Network structure of the Wisconsin Schizotypy Scales—Short Forms: Examining psychometric filtering approaches

By: Alexander P. Christensen, Yoed N. Kenett, Tomaso Aste, Paul J. Silvia, and Thomas R. Kwapil

Abstract:

Schizotypy is a multidimensional construct that provides a useful framework for understanding the etiology, development, and risk for schizophrenia-spectrum disorders. Past research has applied traditional methods, such as factor analysis, to uncovering common dimensions of schizotypy. In the present study, we aimed to advance the construct of schizotypy, measured by the Wisconsin Schizotypy Scales–Short Forms (WSS-SF), beyond this general scope by applying two different psychometric network filtering approaches—the state-of-the-art approach (lasso), which has been employed in previous studies, and an alternative approach (information-filtering networks; IFNs). First, we applied both filtering approaches to two large, independent samples of WSS-SF data (ns = 5,831 and 2,171) and assessed each approach's representation of the WSS-SF's schizotypy construct. Both filtering approaches produced results similar to those from traditional methods, with the IFN approach producing results more consistent with previous theoretical interpretations of schizotypy. Then we evaluated how well both filtering approaches reproduced the global and local network characteristics of the two samples. We found that the IFN approach produced more consistent results for both global and local network characteristics. Finally, we sought to evaluate the predictability of the network centrality measures for each filtering approach, by determining the core, intermediate, and peripheral items on the WSS-SF and using them to predict interview reports of schizophrenia-spectrum symptoms. We found some similarities and differences in their effectiveness, with the IFN approach's network structure providing better overall predictive distinctions. We discuss the implications of our findings for schizotypy and for psychometric network analysis more generally.

Keywords: Schizotypy | Network analysis | Schizophrenia-spectrum disorders

Article:

Schizotypy is a multidimensional construct that encompasses the subclinical and clinical continuum of schizophrenia-spectrum disorders. At its most extreme manifestations, schizotypy is expressed as full-blown schizophrenia. Converging evidence across behavioral, cognitive, neurobiological, and ambulatory assessment studies supports the overlap of the schizotypy continuum and schizophrenia-spectrum disorders (e.g., Ettinger, Meyhöfer, Steffens, Wagner, & Koutsouleris, 2014; Kwapil & Barrantes-Vidal, 2015). In addition, schizotypy provides early detection of schizophrenia-spectrum liability in nonclinical samples, prior to the onset of psychosis, medication, and stigmatization (Kwapil & Barrantes-Vidal, 2015). Therefore, schizotypy offers a promising framework for understanding the etiology, development, and expression of schizophrenia-spectrum disorders. Several traditional approaches, such as confirmatory factor analysis, have been used to examine the dimensional structure of schizotypy, typically identifying two to five underlying dimensions, with *positive*, *negative*, and *disorganized* schizotypy as the most replicated factors (Gross, Mellin, Silvia, Barrantes-Vidal, & Kwapil, 2014; Kwapil, Barrantes-Vidal, & Silvia, 2008; Raine & Benishay, 1995; Wuthrich & Bates, 2006). Despite these findings, traditional approaches do not account for the nature of the interactions taking place between items that contribute to schizophrenia-spectrum liability.

An increasingly popular approach in studying psychopathology is through network science (Borsboom, 2017). The network approach defines psychopathological disorders and personality traits as complex systems—phenomena that emerge from the causal interactions between symptoms and trait nuances (Borsboom & Cramer, 2013; Schmittmann et al., 2013). Such an approach can offer a unique perspective for examining schizophrenia-spectrum liability by defining schizotypy items as interacting symptom nuances, and determining which items are most central to the construct. Therefore, in the present study, we apply the network approach to examine the multidimensional structure of schizotypy via the 60-item Wisconsin Schizotypy Scales–Short Forms (WSS-SF; Winterstein et al., 2011). Furthermore, we employ interview measures to evaluate how the WSS-SF's network structure associates with interview-rated symptom measures.

Because the WSS-SF contains 60 items, the WSS-SF network will be convoluted with up to 1,770 possible connections—that is, a possible connection between every item. Therefore, filtering is needed to minimize spurious relations (multiple comparisons problem) and to increase interpretability (induce parsimony). To examine the network structure of schizotypy (via the WSS-SF), we apply two network filtering approaches to minimize spurious edges and maximize interpretability. One filtering approach, the lasso (Epskamp, Borsboom, & Fried, 2018), has been commonly applied in psychopathology research. The other filtering approach, Information Filtering Networks (IFN), has been previously applied to cognitive and neural networks (Kenett, Anaki, & Faust, 2014; Tewarie, van Dellen, Hillebrand, & Stam, 2015). Although the lasso network filtering approach has become popular in psychopathological research, little attention has been given to the limitations of this approach, such as biased comparability, reduced reproducibility, and arbitrary thresholding (Barfuss, Massara, Di Matteo, & Aste, 2016; Forbes, Wright, Markon, & Krueger, 2017; van Wijk, Stam, & Daffertshofer, 2010). Consequently, we compare the performance of these two filtering approaches using two large cross-sectional

datasets of the WSS-SF, and we conclude that the IFN-based filtering approach can circumvent some of the limitations of the popular lasso-based filtering approach.

Wisconsin Schizotypy Scales–Short Forms

The WSS-SF, a widely used set of scales, measures positive and negative schizotypy. The questionnaire includes two negative schizotypy subscales—Physical Anhedonia (diminution of sensory experiences; L. J. Chapman, Chapman, & Raulin, 1976) and Revised Social Anhedonia (disinterest in social experiences; Eckblad, Chapman, Chapman, & Mishlove, 1982) scales—and two positive schizotypy sub-scales—Perceptual Aberration (distortions of body image; L. J. Chapman, Chapman, & Raulin, 1978) and Magical Ideation (delusions and odd beliefs; Eckblad & Chapman, 1983) scales. The traditional factor structure of the WSS-SF has a positive and negative schizotypy factor, with the social anhedonia scale loading onto both factors (Gross, Silvia, Barrantes-Vidal, & Kwapil, 2015; Kwapil et al., 2008).

Kwapil, Barrantes-Vidal, and Silvia (2008) reported that positive and negative factors are differentially associated with interview measures of symptoms and impairment. Positive schizotypy is related to reports of psychotic-like experiences, substance abuse, mood disorders, and hospitalization. Negative schizotypy is associated with negative and schizoid symptoms, and the decreased likelihood of intimate relationships. Both positive and negative schizotypy are linked to poorer overall functioning and to paranoid and schizotypal symptoms (Gross et al., 2015; Kwapil et al., 2008). Furthermore, elevated scores on both dimensions are associated with increased liability of schizophrenia-spectrum disorders and are related to psychopathology, personality, and impaired social functioning (Barrantes-Vidal et al., 2013). Therefore, this factor structure of the WSS-SF appears to be highly reliable and valid, as shown across multiple studies (Gross et al., 2014; Kwapil et al., 2008; Kwapil, Ros-Morente, Silvia, & Barrantes-Vidal, 2012).

Although the factor structure and validity of the WSS-SF have been investigated by traditional methods, a clearer picture of how these items interact with each other and the importance of their interactions is still lacking (Kwapil & Barrantes-Vidal, 2015). For example, despite social anhedonia's positive loadings onto positive schizotypy, there has been little research on which positive schizotypy scale it's most related to. A factor analysis of the Schizotypal Personality Questionnaire (Raine, 1991), social anhedonia, perceptual aberration, and magical ideation scales suggests that the perceptual aberration scale could be connected to impaired social functioning (Wuthrich & Bates, 2006). Moderate correlations between social anhedonia and perceptual aberration seem to support this idea, however, magical ideation is also shown to be moderately related though to a lesser extent (Kwapil et al., 2008). Thus, an open question is what items of social anhedonia are the most connected to positive schizotypy—and to which scale? Another question is what items are more central to the construct? Are these items more related to clinical symptoms and impairment than less central items?

Investigating the underlying structure of the WSS-SF should provide clearer distinctions of the schizotypy construct it measures (Kwapil & Barrantes-Vidal, 2015), and guide the development of future schizotypy scales by identifying items most central to the schizotypy construct measured by the WSS-SF (Gross et al., 2014; Kwapil, Gross, Silvia, Raulin, & Barrantes-Vidal, 2017). Moreover, identifying items that are at the core of schizotypy measurements can

improve the detection and diagnosis of schizophrenia-spectrum psychopathology (Keshavan, Nasrallah, & Tandon, 2011). Psychometric network analysis provides an avenue to directly investigate connections within and between scales as well as the ability to determine which items are most central to the WSS-SF schizotypy construct.

Psychometric network analysis

The computational field of network science has greatly advanced the understanding of complex systems (Barabási, 2016). Such an approach is increasingly applied at the cognitive and psychological levels to quantitatively study cognitive phenomena in both typical and clinical populations (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013; Isvoranu et al., 2017; Karuza, Thompson-Schill, & Bassett, 2016). Recent research has applied psychometric networks (Epskamp, Maris, Waldorp, & Borsboom, in press) to investigate the intricate interactions of psychopathology and personality (Costantini et al., 2017). The network perspective offers a new conceptualization of psychopathology that takes the form of mutual, interacting symptoms through which disorder arises (Borsboom, 2017; Borsboom & Cramer, 2013; Fried et al., 2017).

Although network psychometric analysis can investigate the intricate interactions between the items of the WSS-SF, interpreting these interactions can be difficult. Networks contain multiple connections across all possible pairs of variables (e.g., symptoms, items) included in the model and therefore are likely to have spurious edges (i.e., multiple comparisons problem). Thus, filtering is necessary to minimize spurious connections and to increase the interpretability of the network. This, however, introduces a problem known as *sparse structure learning* (Zhou, 2011): How best to reduce the complexity and dimensionality of the network while retaining relevant information? To address this problem, different approaches have been developed with the aim of extracting meaningful and parsimonious models (Friedman, Hastie, & Tibshirani, 2008; Molinelli et al., 2013; Zhou, 2011).

Regression-based filtering approach

To date, regression-based filtering methods—removing edges on the basis of statistical significance or regressing each variable over all other variables—have dominated the psychometric network literature (Epskamp et al., 2018; Fried & Cramer, 2017; van Borkulo et al., 2014). Several different regression-based approaches have been developed, including the removal of edges below a certain threshold (Borsboom & Cramer, 2013; McNally et al., 2015), false discovery rate (Bringmann, Lemmens, Huibers, Borsboom, & Tuerlinckx, 2015; Bringmann et al., 2013), and the most widely used approach, the *least absolute shrinkage and selection operator* (lasso; Epskamp et al., 2018; van Borkulo et al., 2014).

The lasso filtering approach (Friedman et al., 2008; Tibshirani, 1996) minimizes spurious edges by using fully regressed coefficients (i.e., one variable regressed over all others) and shrinking small coefficients to zero through the application of an ℓ_1-regularization penalty on the estimation of the inverse covariance matrix (Barber & Drton, 2015; Ravikumar, Wainwright, & Lafferty, 2010). The ℓ_1-penalty is determined by the Extended Bayesian Information Criterion (EBIC; Chen & Chen, 2008), which controls for false positives. Because of this, the lasso

approach excels at preventing model overfitting, leading to better specificity (Epskamp & Fried, 2016). In dichotomous data, like that of the WSS-SF, the lasso approach is applied via the Ising model. The Ising model iteratively regresses one variable, using logistic regression, on all other variables and shrinks small coefficients to zero (van Borkulo et al., 2014).

