Common scale development conventions such as selecting variables based on high item–test correlations and contribution to the test's internal consistency (e.g., Cronbach's α; DeVellis, 2017) may also contribute to local independence violations (Hubley et al., 2014).

Issues caused by violations of local independence can range from relatively minor to severe. Common problems that arise from violations include model misspecification (Montoya & Edwards, 2020), biased model parameters (Edwards et al., 2018), and inaccurate estimates of internal structure (Wood et al., 1996). Because of the potential severity of these effects, different approaches to detect local dependence have been developed. These methods aim to identify local independence violations when they arise so that researchers can determine their causes and handle their effects. This paper aims to introduce another approach, based on network psychometrics, to detect local independence violations.

Network psychometrics applies network science and graph theory methods to model psychological constructs (Epskamp et al., 2018). Network models depict variables as nodes (circles) and their relationships as edges (lines). A common approach to model relationships between variables is to obtain their partial correlations conditioned on all other variables (Epskamp & Fried, 2018). This statistical representation is proposed to suggest that network models do not assume latent variables explicitly (Borsboom et al.,

CONTACT Alexander P. Christensen alexpaulchristensen@gmail.com Department of Psychology and Human Development, Vanderbilt University, Nashville, TN 37203, USA.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00273171.2023.2194606.

2021; Guttman, 1953). Nevertheless, they are useful to identify latent factors (Golino et al., 2020; Golino & Epskamp, 2017; Jiménez et al., 2022), derive model parameters that are statistically similar to latent variable models (e.g., loadings; Christensen & Golino, 2021), and test group equivalence (i.e., measurement invariance; Jamison et al., 2022).

Although local independence is a property of latent variable models, network models still face many of the same issues related to strongly overlapping components (e.g., biased parameter estimates; Fried & Cramer, 2017; Hallquist et al., 2021). Network theory proposes that nodes in a network represent causally autonomous components (e.g., symptoms) of a system (e.g., psychopathology; Borsboom, 2017; Cramer et al., 2012). The causal autonomy of a component suggests that two or more components are not strongly related because they are determined by the same cause but instead because they have unique causal dependence (Christensen et al., 2020; Cramer, 2012). Because network models do not assume latent variables (Guttman, 1953), strongly overlapping components may instead represent redundant components of the system (Marinazzo et al., 2022). For the purposes of this paper, we use the latent variable terminology of local dependence because our simulated data are generated using a latent variable model; however, for network models, this dependence is conceptually consistent with redundant components in a system. Regardless of statistical framework, identifying and handling strongly overlapping components in psychometric networks requires careful consideration (Christensen et al., 2020; Fried & Cramer, 2017; Hallquist et al., 2021).

This paper aims to evaluate contemporary approaches to detect local dependence, including a novel network psychometrics approach, so that researchers can make informed decisions about which approaches might work best for their data. The paper is organized as follows: First, network psychometrics is described and a novel local dependence detection method based on this perspective is introduced. Second, contemporary factor analytic approaches to detect local dependence violations are reviewed. Although there is a rich tradition of identifying violations of local independence in the item response theory literature (Chen & Thissen, 1997; Edwards et al., 2018; Yen, 1984), this simulation study seeks to evaluate approaches across continuous and categorical (polytomous and dichotomous) data. For each of these methods, different approaches to determine which pairs of variables should be considered locally dependent, such as statistical significance and cutoffs, are discussed. Third, we argue that these contemporary methods must be reevaluated because of limitations in the metrics that have been used in nearly all previous recommendations. Finally, we outline the aims of the simulation which seeks to (1) evaluate contemporary local dependence detection methods across continuous, polytomous (5-point Likert scale), and dichotomous data with skew, (2) compare the novel network psychometrics approach introduced in this paper to these contemporary methods, and (3) identify and determine whether there are cutoffs values that may be useful to the applied researcher.

A network psychometrics method to detect local dependence

The standard approach to model psychological constructs as a network is to estimate a Gaussian graphical model (GGM; Borsboom et al., 2021; Lauritzen, 1996). GGMs represent variables as nodes and the relationships between them as conditional associations. For continuous data, a common approach is to apply the graphical least absolute shrinkage and selection operator (GLASSO; Friedman et al., 2008; Friedman et al., 2014). The GLASSO regularizes the inverse covariance matrix using an ℓ_1-norm penalization on the log-likelihood resulting in conditional relationships between variables that have been reduced with some reducing to zero. The purpose of the regularization is to reduce overfitting when there are more variables than there are cases (Friedman et al., 2008; but see Williams & Rodriguez, 2022) as well as to induce a sparse partial correlation matrix (i.e., reduce small conditional relationships between variables to zero; Epskamp et al., 2017). In psychology, it has become standard to use the extended Bayesian information criterion (EBIC; Chen & Chen, 2008) to select from a range of models estimated with different sparsity (often referred to as EBICglasso; Epskamp & Fried, 2018; Foygel & Drton, 2010). The EBICglasso network estimation method can be extended to polytomous and dichotomous data by computing polychoric and tetrachoric correlations, respectively (as well as Spearman correlation; Isvoranu & Epskamp, 2021).[1]

Despite the popularity of the EBICglasso method to estimate a GGM, there are other approaches. One approach uses maximum likelihood estimation to

[1]In the psychometric network literature, it is more common to use the Ising model for dichotomous data (van Borkulo et al., 2014); however, to demonstrate the generalizability of our approach to detect local dependence, we use the EBICglasso with tetrachoric correlations rather than the Ising model.

estimate networks comprised of non-regularized partial correlations (Williams et al., 2019; Williams & Rast, 2018). Another approach uses Bayesian inference using analytic or posterior sampling to obtain parameter estimates for each partial correlation (Williams, 2021; Williams & Mulder, 2020). The analytic approach uses the maximum a posteriori estimate of the partial correlations (Williams, 2021). In both cases, a credible interval is necessary to determine whether partial correlations are retained or set to zero (most often, the 95% credible interval is used). The Bayesian GGM (BGGM) approach in particular has become popular in the applied literature (e.g., Briganti et al., 2022). The goal of these methods, like EBICglasso, is to identify a sparse network model that adequately captures the conditional relationships between variables.

The sparsity (i.e., proportion of zeros) induced in the partial correlation matrix is what characterizes a network model. As a network, statistical measures from graph theory can be applied to quantify the network's structural and topological features. Centrality measures, which quantify a node's relative position in the network, are the most commonly used measures in psychology (e.g., *node strength* or the sum of a node's connections; Bringmann et al., 2019). Psychometric network modeling has rarely ventured beyond centrality measures despite the swath of other measures that exist (Letina et al., 2019; Rubinov & Sporns, 2010), and many of which may be useful for different psychometric procedures that are common in traditional psychometrics such as local dependence detection (Fried & Cramer, 2017).

One such measure is weighted topological overlap (wTO; Zhang & Horvath, 2005). wTO quantifies the extent to which nodes in a network "overlap" by computing the similarity between a pair of nodes' shared connections (e.g., edge weights, signs, quantity). In biological networks, this measure has been used to identify shared genetic expression of proteins (Nowick et al., 2009) and hierarchical organization of metabolic pathways (Ravasz et al., 2002). In psychological networks, wTO can identify two or more variables that are highly related and have roughly the same relations (sign and size of relation) to other variables (three or more may represent a latent factor; Golino & Epskamp, 2017; Golino et al., 2020). Two variables that have high overlap, especially above and beyond other variables, would instead represent conditional information that is roughly redundant (Christensen et al., 2020). In other words, variables that have substantially high overlap are conceptually consistent with variables that are locally dependent (Fried & Cramer, 2017). Determining what suffices as "substantially high," however, requires further investigation.

The theoretical notion of using partial correlation networks and wTO for local dependence detection starts with partial correlations. Consider a unidimensional latent variable model where the local independence assumption holds—that is, after accounting for the latent variable, variables are unrelated to each other. Partial correlations represent the remaining covariance between two variables after they are conditioned on all other variables. The unique covariance between any two variables represents their shared covariance that is not explained by any other variables. Given a unidimensional latent variable model, this unique covariance will tend to zero as the number of variables increases to infinity (Guttman, 1953; Waldorp & Marsman, 2021). Larger partial correlations, however, represent unique pairwise dependence between two variables that exceeds what is expected if the assumed unidimensional latent variable model is true. Unique dependence that exceeds what might be expected from sampling error are the basis for the anti-image partial correlation approach that we'll discuss in the next section (Ferrando et al., 2022).

Our proposed method extends this partial correlation approach by using regularization, a technique often employed in network psychometrics. (G)LASSO regularization shrinks partial correlation coefficients, setting some to zero. A benefit of shrinking coefficients is that it potentially avoids an upward bias of partial correlation coefficients when there are few variables per factor (McDonald, 1985). As Ferrando et al. (2022) find, partial correlations tend to overestimate the true values in these conditions, resulting in the detection of more locally dependent pairs than there really are.

The wTO measure extends the estimate of similarity of two variables beyond their pairwise partial correlation. wTO quantifies the extent to which two variables have similar partial correlations to other variables in a sparse matrix (i.e., network). A potential advantage of a sparse matrix relative to the full partial correlation matrix is that variables that have high partial correlations but do not share similar non-zero values to other variables (i.e., unique predictive utility for different variables) will have lower wTO values and are therefore less likely to be considered locally dependent relative to the full matrix where all elements of the matrix are non-zero (leading to higher wTO values).

Contemporary methods to detect local dependence

Nearly all contemporary methods used to detect local dependence start by fitting a latent variable model (e.g., factor analysis, item response theory, structural equation model) to the data. The model used in these methods tends to be data specific. Categorical data (e.g., dichotomous and polytomous) tend to be modeled using item response theory (Chen & Thissen, 1997; Edwards et al., 2018) whereas continuous data tend to be modeled using factor analysis (e.g., Ferrando et al., 2022; Whittaker, 2012). Factor analytic methods can generalize to categorical data using polychoric (ordinal) and tetrachoric (binary) correlations (Ferrando et al., 2022) while structural equation modeling can use different estimators such as robust maximum likelihood (MLR) for continuous data and weighted least squares means and variances (WLSMV) for categorical data.[2]

Exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) is a relatively recent addition to the structural equation modeling tradition. A benefit of ESEM relative to other approaches such as confirmatory factor analysis is that only the number of dimensions need to be specified to estimate the model. This flexibility allows researchers to perform local dependence detection in more exploratory settings. Modification indices are often used in ESEM to determine parameters that can be included in the model based on improved fit. While these indices are a general metric for any parameter that can improve fit, they can represent violations of local dependence when evaluated in the residual correlation matrix. A correlated residual that significantly improves fit can be used as a potential indicator of local dependence.

Significance of modification indices are based on change in χ^2 with one degree of freedom (i.e., $\Delta\chi_2 \geq 3.841$). A limitation of the significance approach is that it is not only sensitive to the size of the model misspecification but also other characteristics of the model (Saris et al., 1987). One solution that appears to overcome this issue is to estimate the standardized expected parameter change (SEPC), which provides a direct estimate of the size of the misspecification (Saris et al., 2009). Some evidence suggests that a SEPC of 0.20 is substantial enough to warrant model modification for any fixed parameter in the model (Saris et al., 2009; Whittaker, 2012). Whether significance of correlated residual modification indices are sufficient to detect local dependence or a cutoff value of 0.20 in SEPC is optimal to detect local dependence remain open questions.

Ferrando et al. (2022) recently conducted a simulation that compared several local dependence detection methods from the factor analytic perspective. The first method was from traditional factor analysis of the correlations between the fitted residuals. The second method used anti-image partial correlations (Guttman, 1953) or the full non-regularized partial correlation matrix. This second approach directly aligns with our proposed method but does not induce sparsity (or regularization) and does not apply the graph theoretic measure of wTO. The third method mirrored SEPC but used exploratory factor analysis: changes in residual correlation (**E**xpected **RE**sidual correlation direct **C**hange; EREC) when freely estimated. Another method, following similar lines of EREC, estimated the expected change in communality (**E**xpected commu**N**al**I**ty **D**ir**E**ct change index; ENIDE).

Their studied contributed a novel procedure to determine statistical significance for which correlated residuals, partial correlations, or EREC values were substantial enough to signal local dependence. This procedure followed the parallel analysis approach of estimating the empirical parameters and comparing them to a sampling distribution where data are generated by shuffling the values of each variable. This procedure proceeds iteratively for 500 times with each method being applied to the resampled data. ENIDE follows a modified version of this procedure that we don't discuss here but instead recommend interested readers to Ferrando et al. (2022). Afterwards, the mean of the mean or mean of the 95th percentile of the values obtained from the sampling distribution of parameters are obtained, representing the cutoff value to determine statistically meaningful parameter deviations. The statistical basis for this procedure is to determine what parameter values are greater than would be expected from sampling error.

In this study, we focus on the partial correlation method given its direct relation to our proposed method. It's important to note that the ESEM with SEPC differs only slightly from EREC in the latent variable model used (ESEM and exploratory factor analysis, respectively) and procedure used to determine significance (modification indices and resampling procedure, respectively). Because of the similarity of these approaches, we expected that ESEM

[2]Categorical factor analysis and structural equation modeling are alternative parameterizations of item response theory models (Muraki & Carlson, 1995). The main difference in these models is the estimation procedures rather than the models themselves (limited information estimation procedures for the former; full-information for the latter). We thank the anonymous reviewer for pointing out these similarities.

with SEPC can be considered a rough approximation of how the EREC approach would perform.

In Ferrando et al. (2022) study, the partial correlation method performed comparably well to the EREC method and better than the fitted residuals and ENIDE methods based on sensitivity and specificity. In terms of the mean and 95th percentile criteria, we opted to use the 95th percentile criterion because it (1) tends to be more conservative (generally produces higher cutoff values) and (2) aligns most closely to traditional null hypothesis significance testing standards. We note that the mean criterion was comparable in sensitivity and specificity to the 95th percentile criterion in their study. Similar to the ESEM with SEPC method, we also sought to evaluate whether this statistical approach to determine a cutoff was sufficient or whether a single value cutoff would perform better.

Limitations of previous evaluation metrics

Previous simulation studies, such as Ferrando et al. (2022), have used sensitivity (power) and specificity (1 - type I error) to evaluate local dependence detection methods (Chen & Thissen, 1997; Edwards et al., 2018; Houts & Edwards, 2013; Yen, 1984). Sensitivity reflects the method's ability to detect a locally dependent pair of variables when they are truly locally dependent, and specificity reflects the method's ability to detect a non-locally dependent pair of variables when they are truly not locally dependent (Ferrando et al., 2022). To our knowledge, nearly all simulation studies conducted on local dependence detection methods have used sensitivity and specificity as their primary metrics to recommend a method. Although sensitivity and specificity are intuitive and commonly used metrics, they are poor metrics to evaluate the effectiveness of a local dependence detection method.

To see why, we first define the confusion matrix in the context of a local dependence simulation. A true positive (TP) is when a pair of variables are simulated and estimated to be locally dependent. A true negative (TN) is when a pair of variables are simulated and estimated to be not locally dependent. A false positive (FP) is when a pair of variables are estimated to be locally dependent but are simulated as not locally dependent. Conversely, a false negative (FN) is when a pair of variables are simulated as locally dependent but are estimated to be not locally dependent. Sensitivity (power) is defined as TPs divided by the sum of TPs and FNs ($\frac{TP}{TP+FN}$); specificity (1 - type I error) is defined as TNs divided by the sum of TNs and FPs ($\frac{TN}{TN+FP}$).

With these metrics defined, we move on to a hypothetical example. Assume a simulated dataset from a unidimensional factor model with 20 variables and 1 locally dependent pair of variables. Each pair of variables could be locally dependent, so there are a total of 190 possible pairs (i.e., $\frac{20\times(20-1)}{2} = 190$). Further assume that a method estimates the pair of locally dependent variables correctly ($TP = 1$) but it also estimates 9 other pairs of variables to be locally dependent ($FP = 9$). This method therefore estimates a total of 10 sets of variables as locally dependent. The remaining 180 possible variable pairs are correctly identified as not locally dependent ($TN = 180$ and $FN = 0$).

Sensitivity of this method would be perfect (i.e., $\frac{1}{1+0} = 1.00$), specificity is also high (i.e., $\frac{180}{180+9} = 0.95$), and type I error is within null hypothesis statistical testing standards (i.e., $1 - 0.95 = 0.05$). Based on the usual sensitivity and specificity metrics, the method appears to work quite well. Consider, however, that the method is detecting 10 locally dependent pairs of variables and only 1 of them is truly locally dependent. In an applied setting, there is no way to know what pair is truly locally dependent.

Rather than relying on sensitivity and specificity, other confusion matrix metrics provide better insight into the hypothetical method's performance. False discovery rate (FDR; $\frac{FP}{FP+TP}$) quantifies the number of FPs that can be expected given the total number of locally dependent pairs estimated. In our example, FDR equals $\frac{9}{1+9} = 0.90$ or 90% of the estimated locally dependent pairs are not truly locally dependent. Another metric is the critical success index (CSI; $\frac{TP}{TP+FP+FN}$), which captures the overall performance of a method. The CSI is a direct measure of what a researcher wants to know in the context of detecting locally dependent pairs: The estimated locally dependent pairs are truly locally dependent accounting for the overestimation of locally dependent pairs (FP) and any locally dependent pairs that were not detected (FN). In our example, the CSI shows that the method's performance is abysmal: $\frac{1}{1+9+0} = 0.10$.

The same issues arise in conditions where there are no locally dependent pairs of variables. Specificity in these conditions will almost always be high because there will almost always be relatively few FPs to a large number of TNs. In these conditions, the number of FPs offers a better and more direct measure of performance. Specifically, FP provides the number of local dependencies a method tends to overestimate.