The regularized fully regressed coefficients of the lasso induce parsimony by retaining fewer—sparser—connections in the network (Epskamp & Fried, 2016). Furthermore, these connections represent conditionally dependent variables: they are the association between two variables after controlling for all other variables in the network, and the lack of a connection represents conditional independence between two variables. This conditionality grants the possibility of a causal pathway existing between connected nodes (Epskamp & Fried, 2016). Thus, the advantages of the lasso approach are that the network specificity is high, the resulting network is sparse and more interpretable, and the connections can be interpreted as possible casual pathways between variables.

Although the lasso approach is state-of-the-art, with new algorithmic techniques constantly being developed (Costantini et al., 2017; Danaher, Wang, & Witten, 2014; Friedman et al., 2008; Ravikumar, Wainwright, Raskutti, & Yu, 2011), its limitations are receiving a growing amount of attention (Barfuss et al., 2016; Forbes et al., 2017; Guloksuz, Pries, & van Os, 2017; Wichers, Wigman, Bringmann, & de Jonge, 2017). The first limitation is the comparability of lasso-based networks between cross-sectional samples, and between clinical and nonclinical samples, is often biased because the number of edges included in lasso-based networks is a function of sample size—smaller samples retain fewer edges than larger samples (Epskamp & Fried, 2016; Fried et al., 2017). This variation in edge densities—the proportion of edges retained in the network over all possible edges—alters the structure of the network, making cross-sectional and psychopathological differences prone to be confounded with differences in edge density (Tewarie et al., 2015; van Wijk et al., 2010). For example, when a greater number of edges are in the network, the clustering coefficient (CC; the extent to which two neighbors of a node will be neighbors themselves) increases and the average shortest path length (ASPL; the mean shortest distance over all pairs of nodes) decreases. This means that in two WSS-SF networks node A (e.g., symptom A) could be significantly closer to (or further away from) node B (e.g., symptom B), suggesting that the potential of node A to influence node B has been altered. Therefore, differences in the ASPL and CC have the potential to meaningfully alter the interpretation of a symptom's potential to influence other symptoms. These attributes of networks are considered global characteristics because they measure the overall structure of the network.

Traditional psychometric network analysis has given little attention to these global characteristics and have instead focused on local network characteristics termed as centrality measures—measurements of a node's influence in the network. Notably, network centrality measures, such as betweenness centrality (the extent a node lies on paths between other nodes) and closeness centrality (average distance from all other nodes), are based on the distances between nodes in the network and are influenced by variation in edge density (Stam et al., 2014). Because edge densities alter the ASPL, varying distances between nodes, it is likely that the reproducibility of betweenness and closeness centrality are also affected. Indeed, a recent article opted to only discuss node strength—the sum of all edge weights connected to a node—because of the low reliability often reported for betweenness and closeness centrality measures (Fried et

al., 2017). Considering that only a few studies have examined global network characteristics (Boschloo, Schoevers, van Borkulo, Borsboom, & Oldehinkel, 2016; Costantini et al., 2017), it is possible that the low reliability of these centrality measures is related to differences in global network characteristics such as the ASPL.

In addition to decreased reliability of betweenness and closeness centrality measures, the average degree—the mean number of connections a node has in the network—of the network could also be altered due to differences in edge density (van Wijk et al., 2010). Consequently, this has the potential to influence the measurement of node strength by varying the amount of edges that are contributing to the sum of the edge weights. Moreover, greater edge densities could include smaller or larger edge weights, which would alter the average connectivity—mean strength of the edge weights in the network (De Schryver, Vindevogel, Rasmussen, & Cramer, 2015). Therefore, differing edge densities inevitably lead to heterogeneity in global and local (i.e., centrality) network measures, compromising comparability and reducing reproducibility. For the WSS-SF, different sample sizes mean that the lasso-based networks would have a different edge density, leading to different network structures despite measuring the same questionnaire. In short, the global and local network characteristics of the lasso-based networks have variable comparability and reproducibility across independent samples because they are sample size dependent (but see Costantini et al., 2017; van Borkulo et al., 2015).

Another potential pitfall is that conditional independence networks such as lasso-based networks tend to have a greater amount of measurement error—fully regressed coefficients raise the potential of edges between highly correlated variables to be arbitrarily dropped. For example, a symptom network of people scoring high in schizophrenia-spectrum symptoms would have high scores on many of the symptom measures and therefore have small variation in the response patterns between each symptom. Because each symptom is regressed over all others, it's likely that the shared variance between these high scores will be removed, thereby leaving some symptoms completely disconnected. It's unlikely that these symptoms are truly unrelated or disconnected from one another. Therefore, regression-based approaches that induce conditional independence via fully regressed coefficients can reduce the comparability and reproducibility between samples (Forbes et al., 2017).

The shrinkage of correlations below a certain threshold also contributes to reduced reproducibility because variables can be eliminated on the basis of statistical significance rather than theory. For example, some variables or factors may be consistently found to be weakly related but are implicated by theory to be intrinsic in the construct's conceptual hierarchy. Both fully regressed coefficients and the arbitrary selection of a threshold contain the possibility of ignoring hierarchical patterns in the data that are relevant in psychopathological and personality phenomena (Barfuss et al., 2016; Markon, Krueger, & Watson, 2005; McCrae, 2015). For instance, the WSS-SF has relations between positive and negative schizotypy that are relatively weak but are consistent with theory and are modeled in the traditional factor structure (Gross et al., 2015). Therefore, it is possible that connections between the two WSS-SF schizotypy factors would not replicate or that the factors could be arbitrarily separated, despite theoretically being bridged by the social anhedonia scale.

Despite the many different network filtering approaches that exist, so far only regression-based filtering approaches have been applied in the psychometric network literature (Epskamp et al., 2018; Fried & Cramer, 2017; van Borkulo et al., 2014). Thus, although the lasso approach is state-of-the-art, alternative filtering approaches may circumvent its potential pitfalls, such as biased comparability, reduced reproducibility, and the elimination of hierarchical information (Barfuss et al., 2016; Forbes et al., 2017; Tewarie et al., 2015). In the present research, we investigated an alternative filtering approach, *Information Filtering Networks* (IFN; Aste, Di Matteo, & Hyde, 2005; Barfuss et al., 2016; Mantegna, 1999; Massara, Di Matteo, & Aste, 2016; Tumminello, Aste, Di Matteo, & Mantegna, 2005), which is based on topological (structural) constraints. The IFN approach may be able to overcome some of the limitations that are found in the lasso and regression-based approaches.

Information filtering networks

The IFN approach applies a constraint on the structure of the network that reduces it to a subnetwork, which retains the strongest zero-order correlations from the original network. This constraint retains a specific number of connections with defined global topological properties. Two main methods of the IFN approach are the minimum spanning tree (MST; Mantegna, 1999) and the planar maximally filtered graph (PMFG; Tumminello et al., 2005). The MST sorts all edge weights (i.e., correlations) between nodes in a network in a descending order and adds the largest possible edge weight to two nodes such that all the nodes in the network are connected without forming *cliques*—a set of connected nodes. Thus, the MST method retains the most significant connections in the network so that each node has at least one connection to one other node ($n - 1$ edges). The MST has been shown to produce meaningful hierarchical structures in financial systems (Mantegna, 1999) and to be less biased when comparing within- and between-sample neural networks because it avoids some of the methodological limitations (such as those discussed earlier) of other approaches (Tewarie et al., 2015).

The PMFG method similarly sorts edge weights between nodes in a network in a descending order and adds the largest edge weight between two nodes one by one while constraining the subnetwork to be *planar*—the subnetwork can be represented in a way that no edges overlap with one another (Aste et al., 2005; Tumminello et al., 2005). This procedure reveals a subnetwork—a subset of connections ($3n - 6$ edges) from the original network that contain the most important information (correlations) from the original network. The subnetwork is composed of three- and four-node cliques (i.e., a triangle and tetrahedron, respectively). From these cliques, a nested hierarchy develops: dimensionality is reduced in a deterministic manner while retaining local information and the global hierarchical structure of the original network (Song, Di Matteo, & Aste, 2011, 2012). In psychology, the PMFG method has been applied to investigate and compare the structure of semantic memory in typical and clinical populations (Borodkin, Kenett, Faust, & Mashal, 2016; Kenett et al., 2014; Kenett, Gold, & Faust, 2016; Kenett et al., 2013).

More recently, the triangulated maximally filtered graph (TMFG; Massara et al., 2016) method was introduced. Like the PMFG method, the TMFG method filters the network by maintaining planarity, retaining $3n - 6$ edges. However, the way in which the TMFG method adds nodes and edges is more efficient, and it produces comparable or better results than the PMFG method

(Massara et al., 2016). This modified construction adds a node to the center of the three-node cliques by using a score function that maximizes the sum of the three connecting edge weights. With this addition, the three-node cliques become four-node cliques, which possess a *chord*—an edge that is not part of the clique but connects two nodes in the clique, forming two triangles, thus generating what is called a *chordal network* (Lauritzen, Speed, & Vijayan, 1984; Massara et al., 2016).

The advantage of chordal networks is that they perfectly represent the independence assumptions of Markov (i.e., bidirectional or undirected relations) and Bayesian (i.e., directional relations) networks (Koller & Friedman, 2009). This means that there exists a directed network model with conditional independencies identical to those in the undirected network as well as a representation of the edges in the network that can be reduced to a directed network (see Lauritzen & Spiegelhalter, 1988, for more technical details; Pearl, 2014). In this way, a chordal network allows a representation of the whole joint probability distribution and embeds conditional independence within the network structure (Darroch, Lauritzen, & Speed, 1980). Thus, the interpretation of the TMFG method and the lasso approach are equivalent. One limitation of chordal networks, however, is that they may add unnecessary edges in order to satisfy the chordal property (Spiegelhalter, 1987).

Both the IFN and lasso approaches aim to infer a probability distribution that maximizes the likelihood of the observations. Indeed, Barfuss et al. (2016) show that both approaches produce a multivariate normal distribution with the sparse inverse covariance matrix and their performances can be directly compared by comparing the values of their likelihoods in a training dataset to a testing dataset. The main difference between the two approaches is that the IFN aims to maximize the likelihood by retaining the largest correlations, whereas the lasso uses a penalizing term that favors sparsity in the inverse covariance matrix. Thus, IFN networks avoid a potential pitfall of the lasso approach: they use zero-order correlations, instead of fully regressed coefficients, when estimating the conditional independence structure. This means that the reliability of network measures between samples should be greater in IFN-based networks because zero-order correlations are more stable than fully regressed coefficients (Forbes et al., 2017).