The cost of overidentification of local dependence (high FDR) relative to underidentification local dependence (low sensitivity) should be considered by the researcher. On the one hand, identifying many locally dependent pairs can add many additional parameters (e.g., paths in a structural equation model), leading to an overly complex and more unstable model; on the other hand, not identifying the locally dependent pairs that exist can lead to biased parameter estimates and minor factors (Edwards et al., 2018; Wood et al., 1996). In our view, the underidentification of local dependence is more costly and therefore detecting local dependence should be given priority.

Simulation study aims

There were three aims for our simulation study. The first aim was to compare our proposed network psychometrics approach against contemporary methods to detect local dependence. Given there are many methods to estimate networks in psychology, we investigated two of the more commonly applied methods: EBICglasso and BGGM. Using these network structures, we applied the wTO measure to their networks. To determine what suffices as "substantially high" values of wTO, we investigated several cutoff values and selected the best cutoff values based on CSI and FDR to use in the comparison with other methods.

The second aim was to evaluate whether statistical significance approaches for ESEM with SEPC and partial correlations performed better than single cutoff values. For ESEM with SEPC, the cutoff value of 0.20 has been suggested as a generic cutoff for any model parameter (Saris et al., 2009); however, this value, to our knowledge, has not been tested specifically on the correlated residual parameters of the model (i.e., violations of local independence). For partial correlations, Ferrando et al. (2022) put forward a procedure to determine statistically meaningful cutoffs based on sampling error that will change with each sample while others have put forward standard errors to determine statistically significant cutoffs (Mulaik, 2010). In our study, we aimed to test whether the recently proposed resampling procedure to determine a cutoff (Ferrando et al., 2022) would perform better than a single, unchanging cutoff value.

The third and final aim was to evaluate the generalizability of these methods across different data types—continuous, polytomous, and dichotomous data. All methods tested in this study are generally used on continuous data but have the capability, using different estimators (ESEM) and correlations (EBICglasso and partial correlation), to generalize to categorical data. Importantly, we generated data with skew across all data types. To our knowledge, skew has not been added to data generated in simulations evaluating local dependence detection methods despite its prevalence in real-world data. To evaluate our methods, we used FPs for conditions without local dependence and CSI and FDR for conditions with local dependence. To be consistent with previous studies, we also computed sensitivity and specificity for conditions with local dependence. Finally, we demonstrate how these methods perform in empirical data by providing a brief empirical example using the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) dataset collected as a part of the Simons Foundation Autism Research Initiative's Simplex Collection (https://www.sfari.org/).

Simulation

Design

The parameters selected for data generation in this simulation were chosen to represent empirical data that follow conventional psychometric standards of factor models (Comrey & Lee, 2013). Very small (250), small (500), and medium (1000) sample sizes were generated. The population models generated continuous data with factor loadings randomly drawn from a uniform distribution with values between 0.40 and 0.70. Cross-loadings were also randomly drawn from a uniform distribution with values between −0.10 and 0.10. This procedure follows previous simulation work described in García-Garzón et al. (2019). One, two, three, and four factors with six and twelve variables per factor were simulated. Correlations between factors were manipulated to be small (0.25) and moderate (0.50). Six and twelve variables per factor were chosen to generate locally dependent pairs that were proportionate to the number of variables per factor. Specifically, proportions of 0.000, 0.167, and 0.333 were used, which corresponded to 0 and 0, 1 and 2, and 2 and 4 locally dependent pairs of variables per factor, respectively. There were three conditions for the size of the correlated residuals that were added to the population matrix: $0.25(\pm.05)$, $0.35(\pm.05)$, and $0.45(\pm.05)$. These values were randomly drawn from a uniform distribution. The mid-point of the correlated residual ranges are used hereafter for brevity ($0.25 = 0.20$–0.30, $0.35 = 0.30$–0.40, $0.45 = 0.40$–0.50). Skew was randomly drawn from a uniform distribution between −1 and 1 for each variable in the continuous data and was then applied to the corresponding variables in the polytomous and dichotomous data. Skewness of locally dependent variables were made to have the same polarity (i.e., negative or positive skew for both variables), which is most common in empirical data. Polytomous (5-point Likert scale) and dichotomous (binary) data were generated following the procedure in Garrido et al. (2011). All data were generated using the {latentFactoR} package (version 0.0.5; Christensen et al., 2022) in R.

The simulation design allowed for a mixed factorial design: $3 \times 2 \times 2 \times 2 \times 3 \times 3 \times 3$ (number of factors $\times$ variables per factor $\times$ proportion of locally dependent pairs $\times$ correlations between factors $\times$ sample size) resulting in 216 full simulated condition combinations. For the one factor condition, only the 0.25 correlation between factors was generated to avoid oversampling one factor conditions relative to the other conditions. Similarly, for the no local dependence condition, only the 0.25 correlated residual condition was generated to avoid oversampling no local dependence conditions relative to other conditions. These additional conditions brought the total to 294 conditions. For each condition, 100 samples were generated, totaling 29,400 samples across conditions. Finally, all conditions were used for each data type (i.e., continuous, polytomous, and dichotomous) for a final total of 882 conditions and 88,200 samples across conditions. Details about data generation can be found in our Supplementary Information (SI 1).

Local dependency detection methods

Standardized expected parameter change

The SEPC was estimated using ESEM in the *lavaan* package (version 0.6.12; Rosseel, 2012) in R. The population number of factors were used as the number of factors to be estimated with the ESEM. For continuous data, the model was estimated using MLR; for polytomous and dichotomous data, the model was estimated using WLSMV. For the significance approach, correlated residuals with modification indices that were significant based on an alpha of 0.05 were considered to be locally dependent.

We define first the expected parameter change (EPC), which represents the correlated residual when constrained to zero versus estimated freely:

$$EPC = C_{res} - 0,$$

where C_{res} is the freely estimated correlated residual. The modification index is defined by (Saris et al., 2009):

$$MI = (EPC/\sigma)^2,$$

where σ is the standard error of EPC. The confidence interval, and subsequently, significance is defined by a given alpha value (e.g., 0.05):

$$EPC - 1.96\sigma < \theta < EPC + 1.96\sigma.$$

Standardization of the EPC uses the standard deviations of the endogenous and exogenous constructs associated with the parameter of interest (Kaplan, 1989). The SEPC values were obtained using the `modindices` function in the {lavaan} package with the output "sepc.lv", which standardizes the latent (exogenous) variables only. This method is referred to as "SEPC with significance" hereafter.

Partial correlations

Anti-image partial correlations were computed following Mulaik (2010) and Ferrando et al. (2022):

$$\begin{aligned} S^2 &= [diag(R^{-1})]^{-1}, \\ Q &= SR^{-1}S, \\ P &= 2I - Q, \end{aligned}$$

where R is the correlation matrix, Q is the anti-image correlation matrix, and P is the anti-image partial correlation matrix. To generalize correlations to each data type, Pearson's correlation for continuous data, polytomous correlation for polytomous data, and tetrachoric correlation for dichotomous data. To compute the polychoric and tetrachoric correlations, we used the `PolychoricRM` function in the {Turbofuns} package (version 1.0.0; Zhang et al., 2022) in R.

Following Ferrando et al. (2022), we applied the resampling approach to derive a sampling distribution. The (grand) mean of the 95th percentile values of each pairwise partial correlation from the sampling distribution were used as the cutoff to determine which partial correlations were locally dependent. Further, Ledermann's (1937) bound was imposed to limit the maximum number of local dependencies detected. In Ferrando et al. (2022), Ledermann's bound, denoted as g, was subtracted by the number of factors, denoted as r, such that $g - r$ was the limit of local dependencies detected. The population number of factors for each simulated sample was provided to compute this limit. When more local dependencies than this limit were detected (i.e., $n > g - r$), then only up to the n^{th} largest partial correlation was used as the cutoff value. We programmed this approach in R and compared the output with the FACTOR software (version 12.3.1.0; Ferrando Piera & Lorenzo Seva, 2017) to ensure proper implementation. This method is referred to as "partial correlation with significance" hereafter.

Network analysis and weighted topological overlap

Two different network estimation methods were applied to the data: EBICglasso and BGGM. The EBICglasso method was applied to zero-order correlations computed using the same approach as the partial correlations (i.e., Pearson's for continuous, polychoric for polytomous, tetrachoric for dichotomous).[3] The `lambda.min.ratio` parameter was set to 0.10 for all networks (sets the range of the 100 λ values used for the ℓ_1-norm penalty in the GLASSO estimation). The `gamma` parameter was set to 0.50 (controls the EBIC preference for model complexity). If any variables were disconnected in the network, then the `gamma` parameter was decreased by 0.25. If any variables were disconnected with `gamma = 0.25`, then `gamma` was set to 0.00 and the resulting network was used (Golino et al., 2020).