Importantly, the IFN approach avoids another pitfall found in regression-based methods: reduced comparability between samples. Because the number of edges is kept constant and does not vary on the basis of sample size, network measures are less biased when attempting to compare between cross-sectional or clinical and non-clinical samples (Tewarie et al., 2015; van Dellen et al., 2015). In the TMFG method, for example, the WSS-SF network, with 60 variables, will always have 174 edges ($3n - 6$ edges). Thus, although sample size will likely affect the reliability of each network measure's estimate, differences in the network measures will be related to sample heterogeneity (e.g., psychopathological expression) rather than differences in edge density. In contrast, edge density in the lasso approach varies with sample size, so differences in network measures are more likely to be confounded with edge density and sample heterogeneity. Furthermore, in the lasso approach, edge density is typically expected to be greater in clinical samples than in nonclinical or remitting samples (van Borkulo et al., 2015), but unless the sample sizes are exactly equal, it's unclear whether the differences in edge density and network measures are due to differences in sample size, psychopathological expression, or both.

A final advantage of the IFN approach is the development of a hierarchical structure, which complements the structure of many psychopathological and personality phenomena (Markon et al., 2005; McCrae, 2015). This makes IFN-based networks ideal for determining the dimensional nature of the WSS-SF and other psychometric questionnaires because they form a conceptual hierarchy from local connections (items) to global structures (scales, phenomena). For example, items and clusters of items in the WSS-SF represent characteristics that are consistent with symptoms of schizophrenia-spectrum disorders. In this way, WSS-SF items measure overlapping symptom nuances (e.g., "Sometimes I have had feelings that I am united with an object near me"), which form symptom clusters that are representative of a single feature (e.g., uniting percepts with body boundaries) of a larger symptom (e.g., somatic delusions). Although beyond the scope of this article, some clustering techniques do outperform traditional methods of hierarchical clustering, which can be used on the IFN networks to easily extract this hierarchical information (Musmeci, Aste, & Di Matteo, 2015; Song et al., 2011, 2012; Yu et al., 2015).

The present research

One aim of the present study was to analyze the network structure of the WSS-SF and to demonstrate the feasibility of applying the IFN approach in psychometric network analysis. We first applied both lasso and IFN network filtering approaches to analyze the network structure of the WSS-SF in two large, independent samples. We sought to determine which WSS-SF scales would bridge positive and negative schizotypy in order to gain a better understanding of how the weakly related factors are linked. In line with previous research, we expect that the social anhedonia scale will be connected to a positive schizotypy scale for all networks and samples, bridging the negative and positive schizotypy factors (Kwapil et al., 2008; Lewandowski et al., 2006). On the basis of previous correlational and factor-analytic work (Gross et al., 2015; Wuthrich & Bates, 2006), we anticipate that, for all networks and samples, perceptual aberration will have the most connections to social anhedonia.

Another aim of the study was to evaluate the between-sample comparability and reproducibility of both network filtering approaches. Following Forbes et al. (2017), we assessed the comparability of the global network characteristics—ASPL, average connectivity, average degree, CC, and edge density—for both network approaches by examining the similarities and differences between each sample. Moreover, we assessed the reproducibility of the local network characteristics by determining the proportion of edges that replicate (and do not replicate) and by correlating centrality measures between each sample. On the basis of differences in edge density and the type of correlations (i.e., zero-order vs. fully regressed coefficients) used in each network filtering approach's construction, we anticipate that the IFN-based networks will have better comparability (more similar global network characteristics) and more robust reproducibility (better local network characteristics) than the lasso-based networks.

Finally, we sought to determine which items were most central to the WSS-SF's schizotypy construct and investigate whether these items were predictive of schizophrenia-spectrum symptoms. Accordingly, we identified core, intermediate, and peripheral items of Sample 1's network structure for both network filtering approaches. The core, intermediate, and peripheral schizotypy items were then used in multiple regression analyses to predict interview reports of

schizophrenia-spectrum symptoms and overall impairment from a subset of participants in Sample 1. In this way, the predictability of each filtering-based network structure was examined. If the central items identified by each approach are truly more core to the WSS-SF schizotypy phenomenon, then we expect that they should predict schizophrenia-spectrum symptoms at least as well as intermediate and peripheral items.

Table 1. Descriptive statistics for each measure and the respective samples

	Measure Descriptive Statistics					
Sample Size:	$n = 5,831$		$n = 2,171$		$n = 430$	
Statistic	M (SD)	Range	M (SD)	Range	M (SD)	Range
Physical anhedonia	2.09 (2.30)	0–14	2.26 (2.28)	0–14	2.71 (2.95)	0–14
Social anhedonia	1.77 (2.41)	0–15	1.96 (2.42)	0–15	3.01 (3.50)	0–15
Perceptual aberration	1.21 (2.28)	0–15	1.13 (2.15)	0–15	2.32 (3.49)	0–15
Magical ideation	3.26 (2.91)	0–15	3.19 (2.87)	0–15	4.30 (3.74)	0–14
Negative schizotypy total	3.86 (3.80)	0–29	4.22 (3.75)	0–24	5.72 (5.25)	0–26
Positive schizotypy total	4.47 (4.64)	0–30	4.32 (4.46)	0–30	6.61 (6.72)	0–28
Global Assessment Functioning Scale (GAFS)					73.19 (10.13)	6–91
Psychotic-Like Experiences (PLEs)					1.29 (2.07)	0–9
Negative symptoms					3.03 (4.34)	0–22
Schizoid symptoms					0.74 (1.58)	0–11
Schizotypal symptoms					1.13 (1.79)	0–11
Paranoid symptoms					0.83 (1.68)	0–12

Samples 1 and 2 had similar statistics for each schizotypy scale. The subsample had slightly higher means and standard deviations than did the two large samples

Method

Participants

Sample 1 consisted of participants ($M = 19.4$ years, $SD = 3.7$) that were obtained from a large, independent sample of undergraduate students ($n = 6,137$) from the University of North Carolina at Greensboro (UNCG), previously reported in Kwapil et al. (2008). Participants with missing data were excluded ($n = 206$), resulting in a sample consisting of 5,831 participants comparable to the original sample in terms of sex (76% female) and ethnicity (74% Caucasian, 26% African American). Sample 2 consisted of 2,171 undergraduate students from the UNCG who were primarily young adults ($M = 19.6$ years, $SD = 3.3$) reported in Gross, Silvia, Barrantes-Vidal, and

Kwapil (2012). This sample was comparable to the first sample in terms of sex (76% female) and ethnicity (69% Caucasian, 31% African American).

A subset sample of 430 participants was drawn from Sample 1 and was administered structured diagnostic interviews. These participants were either recruited through the university's research pool or were oversampled on the basis of elevated scores on the schizotypy scales, to assure adequate representation of high-schizotypy participants. This sample was comparable to Samples 1 and 2 in terms of demographics (74% female; 74% Caucasian and 26% African American) but had slightly higher schizotypy means than either sample, consistent with the oversampling procedure (Table 1).

Materials

Schizotypy assessment

Schizotypy was measured via the WSS-SF, which contains 60 true–false items measuring positive and negative schizotypy. Positive schizotypy is measured by two subscales assessing perceptual aberration and magical ideation; negative schizotypy is measured by two subscales assessing social and physical anhedonia—all scales have 15 questions each (SI 1). The data were coded in raw matrices such that each row contained all endorsements made by a participant i and each column is a different item j comprising the entire scale. Each cell is either coded as one (participant i endorsed item j) or zero (participant i did not endorse item j). After reverse-scoring as needed, all items were scored so that higher scores reflected higher levels of schizotypy.

Interview measures

The *Global Assessment of Functioning Scale* (GAFS; Endicott, Spitzer, Fleiss, & Cohen, 1976) was used to assess overall functioning of each participant ranging from marked psychopathology at the low end to superior functioning at the high end. The *Wisconsin Manual for Assessing Psychotic-Like Experiences* (L. J. Chapman & Chapman, 1980; Kwapil, Chapman, & Chapman, 1999) was used to measure the deviance of psychotic symptoms across a range of clinical and subclinical deviancy. The *Negative Symptom Manual* (Kwapil & Dickerson, 2001) measures six classes of clinical and subclinical negative symptoms of schizophrenia. Finally, modules of the *International Personality Disorders Examination* (World Health Organization, 1995) were used to provide dimensional ratings of schizoid, schizotypal, and paranoid personality disorder traits.

Network construction

IFN approach

In these networks, nodes represent the different items and edges represent endorsement associations between items—the tendency of the sample to endorse item b, given that item a is endorsed—computed with Pearson's correlations. In binary data, Pearson's correlations are equal to phi coefficients, which are related to a chi-square of a 2×2 contingency table and are interpreted as the association between two variables (Sokal & Rohlf, 1995). Phi coefficients were

used, rather than tetrachoric correlations, because the schizotypy scales are positively skewed, which violates the assumption of a bivariate normal distribution used in the calculation of tetrachoric correlations (Glass & Hopkins, 1970).

Network filtering. The TMFG method (Massara et al., 2016) was applied to construct a subnetwork, from the endorsement association matrix, that captures the most relevant information between nodes that are embedded in the original network and minimizes spurious associations. The resulting subnetwork is a clique-tree composed of four-node cliques connected with three-node cliques, and it retains a total of $3n - 6$ edges from the original network. The TMFG method begins by sorting all edge weights (i.e., the zero-order correlations) in descending order and adding the largest edge weight one by one, on the basis of an iterative construction process of a topologically constrained network (i.e., planar). In this construction, the algorithm adds a node into three-cliques, on the basis of a "T_2 move" (Aste, Gramatica, & Di Matteo, 2012; Massara et al., 2016). The T_2 move inserts a node into any three-clique's center where edges are added to it, forming a tetrahedron and keeping the network planar. When adding these nodes, the algorithm optimizes a score function that ensures the added node has the maximum increase in the sum of the additional edge weights (see SI 2 for further technical details). We computed the TMFG-filtered, weighted adjacency matrix using the NetworkToolbox package[1] (Christensen, 2018) in R (R Core Team, 2017).

lasso approach

The Ising model uses logistic regression to isolate the unique associations between all variables—that is, the association between two nodes is conditioned over all nodes in the network, leaving only the unique variance between the two nodes of reference.[2] These conditioned associations have been interpreted as potential causal pathways in previous lasso-based network models (Epskamp et al., in press); van Borkulo et al., 2014). Logistic regression coefficients are log-odds ratios and not partial correlations—thus, their interpretations are not equivalent but have been referred to as analogous in the literature, so for convenience we will refer to the logistic regression coefficients as fully regressed coefficients. Raw data for both samples were analyzed on the basis of the Ising model, using the IsingFit package (van Borkulo et al., 2014) in R.

Network filtering. For both samples, the eLasso method in the Ising model (van Borkulo et al., 2014) was applied to minimize spurious edges and increase interpretability of the network. The Ising model applies an ℓ_1-regularization penalty on the inverse covariance matrix (Ravikumar et al., 2010; van Borkulo et al., 2014), which is determined by the EBIC (Chen & Chen, 2008), to minimize spurious edges when constructing the networks. The EBIC contains a hyperparameter (γ) that controls how much the model prefers sparsity—ranging from sparse models with less connections ($\gamma > 0$) to dense models with more connections ($\gamma = 0$). In the past,

---

[1] The most up-to-date version of the NetworkToolbox package can be retrieved from https://github.com/AlexChristensen/NetworkToolbox

[2] Notably, these two network filtering approaches differ in the statistics being used (i.e., phi coefficients versus logistic regression with one node regressed over all others). Despite this, they yield the same interpretation of conditional independence. To eliminate any statistical differences between the two approaches, we have applied the glasso approach to phi coefficients. Comparable analyses (SI 4) and discussion (SI 5) are provided in the supplementary materials.

the hyperparameter has been reported to be optimal when equal to .25 (Barber & Drton, 2015; Ravikumar et al., 2010). Thus, on the basis of previous studies, for both samples we used the suggested $\gamma = .25$ setting.

Network measures

Global network characteristics

To investigate the global characteristics of the network, we examined five global network measures: ASPL, average connectivity, average degree, CC, and edge density. The *ASPL* measures the mean shortest distance between any two nodes in the network—smaller values suggest that, on average, symptoms generally are less distant from other symptoms. The *average connectivity* measures the consistency of the edge weights included in the network (De Schryver et al., 2015). This was calculated by summing the absolute sum of all edge weights included in each network and dividing it by the number of edges in the network. The *average degree*—mean number of connections each node has in the network—is computed by summing every node's degree in the network and dividing it by the number of nodes. The *CC* measures the extent in which two neighbors of a node will be neighbors themselves—that is, whether two connected nodes will both be connected to a third node. In this way, the CC represents the cliquishness of the network and indicates the tendency of a symptom to cluster with nearby symptoms. The *edge density* of a network is the proportion of edges included over the total of all possible connections $((n^2 - n)/2)$. The ASPL, average degree, CC, edge density, and average connectivity were computed using the NetworkToolbox package in R whose metrics were adapted from the Brain Connectivity Toolbox in Matlab (Rubinov & Sporns, 2010).

Centrality measures

In line with previous research, we used centrality measures to investigate the local network characteristics. These are considered local because influence is assessed for each node, whereas global measures assess the structure of the entire network. *Betweenness centrality* (BC) measures the extent a node lies on the paths between other nodes (Freeman, 1977). Therefore, items (nodes) with high betweenness values make up the most central elements or "backbone" of the network (Borgatti, 2005). *Closeness centrality* (LC) is the inverse of the average distance from all other nodes (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006; Freeman, 1977). Put simply, closeness centrality is the distance away from the most center point of the network— nodes in the middle of the network will have higher closeness centrality values than the most peripheral nodes. *Node strength* sums the edge weights of all the connections linked to a node so that a high number of connections and high edge weights will have a greater value (Barrat, Barthelemy, Pastor-Satorras, & Vespignani, 2004; Newman, 2004). Thus, node strength centrality measures the number of direct connections *and* the magnitude of those connections. *Degree* (k), the number of connections a node has, is a basic measure of a node's importance, and its distribution reveals important information about the type of network. *Eigenvector centrality* (EC) is the weighted sum of direct and indirect connections of a node and is an index of the quality of connections for each node (Bonacich & Lloyd, 2001). For example, the EC distinguishes a node of low degree that is connected to many high degree nodes and a high degree node that is connected to only low degree nodes (Bonacich, 2007). Thus,

higher EC values are given to nodes that have connections to other central nodes (van Borkulo et al., 2015). The NetworkToolbox package in R was used to compute all measures of centrality.

Statistical analyses

Representation of positive and negative schizotypy

One aim in analyzing the network structure of the WSS-SF is to analyze the network structure of the WSS-SF in two large, independent samples, by both lasso-based and IFN-based approaches, to determine which WSS-SF scales would bridge positive and negative schizotypy. To do so, we examined the number and strength of the connections between the positive and negative schizotypy scales. Moreover, we examined how many edges between the two factors replicated as well as how many nodes—that were connected by those edges—replicated between the two samples for both approaches.

Global network characteristics

ASPL, CC, and average connectivity. To evaluate the differences for the ASPL and CC, we applied independent samples t test analyses for both measures for both approaches. For the ASPL statistical analysis, we used the distance matrix—a matrix with the shortest number of paths from one node to every other node—to calculate the local shortest path length ($ASPL_i$). The $ASPL_i$ is the average distance for each node (i.e., the distance vector) to all other nodes. In this way, the mean of the $ASPL_i$ *is* the global ASPL, which allows the statistical measurement of differences in the global ASPL between samples. Similarly, we computed the local clustering coefficient (C_i), which is the average of each node's clustering coefficient vector. Again, the mean of the C_i *is* the global CC and this allows us to statistically test for differences in the global CC. A significant difference for the ASPL would suggest that on average a node is closer to (or further away from) all other nodes. A significant difference for the CC would suggest that on average a node is more (or less) connected within its local neighborhood. To test for differences in the average strength of the edge weights, we used an independent samples t test analyses on the average connectivity between the two samples for both approaches. A significant difference would suggest that one sample's network retained larger edge weights on average than the other sample's network. Previous work suggests that significant differences in edge strength might suggest clinically relevant differences between samples (van Borkulo et al., 2015). In contrast, no difference would suggest that the edge weights retained in the network are similar and that the samples are similar in psychopathological expression.

Local network characteristics

Local network characteristics were assessed by the number of edges that replicated between both samples' networks, the differences in the strength of the replicated edge weights, and the strength of linear and rank-order centrality correlations.

Edge replication. Edges replicated if they appeared in both samples. The proportion of edges that were replicated were divided by the total number of edges in the respective sample. In the lasso approach, this led to two values. One value for the proportion of edges in Sample 1 (e.g., 400

edges) and another value for the proportion of edges in Sample 2 (e.g., 248 edges). In the IFN approach, edges were constant (i.e., 174 edges), so only one edge replication value was necessary. Next, the *mean absolute edge weight difference of replicated edges* was calculated using only the edges that replicated and then difference between Sample 1 and Sample 2's replicated edges strengths were calculated. Finally, the absolute values were computed and Pearson's correlations were used to determine the similarity between the replicated edge weights (Borsboom et al., 2017). The edge replication measures were computed using the NetworkToolbox package in R.

Between-sample centrality correlations. To examine the reproducibility of centrality values and their rank-order between the samples for both approaches, Pearson's correlation and Kendall's tau-b coefficient were used. Pearson's correlations were used to evaluate the linear consistency of the centrality values, as used in Borsboom et al. (2017). The rank-order correlations were also used because previous research has suggested that centrality indices are often interpreted by their rank (Forbes et al., 2017). Standard guidelines for between-sample reliability have yet to be established in the field of psychometric network analysis.

Prediction of psychopathological interview measures

Hybrid centrality measure. To evaluate the predictability of both filtering approaches' centrality measures, we determined nodes that were most central in Sample 1's network. Sample 1's centrality measures were used to predict interview reports of impairment and schizophrenia-spectrum symptoms because the interviewed sub-sample was drawn from that sample. To assess overall centrality, we used a hybrid centrality measure (Pozzi, Di Matteo, & Aste, 2013). The hybrid centrality measure ranks nodes on the basis of their centrality values across multiple measures of centrality (BC, LC, k, EC, and NS). Based on previous work (Pozzi et al., 2013), we used weighted (w) and unweighted (u) tied rankings (sorted in descending order) of each centrality:

$$ \text{HC} = \frac{\text{BC}^w + \text{BC}^u + \text{LC}^w + \text{LC}^u + k^u + \text{EC}^w + \text{EC}^u + \text{NS}^w - 8}{8 \times (N-1)} $$

The hybrid centrality measure describes highly central nodes with large values and highly peripheral nodes with small values. This hybrid measure is not biased by greater weight given to any one centrality measurement (each centrality measure quantifies different aspects of "centralness") and provides a singular, continuous measure of overall centrality in the network. Such a measure has been shown to provide more consistent and robust results than any single centrality measure in isolation (Pozzi et al., 2013). The NetworkToolbox package in R was used to compute the hybrid centrality measure.

Multiple regression of interview symptoms. The hybrid centrality values were sorted in descending order to determine the core, intermediate, and peripheral WSS-SF items. The top five items for each scale were designated as core, the next five as intermediate, and the last five as peripheral. These breaks (i.e., one third of the items in each scale) were chosen to give an even distribution and gradation of item classifications for each scale. The core, intermediate, and peripheral positive and negative schizotypy items were summed together to create item group

totals. Finally, the positive and negative schizotypy item groups were used in multiple regression analyses to predict psychopathological impairment and several schizophrenia-spectrum symptoms.

Code and scripts

R code for all analyses performed in this study can be found in SI 6. All data and code used in this study is available on the Open Science Framework: https://osf.io/c6rqy/.

Procedure

Participants for both samples completed the Wisconsin Schizotypy Scales (WSS) as part of mass screening assessments (Gross et al., 2012; Kwapil et al., 2008). Participants who received scores of three or greater on an infrequent responding scale were omitted from the analyses (J. P. Chapman & Chapman, 1983). Then, the 60 items that are included in the WSS-SF were extracted from the original WSS. Participants provided informed consent to participate in the study and received course credit for their participation.

Participants from the Kwapil et al. (2008) subsample were administered structured recorded interviews conducted by a licensed clinical psychologist and advanced graduate students in clinical psychology (Kwapil et al., 2008), which lasted approximately 2 h. The interviewers were not aware of the participant's scores on the schizotypy questionnaires. All studies included were approved by the UNCG Institutional Review Board.

Results

Descriptive statistics for each WSS-SF scale, the positive and negative schizotypy factors, and the interview ratings are presented in Table 1. We represented the WSS-SF networks on the basis of the lasso and IFN approaches described above and applied them to both samples, and then visualized the networks using the *qgraph* package (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) in R. Then we computed the various global (ASPL, average connectivity, average degree, CC, and edge density) and local (BC, CC, NS, k, and EC) network measures. Afterward, the edge replication and correlation analyses were computed as described above. Finally, we computed the hybrid centrality measures for both approaches and used the core, intermediate, and peripheral items in multiple regression analyses, predicting interview reports of schizophrenia-spectrum symptoms and psychopathological impairment.