The analytic solution of BGGM was applied to the raw data and used the default estimation settings for each data type: `"continuous"` for continuous, `"ordinal"` for polytomous, and `"binary"` for dichotomous. The BGGM approach requires the selection of a credible interval to determine the partial correlations to retain and set to zero. The standard alpha of 0.05 or credible interval of 0.95 was used in this simulation.

For both the EBICglasso and BGGM networks, the wTO measure was computed using the signed formulation with absolute values taken after. Weighted topological overlap is defined by (Gysi et al., 2018; Nowick et al., 2009):

$$\omega_{ij} = \frac{\sum_u a_{iu}a_{uj} + a_{ij}}{\min\{k_i, k_j\} + 1 - a_{ij}},$$

where a_{ij} represents the edge weight between node i and node j, $\sum_u a_{iu}a_{uj}$ represents the sum of the connections that node i and node j share, and k represents the sum of a node's connections.

Grid search for optimal cutoff values

For all local dependence methods used in this simulation, we performed a grid search to determine cutoffs that might be useful for applied researchers. Unlike statistical significance approaches, cutoffs are based on a single, unchanging parameter value. One limitation of cutoffs are that they are rigid and inflexible to changes in the data (e.g., sample size). A benefit, however, is that they are computationally efficient. For some approaches, like SEPC, a single value cutoff has been recommended over the significance testing approach (e.g., Saris et al., 2009; Whittaker, 2012).

Optimal cutoffs were determined using a grid search over different cutoff points. For each local dependence detection method, we used a range of cutoffs (from 0.15 to 0.40) in increments of 0.05 with the goal of identifying an "arc" or pattern where there was a clear increase and later decrease in performance of each method. The peak of this arc was determined to be the "optimal" cutoff because it maximized the performance of the method relative to other cutoffs. For conditions with no local dependence, we used FPs as our metric of performance; for conditions with local dependence, we used CSI as our metric of performance. This grid search is provided in the Supplemental Information (SI 2). Our grid search identified the following cutoffs for overall performance across data types: SEPC = 0.25, partial correlation = 0.35, wTO with BGGM = 0.25, and wTO with EBICglasso = 0.25. These cutoffs are used as a comparison to the significance-based approaches in the simulation. Hereafter, the methods are referred to "SEPC with cut-off", "partial correlation with cut-off", "wTO with BGGM", and "wTO with EBICglasso", respectively.

Data analysis

We used R (version 4.2.2; R Core Team, 2022) for our simulation, analyses, and the *papaja* package (version 0.1.1; Aust & Barth, 2022) for our manuscript preparation. Figures were created using *ggplot2* (version 3.4.0; Wickham, 2016) and *ggpubr* (version 0.5.0; Kassambara, 2018). All data and R scripts can be found on the Open Science Framework.

Results

The results start with the overall performance across data types and are broken down by local dependence and no local dependence conditions. After, results are presented for each data type to provide more nuanced information about which method should be favored. ANOVAs on the local dependence conditions (using CSI) and no local dependence conditions (using FP) were used to determine how each method was affected by the different conditions. Only main effects that reached at least a large effect size ($\eta_p^2 \geq 0.14$; Cohen,

[3]Recent research has demonstrated that Spearman's correlation may actually be more appropriate for ordinal data when using the EBICglasso method (Isvoranu & Epskamp, 2021). We present the results of Spearman's correlation in the Supplemental Information (SI 3). We find that although Spearman's correlation performs about as well as Pearson's and polychoric correlations, they do not fare as well as tetrachoric correlations in dichotomous data. For this reason, we present only the results of EBICglasso using Pearson's, polychoric, and dichotomous correlations in the main text.

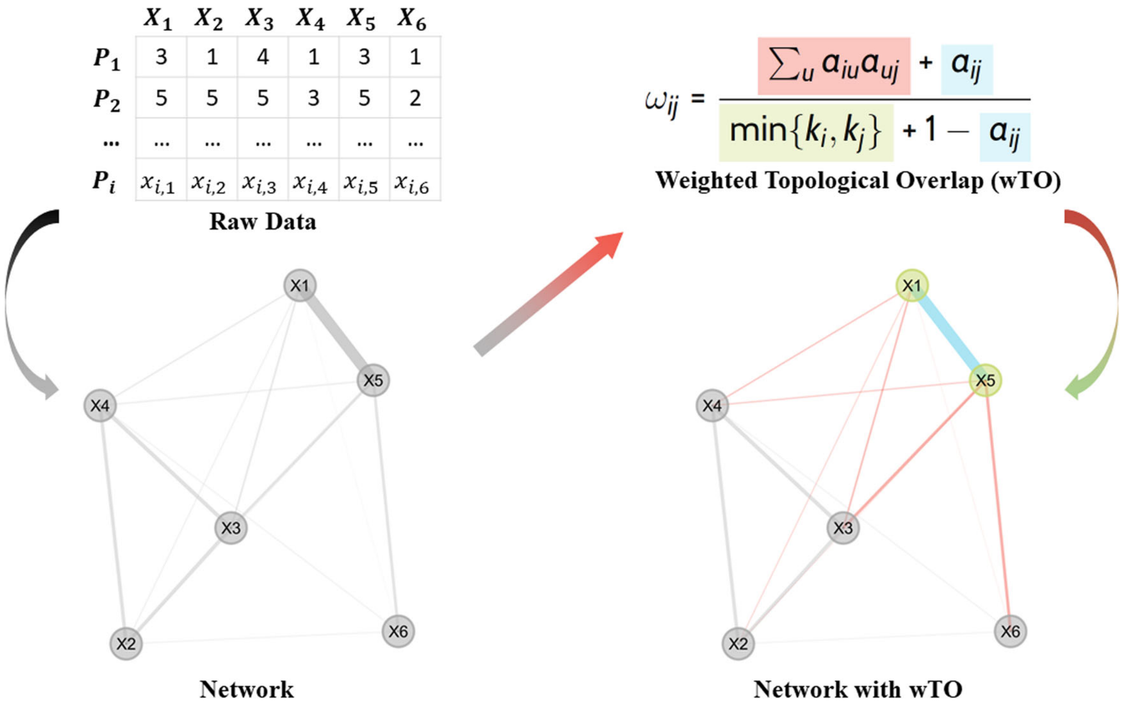

Figure 1. Depiction of unique variable analysis process. The raw data are used to estimate a network. Then, weighted topological overlap (wTO) is applied to the network. The colors are coordinated between the wTO equation and network with wTO: green = target nodes i (X_1) and j (X_5) where the minimum of their number of connections is obtained, red = shared edges to other nodes in the network, and blue = edge between node i and j.

1988) are reported. All large effects (including interactions involving up to three conditions) are presented in figures in the supplementary information (SI 4). To aid interpretation of all results, an interactive {Shiny} application (version 1.7.4; Chang et al., 2022) in R was created: https://alex-christensen.shinyapps.io/local_dependence_results/. This application allows for all possible method, data type, and condition combinations to be visualized and we encourage the interested reader to visualize any main effects or interactions in the application.

Table 1. Overall local dependence performance.

Method	CSI	FDR	Sensitivity	Specificity
wTO (BGGM)	0.861	0.026	0.882	**0.999**
wTO (EBICglasso)	**0.880**	**0.020**	0.895	**0.999**
Partial R (Significance)	0.370	0.615	0.930	0.971
Partial R (Cutoff)	0.787	0.167	**0.939**	0.942
SEPC (Significance)	0.205	0.751	0.760	0.924
SEPC (Cutoff)	0.738	0.130	0.843	0.994

Note. Shaded cells indicate top three best values and bolded text indicates best value. CSI = critical success index, FDR = false discovery rate, wTO = weighted topological overlap, R = correlation, SEPC = standardized expected parameter change.

Overall

Local dependence

Overall, for conditions with local dependence, wTO with EBICglasso had the highest CSI (0.880) and lowest FDR (0.020) followed by wTO with BGGM (CSI = 0.861 and FDR = 0.026) and partial correlation with cutoff (CSI = 0.787 and FDR = 0.167; Table 1). In terms of sensitivity or detection of true local dependence, partial correlation with cutoff was the best (0.939) followed by partial correlation with significance (sensitivity = 0.930) and wTO with EBICglasso (sensitivity = 0.895). Despite strong performance detecting local dependence, partial correlation with significance had the second worst FDR (0.615). This FDR value suggests that over 60% of positives detected are false positive meaning that it's detecting nearly all locally dependent variables but at the cost of vastly overestimating the number of true locally dependent variables.

From the ANOVA, there was a main effect of data type for partial correlation with cutoff ($\eta_p^2 = 0.54$), SEPC with cutoff ($\eta_p^2 = 0.18$), and wTO with BGGM ($\eta_p^2 = 0.23$). For all three methods, CSI decreased from continuous to polytomous to dichotomous data. Three methods had a main effect of residual correlation size: SEPC with cutoff ($\eta_p^2 = 0.31$), wTO with BGGM ($\eta_p^2 = 0.33$), and wTO with EBICglasso ($\eta_p^2 = 0.36$). For all three methods, CSI increased as the size of the correlated residual increased. Partial correlation with cutoff ($\eta_p^2 = 0.35$) and SEPC with cutoff ($\eta_p^2 = 0.14$) had a main effect of sample size such that CSI increased as sample size increased. Two methods had a main effect of number of variables per factor: partial correlation

Table 2. Continuous local dependence performance.