Representation of positive and negative schizotypy

IFN approach. In both samples, the social anhedonia scale was connected with positive schizotypy, which is consistent with traditional factor analysis findings (Gross et al., 2012, 2015). Between both networks, six edges connected social anhedonia and perceptual aberration. As predicted, social anhedonia was connected to positive schizotypy via perceptual aberration (Fig. 1). This finding is consistent with previous correlation results (Gross et al., 2012) and with the factor analysis findings using the three scales of the WSS and the Schizotypal Personality Questionnaire (Wuthrich & Bates, 2006). The average edge strengths of the six edges

between the positive and negative schizotypy factors were relatively small (Sample 1, $M =$ .13, $SD = .03$; Sample 2, $M = .15$, $SD = .02$) as compared to the average edge strengths of the networks (Sample 1, $M = .25$, and Sample 2, $M = .24$). Of the six edges, four replicated with an average absolute edge weight difference of .01 ($SD = .014$). Moreover, in both networks the same three social anhedonia items (SA02, SA07, SA08) were connected to two of the three common perceptual aberration items (PB03 and PB07), which meant that five of the seven nodes connecting between the two factors replicated (71.4%).
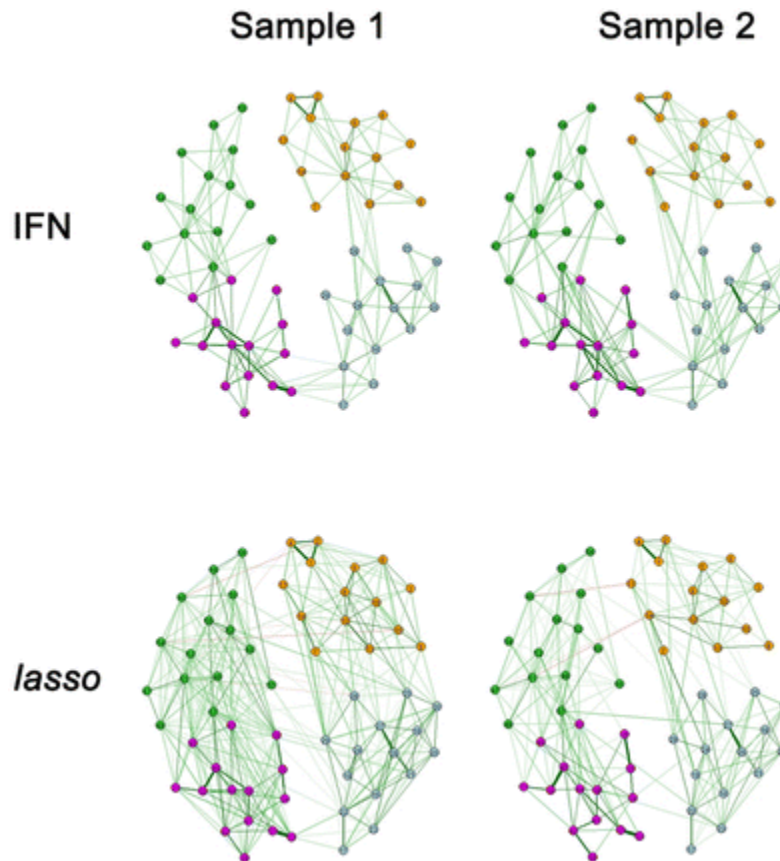

Fig. 1. A visualization comparison of the WSS-SF networks structure for the samples of the IFN-based and lasso-based methods. Nodes are identified by color and scale: physical anhedonia (orange; 1–15), social anhedonia (blue; 16–30), perceptual aberration (purple; 31–45), and magical ideation (green; 46–60). Edge thickness depicts the strength of the edge weights

lasso approach. Beginning with Sample 1, there were 27 (seven negative) connections between positive and negative schizotypy. Similar to the IFN-based networks, the strength of the bridging edges—connections between the positive and negative schizotypy factors—was relatively small ($M = .14$, $SD = .07$) as compared to the average connectivity of the network ($M = .36$). Of the 27 edges, eight (five negative) edges were between physical anhedonia and magical ideation, six edges were between social anhedonia and perceptual aberration, and the remaining 13 (two negative) edges were between social anhedonia and magical ideation. Consistent with traditional analyses, social anhedonia had most of the connections (19 total) with positive schizotypy (Gross et al., 2012; Kwapil et al., 2008). But contrary to our prediction, magical ideation had more edges connected to social anhedonia than to perceptual aberration.

In Sample 2, there were nine (three negative) connections between positive and negative schizotypy. Again, the average weight of these edges ($M = .23$, $SD = .11$) was smaller than the average connectivity of the network ($M = .44$). Of the nine edges, three negative edges were between physical anhedonia and magical ideation, two edges were between social anhedonia and perceptual aberration, and four edges were between social anhedonia and magical ideation. Similar to Sample 1, social anhedonia had the most connections to the positive schizotypy items (six total), and the magical ideation scale had the most connections to social anhedonia. Notably, none of the bridging edges from Sample 1 replicated in Sample 2. Moreover, of the 34 nodes (16 negative and 18 positive schizotypy) that were connected by these edges, only eight replicated (23.5%; SA06, SA08, SA10, MI01, MI05, MI06, MI12, and MI14).

Table 2. Summary of global and local characteristics for both network filtering approaches across all samples

	lasso		IFN	
	$n = 5,831$	**$n = 2,171$**	**$n = 5,831$**	**$n = 2,171$**
Edge density (# of edges)	.226 (400)	.140 (248)	.098 (174)	.098 (174)
ASPL	2.09	2.55	3.75	3.62
CC	.56	.49	.69	.71
Average degree	13.33	8.27	5.80	5.80
Average connectivity (*SD*)	.35 (.34)	.43 (.40)	.25 (.08)	.24 (.08)
Changes in Estimated Edges				
Proportion of replicated edges (%) [a]	204 (51)		204 (82.3)	108 (62.1)
Proportion of non-replicated edges (%) [a]	196 (49)		44 (17.7)	66 (37.9)
Mean absolute edge weight difference of replicated edges (*SD*)	.181 (.154)		.026 (.020)	
Pearson's correlation between replicated edge weights	.84***		.93***	
	Correlations		**Correlations**	
Node Centrality	**Kendall's tau**	**Pearson's**	**Kendall's tau**	**Pearson's**
Betweenness	.38***	.24†	.54***	.74***
Closeness	.45***	.59***	.66***	.84***
Node Strength	.58***	.79***	.63***	.79***
Degree	.24**	.42***	.55***	.70***
Eigenvector	.89***	.96***	.82***	.94***
Hybrid	.34***	.52***	.58***	.79***

[a]Proportion of replicated and non-replicated edges differed between samples for the Ising model because of the differences in edge densities. †$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Global characteristics

IFN approach. The IFN-based networks had identical edge densities for both samples, .098 (174 edges; Table 2). Since the edge density was constant between samples, the global characteristics of the networks were considerably consistent. The ASPL, $t(118) = 1.157$, $p = .250$, and CC, $t(118) = -0.259$, $p = .796$, did not differ significantly between the two networks. Because average degree is a function of edge density, it also did not differ between the two samples (5.80). However, the standard deviations were slightly different between the two samples (Sample 1 = 2.81, Sample 2 = 3.04). Finally, the average connectivities hardly differed between the two samples, $t(346) = 0.858$, $p = .39$, suggesting that the average edge weights included in

the networks between the two samples were highly similar and that both samples had similar psychopathological expressions.

lasso approach. The lasso-based networks had different edge densities between the samples: Sample 1 had an edge density of .226 (400 edges), and Sample 2 had an edge density of .140 (248 edges; Table 2). As expected, the differences in edge density affected the global network measures. First, the ASPL, $t(118) = -9.356, p < .001$, and the CC, $t(118) = 2.136, p = .035$, were significantly different between the two networks. This means that the ASPL significantly decreased and the CC significantly increased when more edges were retained in the network (i.e., Sample 1's network). Next, the average degrees were different between the samples, with 13.33 ($SD = 4.48$) for Sample 1 and 8.27 ($SD = 3.51$) for Sample 2. Finally, the average connectivities differed significantly between the networks, $t(646) = -2.829, p = .005$, suggesting that Sample 1's network retained smaller edge weights than did Sample 2's network. Overall, the global network characteristics for the lasso-based networks were significantly different between the samples, whereas the IFN-based networks did not differ, suggesting that the lasso-based networks had lower comparability between the samples.

Changes in estimated edges

IFN approach. Because the edge densities were equivalent between the samples, only one replication percentage was produced. Samples 1 and 2 had 108 of 174 edges replicate (62.1%). Of the edges that replicated, the mean absolute edge weight difference was relatively small, .026. This is further reflected by the high correlation between the replicated edge weights between the two samples, $r(106) = .93$. Moreover, these results support the finding that the average connectivities hardly differed.

lasso approach. Starting with Sample 1, the proportion of edges that replicated in Sample 2 was 204 of 400 edges (51%). However, the same 204 replicated edges in Sample 2 revealed 82.3% of the edges that replicated. Of the edges that replicated, we calculated the mean absolute edge weight difference between the samples (*SD*s are reported in Table 2). Between Samples 1 and 2, there was an average edge weight difference for replicated edges of .181. We also found a large correlation between the replicated edge weights, $r(202) = .84$. The IFN-based networks, however, had a numerically larger correlation for the replicated edge weights.

Node centrality

IFN approach. The between-sample centrality correlations were fairly consistent across all centrality measures (Table 2). Pearson's centrality correlations ranged from .70 for degree to .94 for eigenvector centrality. Kendall's tau centrality correlations ranged from .54 for betweenness to .82 for eigenvector centrality. For both correlation types, betweenness centrality and degree had similar correlations, whereas closeness centrality and node strength were slightly higher. Eigenvector centrality had the highest correlations of all centrality measures. Considering the large differences in sample size, we take these findings as indicating strong between-sample reliability.

lasso approach. The between-sample centrality correlations for the lasso-based networks had a larger range. Pearson's centrality correlations ranged from .24 for betweenness to .96 for eigenvector centrality (Table 2). Kendall's tau centrality correlations ranged from .24 for degree to .89 for eigenvector centrality. Eigenvector centrality was the most stable, followed by node strength centrality. The high correlations of node strength are consistent with previous findings (Epskamp et al., 2018; Fried et al., 2018). Betweenness centrality had a marginally significant Pearson's correlation ($r = .24$, $p = .071$) and a slightly larger correlation for Kendall's tau ($r = 38$). Similarly, closeness centrality's correlations were moderately related between the two samples. The results for betweenness and closeness centrality are consistent with previous findings, which suggests that they are less reliable than node strength (Epskamp et al., 2018). We found small correlations for degree, which may have been due to the difference in average degrees between the two networks. In summary, the IFN-based networks produced correlations equal to or higher than the lasso correlations across all centrality measures except for eigenvector centrality.

Predicting interview reports of schizophrenia-spectrum symptoms

To assess the predictability of the centrality of items in each network filtering approach, we applied multiple-regression analyses to predict the interview measures of psychopathological impairment and schizophrenia-spectrum symptoms for the 430 participants using the core, intermediate, and peripheral schizotypy items from Sample 1 (see SI 3 for items). Using the hybrid centrality measure, which gave us an overall centrality score, we determined the core, intermediate, and peripheral items as described above.

IFN approach. Consistent with our hypotheses, as predictors the positive and negative core schizotypy items were equal to or better than the intermediate and peripheral items for all measures of schizophrenia-spectrum pathology and overall functioning (Table 3). The positive core schizotypy items significantly predicted diminished overall functioning ($\beta = -.267$, $p = .002$), with marginal effects for positive ($\beta = -.156$, $p = .071$) and negative ($\beta = -.124$, $p = .066$) intermediate schizotypy items. For psychotic-like experiences, both the positive core ($\beta = .332$, $p < .001$) and intermediate ($\beta = .214$, $p = .008$) schizotypy items had significant effects, with positive core schizotypy items having a larger effect. Negative and schizoid symptoms were significantly predicted by the negative core ($\beta = .275$ and $\beta = .271$, respectively; both $ps < .001$) and intermediate ($\beta = .187$, $p = .003$, and $\beta = .133$, $p = .046$, respectively) schizotypy items, with negative core schizotypy items having slightly larger effects for both symptoms. Both positive ($\beta = .234$, $p = .007$) and negative ($\beta = .137$, $p = .042$) core schizotypy items were significantly associated with schizotypal symptoms. Moreover, positive intermediate schizotypy items were also significantly related to schizotypal symptoms ($\beta = .175$, $p = .039$). Finally, only positive core schizotypy items predicted paranoid symptoms ($\beta = .183$, $p = .049$). Overall, these results confirm that the IFN approach's centrality measures have predictive validity; core schizotypy items are equal to or better than the intermediate and peripheral schizotypy items for predicting schizophrenia-spectrum symptoms. Moreover, intermediate schizotypy items were more related to impairment and symptoms than are peripheral items, but to a lesser degree than the core schizotypy items.

Table 3. Multiple regressions of psychopathology and schizophrenia-spectrum symptoms on positive and negative schizotypy's core, intermediate, and peripheral item groups for the IFN-based network

IFN		Interview Criterion ($df = 423$)					
Schizotypy	Predictors (β)	GAFS	PLEs	Negative	Schizoid	Schizotypal	Paranoid
Positive	**Core**	$-.267^{**}$	$.332^{***}$	.100	.139	$.234^{**}$	$.183^{*}$
Positive	**Intermediate**	$-.156^{\dagger}$	$.214^{**}$	.055	.010	$.175^{*}$	.059
Positive	**Peripheral**	.076	.007	$-.118$	$-.093$	.014	$-.039$
Negative	**Core**	$-.076$	.064	$.275^{***}$	$.271^{***}$	$.137^{*}$	.015
Negative	**Intermediate**	$-.124^{\dagger}$	.050	$.187^{**}$	$.133^{*}$	$-.015$	.098
Negative	**Peripheral**	$-.052$	$-.062$	.076	.058	.063	.081
	Adj. R^2	$.162^{***}$	$.278^{***}$	$.247^{***}$	$.185^{***}$	$.184^{***}$	$.059^{***}$

All beta coefficients are standardized; GAFS, Global Adjustment Functioning Scale; PLEs, psychotic-like experiences. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

lasso approach. The positive core schizotypy items performed consistent with our predictions, associating with impairment and nearly all schizophrenia-spectrum symptoms except for negative symptoms (Table 4). Contrary to our predictions, however, the negative core schizotypy items were not as effective as the negative intermediate and peripheral schizotypy item groups at predicting some schizophrenia-spectrum symptoms. Both positive ($\beta = -.275$, $p = .001$) and negative ($\beta = -.162$, $p = .017$) core schizotypy items were related to overall psychopathological impairment. Psychotic-like experiences were predicted by positive core ($\beta = .433$, $p < .001$) and intermediate ($\beta = .156$, $p = .038$) schizotypy items. All negative schizotypy item groups were positively related to negative symptoms, with the effect sizes descending from peripheral ($\beta = .201$, $p = .002$) to intermediate ($\beta = .185$, $p = .004$) to core ($\beta = .166$, $p = .010$) items. Unexpectedly, positive peripheral schizotypy items were negatively related with negative symptoms ($\beta = -.144$, $p = .044$). Schizoid symptoms were positively related to positive core ($\beta = .194$, $p = .021$), negative intermediate ($\beta = .220$, $p = .001$), and negative peripheral ($\beta = .177$, $p = .009$) schizotypy items. Notably, negative core schizotypy items did not predict schizoid symptoms ($\beta = .073$, $p = .272$). Similar to the results for negative symptoms, positive peripheral schizotypy items were negatively associated with schizoid symptoms ($\beta = -.166$, $p = .025$). Positive core ($\beta = .277$, $p = .001$) and intermediate ($\beta = .168$, $p = .035$) schizotypy items predicted schizotypal symptoms, with positive core items having a larger effect. Finally, paranoid symptoms were related to positive ($\beta = .232$, $p = .010$) and negative ($\beta = .173$, $p = .015$) core schizotypy items, as well as marginally related to positive intermediate items ($\beta = .153$, $p = .072$). Surprisingly, positive peripheral schizotypy items had a significant negative relationship with paranoid symptoms ($\beta = -.195$, $p = .013$).

Overall, the lasso results were mixed. On the one hand, the positive schizotypy item groups for the lasso-based network had the expected predictive distinctions of overall centrality—that is, between core, intermediate, and peripheral items. On the other hand, the negative schizotypy item groups exhibited the opposite distinctions for the expected relationships with the negative schizophrenia-spectrum symptoms.

Table 4. Multiple regressions of psychopathology and schizophrenia-spectrum symptoms on positive and negative schizotypy's core, intermediate, and peripheral item groups for the lasso-based network

lasso		Interview Criterion (df = 423)					
Schizotypy	Predictors (β)	GAFS	PLEs	Negative	Schizoid	Schizotypal	Paranoid
Positive	**Core**	− .275***	.433***	.068	.194*	.277***	.232**
Positive	**Intermediate**	− .017	.156*	.123	.038	.168*	.153†
Positive	**Peripheral**	− .053	− .030	− .144*	− .166*	− .012	− .195*
Negative	**Core**	− .162*	− .056	.166**	.073	.008	.173*
Negative	**Intermediate**	− .071	.080	.185**	.220***	.105	− .005
Negative	**Peripheral**	− .033	.040	.201**	.177**	.080	.028
	Adj. R^2	.156***	.285***	.246***	.192***	.185***	.082***

All beta coefficients are standardized; GAFS, Global Adjustment Functioning Scale; PLEs, psychotic-like experiences. †p < .10; *p < .05; **p < .01; ***p < .001

Discussion

In the present study, we conducted the first network analysis of the WSS-SF to investigate the underlying structure of its multidimensional schizotypy continuum. This was achieved by applying two different network methodologies: one popular approach in psychometric network analysis (lasso; Epskamp et al., 2018), and an alternative approach (IFN) that has been applied in cognitive (Borodkin et al., 2016; Kenett, Beaty, Silvia, Anaki, & Faust, 2016; Kenett et al., 2013), neural (Tewarie et al., 2015; van Dellen et al., 2015), and financial market networks (Massara et al., 2016; Tumminello et al., 2005). To evaluate the IFN approach, we compared its performance with the current state-of-the-art approach, the lasso approach (Epskamp et al., in press); van Borkulo et al., 2014). Each approach was evaluated on its representation of the WSS-SF's schizotypy continuum and assessed on its between-sample performance on the basis of global and local network characteristics. To examine the predictability of the network structure produced by each approach, we identified core, intermediate, and peripheral items of positive and negative schizotypy, which were used to predict interview reports of schizophrenia-spectrum symptoms.

Overall, we found that both filtering approaches had network connections that were consistent with traditional findings. The IFN approach, however, produced more reliable and theoretically consistent results. Moreover, we found the IFN approach was more consistent in reproducing global and local characteristics of the WSS-SF networks between samples. Finally, both approaches had strong predictive validity for the positive schizotypy continuum, but the IFN-based network had better predictability for the negative schizotypy continuum, too. Thus, we conclude that the network structure of the IFN-based network had better overall predictability than the lasso-based network.

Representation of positive and negative schizotypy

Both network filtering approaches provided results that were consistent with traditional findings—social anhedonia bridged positive and negative schizotypy factors (Gross et al., 2012; Lewandowski et al., 2006). The approaches differed, however, on which positive scale had the most connections to social anhedonia. The IFN-based networks revealed that perceptual

aberration had the most connections to social anhedonia, which is consistent with previous correlational findings (Gross et al., 2012). Moreover, these results also support past theoretical schizotypy interpretations, which suggest that intrapersonal body distortions affect interpersonal social relationships (Wuthrich & Bates, 2006). The lasso-based networks revealed that magical ideation had the most connections to social anhedonia. Magical ideation has also been shown to be moderately related to social anhedonia but to a lesser degree than perceptual aberration (Gross et al., 2012). Therefore, these results are not incompatible and may in fact complement each other. A possible explanation for why social anhedonia had more connections with different positive schizotypy scales could be that the perceptual aberration scale has stronger zero-order correlations, but when items are conditioned over all other items, the common variance shared by the perceptual aberration correlations is removed and magical ideation items are more uniquely related.

Notably, the zero-order correlations for perceptual aberration in the IFN-based networks were consistent between the two samples, but the unique associations (i.e., fully regressed coefficients) in the lasso-based networks were less reliable and did not replicate. The poor reproducibility of bridging connections has been noted in previous research using traditional conditional independence networks (Forbes et al., 2017). This suggests that bridging edges in the lasso-based networks might be unreliable. One reason for this might be that the zero-order correlations of the bridging edges are typically small (as shown by our results), so conditioning them increases measurement-error and makes them prone to arbitrary retention and removal.

To further advance the understanding of the schizotypy continuum, future work should consider using network analysis on multiple questionnaires like previous factor analysis has done (Wuthrich & Bates, 2006). Such a network could use questionnaires, such as the WSS-SF, the Multidimensional Schizotypy Scale (Kwapil et al., 2017), the Schizotypal Personality Questionnaire (Raine & Benishay, 1995), and the Oxford–Liverpool Inventory of Feelings and Experiences (Mason, Claridge, & Jackson, 1995), to form a nomological network (Cronbach & Meehl, 1955) of schizotypy (Meehl, 1962). The combination of these scales could further develop the measurement of schizotypy by increasing the definitiveness and distinctiveness of its dimensions. Moreover, the conceptual hierarchy produced by the IFN approach's clique structure would be particularly useful for investigating the hierarchical structure of the schizotypy nomological network.

Between-sample network characteristics

Overall, the IFN approach was more consistent across global network characteristics and nearly all local centrality measures. This was expected because the IFN approach constrains the networks to an equal edge density, which maintains a similar network structure and ensures network comparability across independent samples. In contrast, the lasso approach adapts the edge density on the basis of sample size, which alters the network structure. For instance, the larger edge density in the lasso-based Sample 1's network led to a larger CC, smaller ASPL, greater average degree, and smaller average connectivity. In comparison, the IFN-based networks had no such differences. The significant differences in ASPL and CC of the lasso-based networks suggest that nodes have different potentials to influence their neighbors and other nodes in the network. This significantly alters the interpretation of what a node's potential

influence could be. Moreover, the significant difference of average edge strength might be an effect of including more edges that are smaller (rather than clinically relevant differences between samples) because the sample size suggests that these edges are no longer considered false positives. Thus, although the lasso-based network's edge densities are adapted to produce fewer false positives in the specific samples, the comparability and reproducibility of the network measures between samples are influenced.

The edge replication proportions, for instance, are difficult to compare for cross-sectional lasso-based networks. As we show, cross-sectional samples with large sample size differences can skew and complicate the true proportion of edges that replicate. On the one hand, if a researcher uses the largest sample as the baseline model, like past research has done (Forbes et al., 2017), then the edge reproducibility between samples becomes inflated because there are more edges on the whole to replicate. On the other hand, comparing a smaller sample to a larger sample likely underestimates the edge replication proportion. The IFN approach, however, produces a single measurement, which makes the edge replication proportion comparable and direct. In addition, despite the lasso approach allowing fewer false positives, both lasso-based networks had a greater number of edges than the IFN-based networks. Thus, the IFN-based networks were more parsimonious (i.e., fewer number of edges) and had better global network comparability.