Method	CSI	FDR	Sensitivity	Specificity
wTO (BGGM)	**0.939**	0.034	0.972	0.998
wTO (EBICglasso)	0.914	0.004	0.918	**1.000**
Partial R (Significance)	0.414	0.585	**0.996**	0.973
Partial R (Cutoff)	0.936	**0.003**	0.939	**1.000**
SEPC (Significance)	0.192	0.674	0.567	0.938
SEPC (Cutoff)	0.843	0.014	0.857	0.999

Note. Shaded cells indicate top three best values and bolded text indicates best value. CSI = critical success index, FDR = false discovery rate, wTO = weighted topological overlap, R = correlation, SEPC = standardized expected parameter change.

Table 3. Polytomous local dependence performance.

Method	CSI	FDR	Sensitivity	Specificity
wTO (BGGM)	**0.912**	0.025	0.933	**0.999**
wTO (EBICglasso)	0.901	**0.011**	0.911	**0.999**
Partial R (Significance)	0.391	0.607	**0.991**	0.971
Partial R (Cutoff)	0.892	0.058	0.944	0.997
SEPC (Significance)	0.188	0.807	0.811	0.900
SEPC (Cutoff)	0.763	0.100	0.856	0.996

Note. Shaded cells indicate top three best values and bolded text indicates best value. CSI = critical success index, FDR = false discovery rate, wTO = weighted topological overlap, R = correlation, SEPC = standardized expected parameter change.

with significance ($\eta_p^2 = 0.31$) and SEPC with significance ($\eta_p^2 = 0.26$). For both methods, CSI decreased as the number of variables per factor increased. Finally, partial correlation with significance had a main effect of number of factors ($\eta_p^2 = 0.71$) such that CSI decreased as factors increased.

No local dependence

Overall, for conditions with no local dependence, wTO with EBICglasso had the lowest FP (0.085) followed by wTO with BGGM (FP = 0.113) and SEPC with cutoff (FP = 0.296). After these three, there was a substantial jump in false positives: SEPC with significance (FP = 5.684), partial correlation with significance (FP = 12.202), and partial correlation with cutoff (FP = 23.510).

From the ANOVA, there were three methods with a main effect of data type: partial correlation with cutoff ($\eta_p^2 = 0.28$), SEPC with significance ($\eta_p^2 = 0.25$), and SEPC with cutoff ($\eta_p^2 = 0.32$). For partial correlation with cutoff and SEPC with significance, FP increased as data type went from continuous to polytomous to dichotomous (values are reported in their respective data type sections). For SEPC with significance, FP decreased from continuous to polytomous but increased from polytomous to dichotomous. All methods but wTO with BGGM and wTO with EBICglasso had a main effect of sample size: partial correlation with cutoff ($\eta_p^2 = 0.17$), SEPC with significance ($\eta_p^2 = 0.56$), SEPC with cutoff ($\eta_p^2 = 0.31$), and partial correlation with significance ($\eta_p^2 = 0.27$). For the former three, FP decreased as sample size increased. For partial correlation with significance, FP increased as sample size increased.

The two significance methods had main effects for number of variables per factor and number of factors: partial correlation with significance ($\eta_p^2 = 0.86$ and $\eta_p^2 = 0.92$, respectively) and SEPC with significance ($\eta_p^2 = 0.60$ and $\eta_p^2 = 0.64$, respectively). For both methods, FP increased as variables per factors increased as well as when the number of factors increased.

Continuous data

Local dependence

Focusing on continuous data, wTO with BGGM had the highest CSI (0.939 and FDR = 0.034) followed by partial correlation with cutoff, which had the lowest FDR (CSI = 0.936 and FDR = 0.003), and wTO with EBICglasso (CSI = 0.914 and FDR = 0.004; Table 2). In terms of sensitivity, partial correlation with significance performed best (0.996) followed by wTO with BGGM (0.972) and partial correlation with cutoff (0.939). Once again, however, sensitivity must be contextualized by FDR, which demonstrated that partial correlation with significance substantially overestimated the number of local dependencies (FDR = 0.585).

From the ANOVA, there was a main effect of correlated residual size for partial correlated with cutoff ($\eta_p^2 = 0.27$), SEPC with cutoff ($\eta_p^2 = 0.47$), and wTO with EBICglasso ($\eta_p^2 = 0.35$). All three methods had CSI increase as the size of the correlated residual increased. This effect was primarily driven by the smallest correlated residual size (0.25). There were two methods with a main effect of the number of variables per factor: partial correlation with significance ($\eta_p^2 = 0.24$) and SEPC with significance ($\eta_p^2 = 0.52$). For both methods, CSI decreased as the number of variables per factor increased. SEPC with significance had a main effect of sample size ($\eta_p^2 = 0.19$) such that CSI increased as sample size increased. Partial correlation with significance had a main effect of number of factors ($\eta_p^2 = 0.75$) such that CSI decreased as the number of factors increased.

No local dependence

Focusing on continuous data with no local dependence, SEPC with cutoff had the lowest FP (0.000) followed by partial correlation with cutoff (FP = 0.010), wTO with EBICglasso (FP = 0.016), and wTO with BGGM (FP = 0.159). Then, there was a substantial jump in false positives: SEPC with significance (FP = 7.092) and partial correlation with significance (FP = 12.468).

Table 4. Dichotomous local dependence performance.

Method	CSI	FDR	Sensitivity	Specificity
wTO (BGGM)	0.732	**0.020**	0.742	**0.999**
wTO (EBICglasso)	**0.824**	0.045	0.856	0.998
Partial R (Significance)	0.306	0.657	0.804	0.969
Partial R (Cutoff)	0.533	0.441	**0.936**	0.829
SEPC (Significance)	0.234	0.752	0.901	0.933
SEPC (Cutoff)	0.609	0.276	0.818	0.988

Note. Shaded cells indicate top three best values and bolded text indicates best value. CSI = critical success index, FDR = false discovery rate, wTO = weighted topological overlap, R = correlation, SEPC = standardized expected parameter change.

From the ANOVA, there were two methods with a main effect of sample size: SEPC with significance ($\eta_p^2 = 0.15$) and wTO with BGGM ($\eta_p^2 = 0.14$). Both methods showed a decreased in FP as sample size increased with effects mainly driven by the smallest sample size (250). Partial correlation with significance and SEPC with significance had main effects of number of variables per factor ($\eta_p^2 = 0.88$ and $\eta_p^2 = 0.58$, respectively) and number of factors ($\eta_p^2 = 0.94$ and $\eta_p^2 = 0.69$, respectively). For both methods, FP increased as the number of variables per factor increased as well as when the number of factors increased.

Polytomous data

Local dependence

Focusing on polytomous data, wTO with BGGM had the highest CSI (0.912 and FDR = 0.025) followed by wTO with EBICglasso, which had the lowest FDR (CSI = 0.901 and FDR = 0.011), and partial correlation with cutoff (CSI = 0.892 and FDR = 0.058; Table 3). In terms of sensitivity, partial correlation with significance performed best (0.991 and FDR = 0.607) followed by partial correlation with cutoff (0.944) and wTO with BGGM (0.933).

From the ANOVA, there was a main effect for correlated residual size for all methods except partial correlation with significance: partial correlation with cutoff ($\eta_p^2 = 0.14$), SEPC with significance ($\eta_p^2 = 0.29$), SEPC with cutoff ($\eta_p^2 = 0.28$), wTO with BGGM ($\eta_p^2 = 0.24$), and wTO with EBICglasso ($\eta_p^2 = 0.35$). For SPEC with significance, CSI decreased as the size of the correlated residual increased. For SEPC with cutoff, CSI increased from 0.25 to 0.35 but decreased from 0.35 to 0.45. For partial correlation with cutoff, wTO with BGGM, and wTO with EBICglasso, CSI increased as the size of the correlated residual increased.

Partial correlation with cutoff ($\eta_p^2 = 0.18$) and SEPC with significance ($\eta_p^2 = 0.44$) had a main effect of sample size. The CSI for partial correlation increased as sample size increased; the CSI for SEPC with significance decreased as sample size increased. Partial correlation with significance ($\eta_p^2 = 0.29$) and SEPC with significance ($\eta_p^2 = 0.15$) had a main effect of number of variables per factor such that CSI decreased as the number of variables per factor increased. Partial correlation with significance also had a main effect of number of factors ($\eta_p^2 = 0.73$) such that CSI decreased as the number of factors increased.