Global network differences may have also affected the reproducibility of local network measures (i.e., centrality) in the lasso-based networks. For example, there were low correlations for betweenness centrality, closeness centrality, and degree. Previous research using the lasso approach has continuously found that betweenness and closeness centrality have low correlations and reliability (Epskamp et al., 2018; Forbes et al., 2017). It's likely that the low reliability of these measures is due to variations in the global network structure (e.g., varying ASPL values). Notably, the consistency of the global network structure in the IFN-based networks seemed to improve the reproducibility of the centrality measures. The IFN approach had numerically larger between-sample centrality correlations for all measures except for eigenvector centrality. The zero-order correlations of the IFN approach also probably contributed to the higher centrality reproducibility because they included less measurement-error. Therefore, centrality reproducibility in the lasso-based networks likely suffered from two main issues: differences in edge density and fully regressed coefficients. In general, the centrality correlations for both approaches were smaller than previous findings (Borsboom et al., 2017), which was expected due to the large differences in sample size.

Finally, an important finding was that eigenvector centrality had very high between-sample correlations for both network filtering approaches. In a growing field that is concerned about the reliability and reproducibility of networks (Borsboom et al., 2017; Epskamp et al., 2018; Forbes et al., 2017; Fried & Cramer, 2017), the consistency of the eigenvector centrality suggests that it should be included with the other centrality measures. The EC is proportional to the sum of the centrality values of the nodes that it is connected to, which means it is a decent singular measure of overall centrality. Moreover, the EC has important interpretations: it is related to the dimensional structure of the network (Bonacich, 1972) and nodes high in EC have high quality connections, meaning they have a high potential for influence in the network. Because of this, the eigenvector centrality should strongly be considered as a staple centrality measure for future psychometric network analysis.

WSS-SF schizotypy continuum

For the IFN-based network, the interview results revealed that items classified as core items were more strongly associated with impaired functioning and schizophrenia-spectrum symptoms than intermediate and peripheral items. Furthermore, negative schizotypy intermediate items were more related to the negative symptoms and schizoid symptoms than peripheral items. Comparatively, the lasso-based network had similar findings for positive schizotypy, with positive core items having equal to or better associations with impairment and schizophrenia-spectrum symptoms. The negative schizotypy findings, however, were contrary to our expectations: negative intermediate and peripheral schizotypy items had larger effects for negative symptoms and schizoid symptoms than the negative core schizotypy items. These two symptoms are traditionally associated with negative schizotypy (Gross et al., 2012; Kwapil et al., 2008), which means that negative core schizotypy items should have had stronger effects. This suggests that the lasso-based network did not provide as strong of inferential differentiation of overall centrality within its structure for negative schizotypy items when compared to the IFN-based networks.

The IFN-based network distinctions establish a richer conceptualization of the WSS-SF's schizotypy continuum, which provides clearer links to the liability of schizophrenia-spectrum disorders. Our findings indicate that core items have greater clinical relevance than intermediate and peripheral items, and some intermediate items are more related to clinical symptoms than peripheral items. This gradation of the WSS-SF's schizotypy phenomenon increases the specificity of its schizotypy continuum, which is useful for detecting early signs of schizophrenia-spectrum liability before the onset of disorder (Kwapil & Barrantes-Vidal, 2015).

Limitations

The first limitation, pointed out by Guloksuz et al. (2017), pertains to the latent constructs already embedded in the WSS-SF's schizotypy construct. Schizotypy is based on DSM criteria, which confines the understanding of schizophrenia-spectrum disorders to symptoms that are already known. In this way, our network analysis largely supports previous knowledge. Although we extend this knowledge of the WSS-SF's schizotypy continuum by determining specific items that are more relevant to clinical symptoms, future research should expand the network analyses of schizotypy to include behavioral and cognitive measurements. For example, schizotypy has been linked to depression, anxiety, personality, executive control, and memory deficits (Kane et al., 2016; Lewandowski et al., 2006; Sahakyan & Kwapil, 2016). A network including these measures would provide a more holistic perspective for determining the liability of schizophrenia-spectrum disorders.

Next, although the TMFG method of the IFN approach performed well, the three- and four-clique structure imposed by the TMFG method might not be suitable for all types of psychometric networks. Moreover, all of the IFN methods impose a specific network structure (i.e., planar or a tree) on the data. This limitation is not unique to the IFN approach—it holds for most network estimation methods (Epskamp, Kruis, & Marsman, 2017)—but it strongly applies here. The imposed structures of the IFN approach could produce too few or too many

connections than what are in the *true* network structure. On the one hand, the tree and planar constraints are prone to limiting the connections of densely connected nodes by allowing only edges that keep the network a tree or planar. On the other hand, spurious connections may be artificially retained because of the cliques maintained in the PMFG and TMFG structures. However, for the TMFG method specifically, artificial edges may be necessary to maintain its chordality property (Spiegelhalter, 1987). Nevertheless, the IFN approach is not limited to tree (MST) or planar (PMFG and TMFG) structures, which means a variety of networks with different structural properties can be constructed by this approach (Aste et al., 2005). Furthermore, future methodological advancements could be applied to identify the reliability of edges kept in the network (Tumminello, Coronnello, Lillo, Micciche, & Mantegna, 2007) and to reduce the number of false positives included in the IFN networks.

Finally, our dataset was limited to cross-sectional samples, which means our analyses were restricted to the group-level and to a single time point. Despite our findings that the most central nodes in the network predict increased liability for schizophrenia-spectrum symptoms, it does not mean that this holds across individuals. Similarly, singular time points are ineffective at detecting the progression of schizophrenia-spectrum liability over time, which would be valuable for prevention and intervention. Experience sampling, for example, would provide a perspective of schizotypy in daily life, which could offer a developmental time course for future liability and decomposition into disorder (Chun, Barrantes-Vidal, Sheinbaum, & Kwapil, 2017). Therefore, we echo the call for future experimental designs to gather this type of data to provide more detailed insights into the progression of schizophrenia-spectrum liability at the individual-level (Borsboom, 2017; Fried & Cramer, 2017; Guloksuz et al., 2017; Wichers et al., 2017).

Conclusions

In summary, our findings provide improved insight into the connections between the WSS-SF's schizotypy factors and define a more specific continuum of WSS-SF items that are clinically relevant for schizophrenia-spectrum liability. We also examined two network filtering approaches and found that the IFN-based networks revealed more consistent, parsimonious, and predictive results. Moreover, we established the feasibility of the IFN approach by demonstrating its ability to provide less biased comparison and greater reproducibility in cross-sectional samples than the lasso-based networks. Thus, the IFN approach provides an alternative network approach in psychometric network analysis and contributes to the ongoing discussion of network reproducibility in this field.

Author note

References

Aste, T., Di Matteo, T., & Hyde, S. (2005). Complex networks on hyperbolic surfaces. *Physica A*, *346*, 20–26.

Aste, T., Gramatica, R., & Di Matteo, T. (2012). Exploring complex networks via topological embedding on surfaces. *Physical Review E*, *86*, 036109.

Barabási, A.-L. (2016). Network science. Cambridge: Cambridge University Press.

Barber, R. F., & Drton, M. (2015). High-dimensional Ising model selection with Bayesian information criteria. *Electronic Journal of Statistics*, *9*, 567–607.

Barfuss, W., Massara, G. P., Di Matteo, T., & Aste, T. (2016). Parsimonious modeling with information filtering networks. *Physical Review E*, *94*, 062306.

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*, 348–360.

Barrantes-Vidal, N., Gross, G. M., Sheinbaum, T., Mitjavila, M., Ballespí, S., & Kwapil, T. R. (2013). Positive and negative schizotypy are associated with prodromal and schizophrenia-spectrum symptoms. *Schizophrenia Research*, *145*, 50–55.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, *101*, 3747–3752.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*, 175–308.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, *2*, 113–120.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*, 555–564.

Bonacich, P., & Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, *23*, 191–201.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, *27*, 55–71.

Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition*, *156*, 60–70.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*, 5–13.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121.

Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L. J., & Cramer, A. O. J. (2017). False alarm? A comprehensive reanalysis of "Evidence that psychopathology symptom networks have limited replicability" by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, *126*, 989–999. doi: https://doi.org/10.1037/abn0000306

Boschloo, L., Schoevers, R. A., van Borkulo, C. D., Borsboom, D., & Oldehinkel, A. J. (2016). The network structure of psychopathology in a community sample of preadolescents. *Journal of Abnormal Psychology*, *125*, 599–606.

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory–II. *Psychological Medicine*, *45*, 747–757.

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., … Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, *8*, e60188. doi: https://doi.org/10.1371/journal.pone.0060188

Chapman, J. P., & Chapman, L. J. (1983). Reliability and the discrimination of normal and pathological groups. *Journal of Nervous and Mental Disease*, *171*, 658–661.

Chapman, L. J., & Chapman, J. P. (1980). Scales for rating psychotic and psychotic-like experiences as continua. *Schizophrenia Bulletin*, *6*, 476–489.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of Abnormal Psychology*, *85*, 374–382.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-image aberration in schizophrenia. *Journal of Abnormal Psychology*, *87*, 399–407.

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*, 759–771.

Christensen, A.P. (2018). NetworkToolbox: Methods and measures for brain, cognitive, and psychometric network analysis (R package version 1.0.0). Available from https://github.com/AlexChristensen/NetworkToolbox

Chun, C. A., Barrantes-Vidal, N., Sheinbaum, T., & Kwapil, T. R. (2017). Expression of schizophrenia-spectrum personality traits in daily life. *Personality Disorders: Theory, Research, and Treatment*, *8*, 64–74.

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2017). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences* doi: https://doi.org/10.1016/j.paid.2017.06.011

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Danaher, P., Wang, P., & Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society: Series B*, *76*, 373–397.

Darroch, J. N., Lauritzen, S. L., & Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *Annals of Statistics*, *8*, 522–539.

De Schryver, M., Vindevogel, S., Rasmussen, A. E., & Cramer, A. O. J. (2015). Unpacking constructs: A network approach for studying war exposure, daily stressors and post-traumatic stress disorder. *Frontiers in Psychology*, *6*, 1896:1–10. doi: https://doi.org/10.3389/fpsyg.2015.01896

Eckblad, M. L., & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology*, *51*, 215–225.

Eckblad, M. L., Chapman, L. J., Chapman, J. P., & Mishlove, M. (1982). *The Revised Social Anhedonia Scale*. Unpublished test.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, *33*, 766–771.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 195–212. doi: https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*, 1–18.

Epskamp, S., & Fried, E. I. (2016). *A tutorial on regularized partial correlation networks*. arXiv preprint. arXiv:1607.01367

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

Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (in press). Network psychometrics. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development. New York: Wiley.