No local dependence

Focusing on polytomous data with no local dependence, SEPC with cutoff had the lowest FP (0.016) followed by wTO with EBICglasso (FP = 0.040), wTO with BGGM (FP = 0.108), and partial correlation with cutoff (0.306). Then, there was a substantial jump in false positives: SEPC with significance (FP = 3.585) and partial correlation with significance (FP = 12.925).

From the ANOVA, both partial correlation with significance and SEPC with significance had main effects for number of variables per factor ($\eta_p^2 = 0.89$ and $\eta_p^2 = 0.51$, respectively) and number of factors ($\eta_p^2 = 0.94$ and $\eta_p^2 = 0.53$, respectively). For both methods, FP increased as number of variables per factor increased as well as when number of factors increased.

Dichotomous data

Focusing on dichotomous data, wTO with EBICglasso had the highest CSI (0.824 and FDR = 0.045) followed by wTO with BGGM, which had the lowest FDR (CSI = 0.732 and FDR = 0.020), and SEPC with cutoff (CSI = 0.609 and FDR = 0.276; Table 4). In terms of sensitivity, partial correlation with cutoff performed best (0.936 and FDR = 0.441) followed by SEPC with significance (sensitivity = 0.901 and FDR = 0.752) and wTO with EBICglasso (sensitivity = 0.856).

Local dependence

From the ANOVA, there was a main effect of correlated residual size for SEPC with cutoff ($\eta_p^2 = 0.20$), wTO with BGGM ($\eta_p^2 = 0.55$), and wTO with EBICglasso ($\eta_p^2 = 0.39$). For wTO with BGGM and wTO with EBICglasso, CSI increased as size of correlated residual increased. For SEPC with cutoff, CSI increased from 0.25 to 0.35 but decreased from 0.35 to 0.45. Partial correlation with significance ($\eta_p^2 = 0.17$), partial correlation with cutoff ($\eta_p^2 = 0.69$), SEPC with cutoff ($\eta_p^2 = 0.38$), and wTO with

EBICglasso ($\eta_p^2 = 0.15$) had a main effect of sample size such that CSI increased as sample size increased.

Partial correlation with significance and partial correlation with cutoff both had a main effect of number of variables per factor ($\eta_p^2 = 0.40$ and $\eta_p^2 = 0.24$, respectively) and number of factors ($\eta_p^2 = 0.64$ and $\eta_p^2 = 0.34$, respectively). For both methods, CSI decreased as the number of variables per factor increased as well as when the number of factors increased.

No local dependence

Focusing on dichotomous data with no local dependence, wTO with BGGM had the lowest FP (0.071) followed by wTO with EBICglasso (FP = 0.200) and SEPC with cutoff (FP = 0.871). Then, there was a substantial jump in false positives: SEPC with significance (FP = 6.376), partial correlation with significance (FP = 11.213), and partial correlation with cutoff (FP = 70.213).

From the ANOVA, there was a main effect of sample size for partial correlation with significance ($\eta_p^2 = 0.68$), partial correlation with cutoff ($\eta_p^2 = 0.37$), and SEPC with cutoff ($\eta_p^2 = 0.57$). For partial correlation with cutoff and SEPC with cutoff, FP decreased as sample size increased with the effect being primarily driven by the smallest sample size (250). For partial correlation with significance, FP increased as sample size increased.

Partial correlation with significance ($\eta_p^2 = 0.77$), partial correlation with cutoff ($\eta_p^2 = 0.32$), and SEPC with significance ($\eta_p^2 = 0.69$) had a main effect of number of variables per factor such that FP increased as number of variables per factor increased. These methods as well as SEPC with cutoff had a main effect of number of factors: partial correlation with significance ($\eta_p^2 = 0.84$), partial correlation with cutoff ($\eta_p^2 = 0.29$), SEPC with significance ($\eta_p^2 = 0.68$), and SEPC with cutoff ($\eta_p^2 = 0.21$). For all four methods, FP increased as the number of factors increased.

Empirical example

For the empirical example, all methods evaluated in the simulation were applied to the BAPQ (Hurley et al., 2007). The BAPQ was completed by 5,659 people who were fathers and mothers of a child with an autism spectrum disorder. The original internal structure proposed by Hurley et al. (2007) has three factors which capture different aspects of the broad autism phenotype: aloof personality representing a limited interest in or enjoyment of social interactions, rigid personality representing a resistance and/or difficulty adapting to change, and pragmatic language representing deficits in the social use of language leading to difficulties with effective communication and/or conversational reciprocity. Each factor has 12 items (36 items in total), which were responded to using a 6-point Likert scale (for item descriptions, see SI 5).

To apply the SEPC with significance and partial correlation with significance methods, the theoretical number of factors (three) were used. ESEM was estimated using WLSMV (Rhemtulla et al., 2012) and polychoric correlations were used for the other methods. Each method is listed with the number of locally dependent pairs of variables they estimated as locally dependent: SEPC with significance estimated 357 locally dependent pairs, SEPC with cutoff estimated 2 locally dependent pairs, partial correlation with significance estimated 25 locally dependent pairs, partial correlation with cutoff estimated 2 locally dependent pairs, wTO with EBICglasso estimated 2 locally dependent pairs, and wTO with BGGM estimated no locally dependent pairs. Consistent with our simulation study, SEPC with significance and partial correlation with significance had substantially high estimates of locally dependent pairs (likely overestimating the number of locally dependent pairs). In contrast, their cutoff counterparts only estimated two locally dependent pairs of variables.

For SEPC with cutoff, partial correlation with cutoff, and wTO with EBICglasso, the same two pairs of variables were identified as locally dependent: 4 ("It's hard for me to avoid getting sidetracked in conversation") and 32 ("I lose track of my original point when talking to people") as well as 21 ("I can tell when someone is not interested in what I am saying") and 34 ("I can tell when it is time to change topics in conversation"). Both local dependencies were in the pragmatic language factor.

Discussion

This simulation set out with three aims: (1) introduce a novel local dependence detection method based on network psychometrics and compare it to contemporary methods of local dependence detection, (2) evaluate whether statistical significance approaches perform better than cutoff values, and (3) examine the generalizability of these approaches across continuous and categorical data. We found that the network psychometrics local dependence detection method outperformed all contemporary methods. Further, we found that a single, unchanging cutoff substantially outperformed their statistical significance counterparts. Key to these conclusions was the use of the confusion matrix metrics CSI, FDR, and FP rather than the more commonly used sensitivity and specificity.

Overall, wTO with EBICglasso had the highest CSI (0.880), lowest FDR (0.020), and lowest FP (0.085) values of any method tested. When there are local dependencies, on average, wTO with EBICglasso is expected to have 2 false positives out of 100 total positives. When there are no local dependencies, on average, wTO with EBICglasso is expected to suggest 1 locally dependent pair about every 12 samples (or every $\frac{1}{0.085} = 11.77$ no local dependence conditions across all conditions). This combination was closely followed by wTO with BGGM (overall CSI $= 0.861$, overall FDR $= 0.026$, overall FP $= 0.111$), which was only outperformed by wTO with EBICglasso in dichotomous data. Based on the metrics used in this study, wTO with EBICglasso would be the recommended method for general use cases.

For continuous data, these methods and partial correlation with cutoff (0.35) were substantially better than the other methods. Based on the combination of CSI (0.936), FDR (0.003), and FP (0.010), partial correlation with cutoff would be the recommended method. Its CSI was nearly as high as wTO with BGGM (CSI $= 0.939$; the highest CSI for continuous data) but its FDR was one-tenth of wTO with BGGM's FDR. By comparison, partial correlation with cutoff would, on average, have 3 positives out of 1000 positives be false positive relative to wTO with BGGM which would, on average, have 3.4 positives out of 100 positives be false positive. Similarly, its FP in the no local dependence conditions was over fifteen times lower than wTO with BGGM (FP $= 0.158$). wTO with EBICglasso was on par with partial correlation with cutoff for FDR (0.004) and FP (0.016) but had a slightly lower CSI (0.914).

For polytomous data, these same three methods (wTO with BGGM, wTO with EBICglasso, and partial correlation with cutoff), all performed comparably well with respect to CSI (0.912, 0.901, and 0.892, respectively). wTO with EBICglasso had the best FDR (0.011) and FP (0.040) relative to wTO with BGGM (FDR $= 0.025$ and FP $= 0.108$) and partial correlation with cutoff (FDR $= 0.058$ and FP $= 0.306$). Either wTO with BGGM or wTO with EBICglasso would be the recommended method depending on whether identifying all local dependencies or fewer false positives were given preference. For identifying all local dependencies, wTO with BGGM would be preferred given its higher sensitivity (0.933) relative to wTO with EBICglasso (sensitivity $= 0.911$). For identifying fewer false positives, wTO with EBICglasso would be preferred given its over two times lower FDR.