Ettinger, U., Meyhöfer, I., Steffens, M., Wagner, M., & Koutsouleris, N. (2014). Genetics, cognition, and neurobiology of schizotypal personality: A review of the overlap with schizophrenia. *Frontiers in Psychiatry*, *5*, 1–18.

Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology*, *126*, 969–988. doi: https://doi.org/10.1037/abn0000276

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*, 35–41.

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

Fried, E. I., Eidhof, M. B., Palic, S., Costantini, G., Huisman-van Dijk, H. M., Bockting, C. L. H., … Karstoft, K.-I. (2018). Replicability and generalizability of posttraumatic stress disorder (PTSD) networks: A cross-cultural multisite study of PTSD symptoms in four trauma patient samples. *Clinical Psychological Science*. Advance online publication. doi: https://doi.org/10.1177/2167702617745092

Fried, E. I., van Borkulo, C. D., 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*, 1–10. doi: 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*, 432–441.

Glass, G. V., & Hopkins, K. D. (1970). Statistical methods in education and psychology. Boston: Aliyn & Bacon.

Gross, G. M., Mellin, J., Silvia, P. J., Barrantes-Vidal, N., & Kwapil, T. R. (2014). Comparing the factor structure of the Wisconsin Schizotypy Scales and the Schizotypal Personality Questionnaire. *Personality Disorders: Theory, Research, and Treatment*, *5*, 397–405.

Gross, G. M., Silvia, P. J., Barrantes-Vidal, N., & Kwapil, T. R. (2012). Psychometric properties and validity of short forms of the Wisconsin Schizotypy Scales in two large samples. *Schizophrenia Research*, *134*, 267–272.

Gross, G. M., Silvia, P. J., Barrantes-Vidal, N., & Kwapil, T. R. (2015). The dimensional structure of short forms of the Wisconsin Schizotypy Scales. *Schizophrenia Research*, *166*, 80–85.

Guloksuz, S., Pries, L. K., & van Os, J. (2017). Application of network methods for understanding mental disorders: Pitfalls and promise. *Psychological Medicine*, *47*, 2743–2752. doi: https://doi.org/10.1017/S0033291717001350

Isvoranu, A.-M., van Borkulo, C. D., Boyette, L.-L., Wigman, J. T. W., Vinkers, C. H., & Borsboom, D. (2017). A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms. *Schizophrenia Bulletin*, *1*, 187–196.

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General*, *145*, 1017–1048. doi: https://doi.org/10.1037/xge0000184

Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. (2016). Local patterns to global architectures: Influences of network topology on human learning. *Trends in Cognitive Sciences*, *20*, 629–640.

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, *8*, 407:1–16. doi: https://doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., Beaty, R. E., Silvia, P. J., Anaki, D., & Faust, M. (2016). Structure and flexibility: Investigating the relation between the structure of the mental lexicon, fluid intelligence, and creative achievement. *Psychology of Aesthetics, Creativity, and the Arts*, *10*, 377–388.

Kenett, Y. N., Gold, R., & Faust, M. (2016). The hyper-modular associative mind: A computational analysis of associative responses of persons with Asperger Syndrome. *Language and Speech*, *59*, 297–317.

Kenett, Y. N., Wechsler-Kashi, D., Kenett, D. Y., Schwartz, R. G., Ben Jacob, E., & Faust, M. (2013). Semantic organization in children with cochlear implants: Computational analysis of verbal fluency. *Frontiers in Psychology*, *4*, 543:1–11. doi: https://doi.org/10.3389/fpsyg.2013.00543

Keshavan, M. S., Nasrallah, H. A., & Tandon, R. (2011). Moving ahead with the schizophrenia concept: From the elephant to the mouse. *Schizophrenia Research*, *127*, 3–13.

Koller, D., & Friedman, N. (2009). Probabilistic graphical models: Principles and techniques. Cambridge: MIT Press.

Kwapil, T. R., & Barrantes-Vidal, N. (2015). Schizotypy: Looking back and moving forward. *Schizophrenia Bulletin*, *41*, S366–S373.

Kwapil, T. R., Barrantes-Vidal, N., & Silvia, P. J. (2008). The dimensional structure of the Wisconsin schizotypy scales: Factor identification and construct validity. *Schizophrenia Bulletin*, *34*, 444–457.

Kwapil, T. R., Chapman, L. J., & Chapman, J. P. (1999). Validity and usefulness of the Wisconsin Manual for assessing psychotic-like experiences. *Schizophrenia Bulletin*, *25*, 363–375.

Kwapil, T. R., & Dickerson, L. A. (2001) The assessment of negative symptoms in psychosis-prone young adults. Paper presented at the International Congress of Schizophrenia Research, Whistler.

Kwapil, T. R., Gross, G. M., Silvia, P. J., Raulin, M. L., & Barrantes-Vidal, N. (in press). Development and psychometric properties of the Multidimensional Schizotypy Scale: A new measure for assessing positive, negative, and disorganized schizotypy. *Schizophrenia Research*. Advance online publication. doi: https://doi.org/10.1016/j.schres.2017.07.001

Kwapil, T. R., Ros-Morente, A., Silvia, P. J., & Barrantes-Vidal, N. (2012). Factor invariance of psychometric schizotypy in Spanish and American samples. *Journal of Psychopathology and Behavioral Assessment*, *34*, 145–152.

Lauritzen, S., & Spiegelhalter, D. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society: Series B*, *50*, 157–224.

Lauritzen, S. L., Speed, T. P., & Vijayan, K. (1984). Decomposable graphs and hypergraphs. *Journal of the Australian Mathematical Society*, *36*, 12–29.

Lewandowski, K. E., Barrantes-Vidal, N., Nelson-Gray, R. O., Clancy, C., Kepley, H. O., & Kwapil, T. R. (2006). Anxiety and depression symptoms in psychometrically identified schizotypy. *Schizophrenia Research*, *83*, 225–235.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *European Physics Journal*, *11*, 193–197.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139–157.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences*, *18*, 7–13.

Massara, G. P., Di Matteo, T., & Aste, T. (2016). Network filtering for big data: Triangulated maximally filtered graph. *Journal of Complex Networks*, *5*, 161–178.

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*, 97–112.

McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems a network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*, 836–849.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, *17*, 827–838.

Molinelli, E. J., Korkut, A., Wang, W., Miller, M. L., Gauthier, N. P., Jing, X., … Solit, D. B. (2013). Perturbation biology: Inferring signaling networks in cellular systems. *PLoS Computational Biology*, *9*, e1003290.

Musmeci, N., Aste, T., & Di Matteo, T. (2015). Relation between financial market structure and the real economy: Comparison between clustering methods. *PLoS ONE*, *10*, e0116201. doi: https://doi.org/10.1371/journal.pone.0116201

Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E*, *70*, 056131.

Pearl, J. (2014). Probabilistic reasoning in intelligent systems: Networks of plausible inference. San Francisco: Morgan Kaufmann.

Pozzi, F., Di Matteo, T., & Aste, T. (2013). Spread of risk across financial markets: Better to invest in the peripheries. *Scientific Reports*, *3*, 1665:1–7.

R Core Team (2017). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Raine, A. (1991). The SPQ: A scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia Bulletin*, *17*, 555–564.

Raine, A., & Benishay, D. (1995). The SPQ-B: A brief screening instrument for schizotypal personality disorder. *Journal of Personality Disorders*, *9*, 346–355.

Ravikumar, P., Wainwright, M. J., & Lafferty, J. D. (2010). High-dimensional Ising model selection using $\ell 1$-regularized logistic regression. *Annals of Statistics*, *38*, 1287–1319.

Ravikumar, P., Wainwright, M. J., Raskutti, G., & Yu, B. (2011). High-dimensional covariance estimation by minimizing $\ell 1$-penalized log-determinant divergence. *Electronic Journal of Statistics*, *5*, 935–980.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*, 1059–1069.

Sahakyan, L., & Kwapil, T. R. (2016). Positive schizotypy and negative schizotypy are associated with differential patterns of episodic memory impairment. *Schizophrenia Research: Cognition*, *5*, 35–40.

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*, 43–53.

Sokal, R. R., & Rohlf, F. J. (1995). Biometry. New York: WH Freeman.

Song, W.-M., Di Matteo, T., & Aste, T. (2011). Nested hierarchies in planar graphs. *Discrete Applied Mathematics*, *159*, 2135–2146.

Song, W.-M., Di Matteo, T., & Aste, T. (2012). Hierarchical information clustering by means of topologically embedded graphs. *PLoS ONE*, *7*, e31929. doi: https://doi.org/10.1371/journal.pone.0031929

Spiegelhalter, D. J. (1987). Coherent evidence propagation in expert systems. *Statistician*, 201–210.

Stam, C. J., Tewarie, P., Van Dellen, E., van Straaten, E. C. W., Hillebrand, A., & Van Mieghem, P. (2014). The trees and the forest: Characterization of complex brain networks with minimum spanning trees. *International Journal of Psychophysiology*, *92*, 129–138.

Tewarie, P., van Dellen, E., Hillebrand, A., & Stam, C. J. (2015). The minimum spanning tree: An unbiased method for brain network analysis. *NeuroImage*, *104*, 177–188.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 267–288.

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, *102*, 10421–10426.

Tumminello, M., Coronnello, C., Lillo, F., Micciche, S., & Mantegna, R. N. (2007). Spanning trees and bootstrap reliability estimation in correlation-based networks. *International Journal of Bifurcation and Chaos*, *17*, 2319–2329.

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*, 5918:1–10.

van Borkulo, C. D., 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*, 1219–1226.

van Dellen, E., Bohlken, M. M., Draaisma, L., Tewarie, P. K., van Lutterveld, R., Mandl, R., Stam, C. J., & Sommer, I. E. (2015). Structural brain network disturbances in the psychosis spectrum. *Schizophrenia Bulletin*, *42*, 782–789.

van Wijk, B. C. M., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE*, *5*, e13701. doi: https://doi.org/10.1371/journal.pone.0013701

Wichers, M., Wigman, J. T. W., Bringmann, L. F., & de Jonge, P. (2017). Mental disorders as networks: Some cautionary reflections on a promising approach. *Social Psychiatry and Psychiatric Epidemiology*, *52*, 143–145.

Winterstein, B. P., Silvia, P. J., Kwapil, T. R., Kaufman, J. C., Reiter-Palmon, R., & Wigert, B. (2011). Brief assessment of schizotypy: Developing short forms of the Wisconsin Schizotypy Scales. *Personality and Individual Differences*, *51*, 920–924.

World Health Organization. (1995). International personality disorders examination. Geneva: WHO.

Wuthrich, V. M., & Bates, T. C. (2006). Confirmatory factor analysis of the three-factor structure of the schizotypal personality questionnaire and Chapman schizotypy scales. *Journal of Personality Assessment*, *87*, 292–304.

Yu, M., Hillebrand, A., Tewarie, P., Meier, J., van Dijk, B., Van Mieghem, P., & Stam, C. J. (2015). Hierarchical clustering in minimum spanning trees. *Chaos*, *25*, 023107.

Zhou, Y. (2011). *Structure learning of probabilistic graphical models: A comprehensive survey*. arXiv preprint. arXiv:1111.6925