For dichotomous data, wTO with EBICglasso (CSI $= 0.824$, FDR $= 0.045$, FP $= 0.200$) and wTO with BGGM (CSI $= 0.732$, FDR $= 0.020$, FP $= 0.071$) substantially outperformed all other methods (nearest being SEPC with cutoff: CSI $= 0.609$, FDR $= 0.276$, FP $= 0.871$). It's notable that performance was substantially worse for all methods in dichotomous data relative to continuous and polytomous data. This challenge is well-documented in the item response theory literature (Edwards et al., 2018). Still, wTO with EBICglasso was effective and is the clear cut recommended method for dichotomous data based on our simulation.

The above recommendations were hardly inevitable. If focusing solely on sensitivity and specificity, as nearly all other local dependence detection simulations have done (Chen & Thissen, 1997; Edwards et al., 2018; Ferrando et al., 2022; Houts & Edwards, 2013; Yen, 1984), then partial correlation with significance would have been the clear cut choice, posting some of the best sensitivity values in the overall (0.930), continuous (0.996), and polytomous (0.991) data. The CSI and FDR metrics demonstrate that focusing on sensitivity and specificity, in the context of detecting local independence violations, is misleading. Despite high sensitivity or ability to detect most simulated local dependencies, the method had substantially high FDR (overall $= 0.615$, continuous $= 0.585$, polytomous $= 0.607$). With FDR, sensitivity is contextualized in a way that provides information about the tradeoff between detecting all local dependencies with false positives. Even in the best case, partial correlation with significance had an FDR an around 0.580 or 58% of all locally dependent pairs it detected were false positives. In applied settings, it would be difficult to determine which estimated pairs of variables are truly locally dependent given that nearly every other pair would be a false positive. Our study highlights the issue of using sensitivity and specificity metrics only to make recommendations about the effectiveness of a local dependence detection method. We urge future studies to use CSI and FDR in combination with sensitivity and specificity when evaluating local dependence detection methods.

When investigating whether statistical significance or a single cutoff value is sufficient, we found a single cutoff value to be superior. This finding is consistent with Whittaker (2012) who found that a SEPC cutoff worked better than statistical significance based on modification indices in some conditions. Our grid search replicated the SEPC cutoff value of 0.20 identified by Saris et al. (2009) for continuous data (SI 2). Overall, however, we found the SEPC value of 0.25 to be most effective. Similarly, the statistical significance approach proposed by Ferrando et al. (2022) did not perform as well as a cutoff value for partial correlations. Although the approach proposed by Ferrando et al. (2022) is grounded in statistical theory, our simulation demonstrates that the approach is not an effective method to detect locally dependent variables. Both significance approaches were significantly affected by the number of variables per factor and number of factors suggesting that these methods are generally affected by the total number of variables in a dataset. Future studies using significance approaches may want to consider some form of multiple comparisons correction for the number of possible pairs given our findings (e.g., Pérez & Pericchi, 2014).

Although we tested several different local dependence detection methods across a wide range of conditions, there remains several open questions and directions for future research. The use of polychoric and tetrachoric correlations in the EBICglasso is not strictly correct with the estimate of the likelihood expression of the model, which expects a sample covariance matrix, and makes comparisons to the ESEM with SEPC inconsistent. Future work should investigate more appropriate implementations of the GGM estimation using the `ggm` function in the {psychonetrics} package (Epskamp, 2019). In terms of other comparisons, adaptations of approaches applied in this study could be used. The resampling partial correlation approach could be applied to the regularized partial correlations with wTO. Such an approach would not rely on a singular cutoff value as proposed in this study. Similarly, wTO could be applied to the resampling approach of partial correlations where the values larger than the sampling error cutoff are treated as a method to set values in the partial correlation matrix to zero rather than the cutoff for local dependence. After, wTO could be applied to the sparse matrix. There are also other methods proposed by Ferrando et al. (2022), such as EREC and ENIDE, that we did not test in this study. We note that the EREC method closely aligns with the ESEM with SEPC, the main difference being the use of exploratory factor analysis rather than ESEM and the procedure used to determine significance (resampling and modification indices, respectively). Despite these differences, the methods are similar enough to suggest that our results may be roughly indicative of how EREC might have performed in this study. Further, our results suggest that all previous recommendations that have been based on sensitivity and specificity should be reevaluated using CSI and FDR to determine whether they still hold weight.

The issues associated with local independence violations are far reaching. Contemporary psychometric methods, including network psychometrics, are sensitive to the effects of these violations (Edwards et al., 2018; Fried & Cramer, 2017). Researchers should attempt to reduce local dependence as much as possible in psychometric validation. Local dependence often shows up as redundant indicators that have shared semantic reference (i.e., using similar item phrasing or similar item content; Leising et al., 2020; Rosenbusch et al., 2020). Similar item phrasing may be a well-known scale construction strategy, but it reflects either a lack of strong theory to guide the scale construction or a lack of construct bandwidth which result in the development of scales with artificially inflated internal consistency through "cheating by repeating" (Reise et al., 2018). The weighted topological overlap paired with EBICglasso approach, which we call *Unique Variable Analysis*, provides researchers with a robust method to identify whether any local independence violations exist in their measures.

Broadly, local dependence detection should be integrated into other test validation practices (Flake et al., 2017). We view local dependence detection as the first step of structural validation (Christensen et al., 2020). Our hope is that researchers will become more aware about the violations of local independence as it can have dire consequences for psychometric modeling.

Open Scholarship

This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. The data and materials are openly accessible at https://doi.org/10.5281/zenodo.1212328 and https://doi.org/10.17605/osf.io/9w3jy. To obtain the author's disclosure form, please contact the Editor.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: L.E.G. was supported by Grant 2018-2019-1D2-085 from the Fondo Nacional de Innovación y Desarrollo Científico y Tecnológico (FONDOCYT) of the Dominican Republic.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgements: The authors did not preregister the study. All data, code, and materials can be found on the Open Science Framework: https://osf.io/9w3jy/. Alexander P. Christensen https://orcid.org/0000-0002-9798-7037, Luis Eduardo Garrido https://orcid.org/0000-0001-8932-6063, and Hudson Golino https://orcid.org/0000-0002-1601-1447. The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions or the funding agency is not intended and should not be inferred.

ORCID

Alexander P. Christensen http://orcid.org/0000-0002-9798-7037
Luis Eduardo Garrido http://orcid.org/0000-0001-8932-6063
Hudson Golino http://orcid.org/0000-0002-1601-1447

Data availability statement

All data and R scripts can be found on the Open Science Framework: https://osf.io/9w3jy/.

The authors made the following contributions. Alexander P. Christensen: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization, Writing - Original Draft, Writing - Review & Editing; Luis Eduardo Garrido: Conceptualization, Formal Analysis, Supervision, Writing - Review & Editing; Hudson Golino: Conceptualization, Formal Analysis, Data Curation, Supervision, Writing - Review & Editing.

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 397–438. https://doi.org/10.1080/10705510903008204

Aust, F., & Barth, M. (2022). *papaja: Create APA manuscripts with R Markdown*. https://github.com/crsh/papaja

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*, 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. (2021). 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

Briganti, G., Williams, D. R., Mulder, J., & Linkowski, P. (2022). Bayesian network structure and predictability of autistic traits. *Psychological Reports*, *125*(1), 344–357. https://doi.org/10.1177/0033294120978159

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T., & Snippe, E. (2019). What do centrality measures measure in psychology networks? *Journal of Abnormal Psychology*, *128*, 892–903. https://doi.org/10.1037/abn0000446

Chang, W., Cheng, J. J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2022). *shiny: Web Application Framework for R*. https://CRAN.R-project.org/package=shiny

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*, 759–771. https://doi.org/10.1093/biomet/asn034

Chen, W.-H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, *22*(3), 265–289. https://doi.org/10.3102/10769986022003265

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

Christensen, A. P., Golino, H., & Silvia, P. J. (2020). A psychometric network perspective on the validity and validation of personality trait questionnaires. *European Journal of Personality*, *34*, 1095–1108. https://doi.org/10.1002/per.2265

Christensen, A. P., Nieto Canaveras, M. D., Golino, H., Garrido, L. E., Jimenez, M., Abad, F., Garcia-Garzon, E., & Franco, V. R. (2022). *latentFactoR: Data simulation based on latent factors*.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis* (2nd ed.). Psychology Press.

Cramer, A. O. J. (2012). Why the item "23 + 1" is not in a depression questionnaire: Validity from a network perspective. *Measurement: Interdisciplinary Research & Perspective*, *10*, 50–54. https://doi.org/10.1080/15366367.2012.681973

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*, 414–431. https://doi.org/10.1002/per.1866

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.

Edwards, M. C., Houts, C. R., & Cai, L. (2018). A diagnostic procedure to detect departures from local independence in item response theory models. *Psychological Methods*, *23*(1), 138–149. https://doi.org/10.1037/met0000121

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

Epskamp, S., Kruis, J., & Marsman, M. (2017). Estimating psychopathological networks: Be careful what you wish for. *PLoS One*, *12*(6), e0179891. https://doi.org/10.1371/journal.pone.0179891

Epskamp, S., Maris, G., Waldrop, L. J., & Borsboom, D. (2018). Network psychometrics. In P. Irwing, D. Hughes, & T. Booth (Eds.), *The Wiley handbook of psychometric testing, 2 volume set: A multidisciplinary reference on survey, scale and test development* (pp. 953–986). Wiley. https://doi.org/10.1002/9781118489772.ch30

Epskamp, S. (2019). *psychonetrics: Structural equation modeling and confirmatory network analysis*. R package verson 0.3.3. https://CRAN.R-project.org/package=psychonetrics

Ferrando Piera, P. J., & Lorenzo Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, *29*(2), 236–240. https://doi.org/10.7334/psicothema2016.304

Ferrando, P. J., Hernandez-Dorado, A., & Lorenzo-Seva, U. (2022). Detecting correlated residuals in exploratory factor analysis: New proposals and a comparison of procedures. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(4), 630–638. https://doi.org/10.1080/10705511.2021.2004543

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*, 370–378. https://doi.org/10.1177/1948550617693063

Foygel, R., & Drton, M. (2010). Extended Bayesian information criteria for Gaussian graphical models. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems* (pp. 604–612). Retrieved from https://papers.nips.cc/paper/4087-extended-bayesian-information-criteria-for-gaussian-graphical-models

Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999–1020. https://doi.org/10.1177/1745691617705892

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

Friedman, J., Hastie, T., & Tibshirani, R. (2014). *glasso: Graphical lasso – estimation of Gaussian graphical models*. https://CRAN.R-project.org/package=glasso

García-Garzón, E., Abad, F. J., & Garrido, L. E. (2019). Improving bi-factor exploratory modeling: Empirical target rotation based on loading differences. *Methodology*, *15*, 45–55. https://doi.org/10.1027/1614-2241/a000163

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. *Educational and Psychological Measurement*, *71*, 551–570. https://doi.org/10.1177/0013164410389489

Golino, H., & Epskamp, S. (2017). Exploratory Graph Analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One*, *12*, e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*, 292–320. https://doi.org/10.1037/met0000255

Guttman, L. (1953). Image theory for the structure of quantitative variates. *Psychometrika*, *18*, 277–296.

Gysi, D. M., Voigt, A., de Miranda Fragoso, T., Almaas, E., & Nowick, K. (2018). wTO: An R package for computing weighted topological overlap and a consensus network with integrated visualization tool. *BMC Bioinformatics*, *19*, 392. https://doi.org/10.1186/s12859-018-2351-7

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

Houts, C. R., & Edwards, M. C. (2013). The performance of local dependence measures with psychological data. *Applied Psychological Measurement*, *37*(7), 541–562. https://doi.org/10.1177/0146621613491456

Hubley, A. M., Zhu, S. M., Sasaki, A., & Gadermann, A. M. (2014). Synthesis of validation practices in two assessment journals: Psychological Assessment and the European Journal of Psychological Assessment. In B. Zumbo & E. Chan (Eds.), *Validity and validation in social, behavioral, and health sciences* (pp. 193–213). Springer.

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679–1690. https://doi.org/10.1007/s10803-006-0299-3

Isvoranu, A.-M., & Epskamp, S. (2021). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods*, https://doi.org/10.1037/met0000439

Jamison, L., Christensen, A. P., & Golino, H. (2022). Metric invariance in exploratory graph analysis via permutation testing. PsyArXiv. https://doi.org/10.31234/osf.io/j4rx9

Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2022). Dimensionality assessment in generalized bi-factor structures: A network psychometrics approach. PsyArXiv. https://doi.org/10.31234/osf.io/2ujdk

Kaplan, D. (1989). Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behavioral Research*, *24*(3), 285–305. https://doi.org/10.1207/s15327906mbr2403_2

Kassambara, A. (2018). *ggpubr: 'ggplot2' based publication ready plots*. https://CRAN.R-project.org/package=ggpubr

Lauritzen, S. L. (1996). *Graphical models*. Clarendon Press.

Ledermann, W. (1937). On the rank of the reduced correlational matrix in multiple-factor analysis. *Psychometrika*, *2*(2), 85–93. https://doi.org/10.1007/BF02288062

Leising, D., Burger, J., Zimmermann, J., Bäckström, M., Oltmanns, J., & Connelly, B. (2020). Why do items correlate with one another? A conceptual analysis with relevance for general factors and network models. PsyArXiv, https://doi.org/10.31234/osf.io/7c895

Letina, S., Blanken, T. F., Deserno, M. K., & Borsboom, D. (2019). Expanding network analysis tools in psychological networks: Minimal spanning trees, participation coefficients, and motif analysis applied to a network of 26 psychological attributes. *Complexity*, *2019*, 1–27. https://doi.org/10.1155/2019/9424605

Marinazzo, D., Van Roozendaal, J., Rosas, F. E., Stella, M., Comolatti, R., Colenbier, N., Stramaglia, S., & Rosseel, Y. (2022). *An information-theoretic approach to hypergraph psychometrics*. https://arxiv.org/abs/2205.01035

McDonald, R. P. (1985). *Factor analysis and related methods*. Psychology Press.

Montoya, A. K., & Edwards, M. C. (2020). The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. *Educational and Psychological Measurement*, *81*(3), 413–440. https://doi.org/10.1177/0013164420942899

Mulaik, S. A. (2010). *Foundations of factor analysis* (2nd ed.). CRC press. https://doi.org/10.1201/b15851

Muraki, E., & Carlson, J. E. (1995). Full-information factor analysis for polytomous item responses. *Applied Psychological Measurement*, *19*(1), 73–90. https://doi.org/10.1177/014662169501900109

Nowick, K., Gernat, T., Almaas, E., & Stubbs, L. (2009). Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22358–22363. https://doi.org/10.1073/pnas.0911376106

Pérez, M.-E., & Pericchi, L. R. (2014). Changing statistical significance with the amount of information: The adaptive α significance level. *Statistics & Probability Letters*, *85*, 20–24. https://doi.org/10.1016/j.spl.2013.10.018

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A.-L. (2002). Hierarchical organization of modularity in metabolic networks. *Science (New York, N.Y.)*, *297*(5586), 1551–1555. https://doi.org/10.1126/science.1073374

Reise, S. P., Bonifay, W., & Haviland, M. G. (2018). Bifactor modelling and the evaluation of scale scores. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing* (pp. 675–707). Wiley. https://doi.org/10.1002/9781118489772.ch22

Revelle, W. (2017). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315

Rosenbusch, H., Wanders, F., & Pit, I. L. (2020). The Semantic Scale Network: An online tool to detect semantic overlap of psychological scales and prevent scale redundancies. *Psychological Methods*, *25*(3), 380–392. https://doi.org/10.1037/met0000244

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

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, *17*, 105–129. https://doi.org/10.2307/271030

Saris, W., E., Satorra, A., Veld, W., & M., van der. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 561–582. https://doi.org/10.1080/10705510903203433

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, *4*, 5918. https://doi.org/10.1038/srep05918

Waldorp, L., & Marsman, M. (2021). Relations between networks, regression, partial correlation, and the latent variable model. *Multivariate Behavioral Research*, *57*(6), 994–1006. https://doi.org/10.1080/00273171.2021.1938959

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *The Journal of Experimental Education*, *80*(1), 26–44. https://doi.org/10.1080/00220973.2010.531299

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://ggplot2-book.org/

Williams, D. R. (2021). Bayesian estimation for Gaussian graphical models: Structure learning, predictability, and network comparisons. *Multivariate Behavioral Research*, *56*(2), 336–352. https://doi.org/10.1080/00273171.2021.1894412

Williams, D. R., & Mulder, J. (2020). BGGM: Bayesian Gaussian graphical models in R. *Journal of Open Source Software*, *5*(51), 2111. https://doi.org/10.21105/joss.02111

Williams, D. R., & Rast, P. (2018). Back to the basics: Rethinking partial correlation network methodology. *Journal of Mathematical and Statistical Psychology*, *73*(2), 187–212. https://doi.org/10.1111/bmsp.12173

Williams, D. R., Rhemtulla, M., Wysocki, A. C., & Rast, P. (2019). On nonregularized estimation of psychological networks. *Multivariate Behavioral Research*, *54*(5), 719–750. https://doi.org/10.1080/00273171.2019.1575716

Williams, D. R., & Rodriguez, J. E. (2022). Why overfitting is not (usually) a problem in partial correlation networks. *Psychological Methods*, *27*(5), 822–840. https://doi.org/10.1037/met0000437

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological*

Methods, *1*(4), 354–365. https://doi.org/10.1037/1082-989X.1.4.354

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, *8*, 125–145. https://doi.org/10.1177/014662168400800201

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*, 17. https://doi.org/10.2202/1544-6115.1128

Zhang, G., Trichtinger, L. A., Lee, D., & Jiang, G. (2022). PolychoricRM: A computationally efficient r function for estimating polychoric correlations and their asymptotic covariance matrix. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(2), 310–320. https://doi.org/10.1080/10705511.2021.1929996