Journal of Research in Personality 54 (2015) 13–29

Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

State of the aRt personality research: A tutorial on network analysis of personality data in R

Giulio Costantini [a,*,1], Sacha Epskamp [b,1], Denny Borsboom [b], Marco Perugini [a], René Mõttus [c,d], Lourens J. Waldorp [b], Angélique O.J. Cramer [b]

[a] Department of Psychology, University of Milan-Bicocca, Piazza dell'Ateneo Nuovo 1 (U6), 20126 Milan, Italy
[b] Department of Psychological Methods, University of Amsterdam, Weesperplein 4, 1018 XA Amsterdam, The Netherlands
[c] Department of Psychology, University of Edinburgh, George Square 7, EH8 9JZ Edinburgh, Scotland, UK
[d] Department of Psychology, University of Tartu, Näituse 2, 50409 Tartu, Estonia

ARTICLE INFO

Article history:
Available online 15 July 2014

Keywords:
Network analysis
Psychometrics
Latent variables
Centrality
Clustering
Personality traits
HEXACO

ABSTRACT

Network analysis represents a novel theoretical approach to personality. Network approaches motivate alternative ways of analyzing data, and suggest new ways of modeling and simulating personality processes. In the present paper, we provide an overview of network analysis strategies as they apply to personality data. We discuss different ways to construct networks from typical personality data, show how to compute and interpret important measures of centrality and clustering, and illustrate how one can simulate on networks to mimic personality processes. All analyses are illustrated using a data set on the commonly used HEXACO questionnaire using elementary R-code that readers may easily adapt to apply to their own data.

0. Introduction

A network is an abstract model composed of a set of nodes or vertices, a set of edges, links or ties that connect the nodes, together with information concerning the nature of the nodes and edges (e.g., De Nooy, Mrvar, & Batagelj, 2011). Fig. 1 reports the example of a simple network, with six nodes and seven edges. The nodes usually represent entities and the edges represent their relations. This simple model can be used to describe many kinds of phenomena, such as social relations, technological and biological structures, and information networks (e.g., Newman, 2010, Chapters 2–5). Recently networks of relations among thoughts, feelings and behaviors have been proposed as models of personality and of psychopathology: in this framework, traits have been conceived of as emerging phenomena that arise from such networks (Borsboom & Cramer, 2013; Cramer et al., 2012a; Schmittmann et al., 2013). An R package, *qgraph*, has been developed for the specific purpose of analyzing personality and psychopathology data (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012).

The aim of this contribution is to provide the reader with the necessary theoretical and methodological tools to analyze personality data using network analysis, by presenting key network concepts, instructions for applying them in R (R Core Team, 2013), and examples based on simulated and on real data. First, we show how a network can be defined from personality data. Second, we present a brief overview of important network concepts. Then, we discuss how network concepts can be applied to personality data using R. In the last part of the paper, we outline how network-based simulations can be performed that are specifically relevant for personality psychology. Both the data and the R code are available for the reader to replicate our analyses and to perform similar analyses on his/her own data.

1. Constructing personality networks

A typical personality data set consists of cross-sectional measures of multiple subjects on a set of items designed to measure several facets of personality. In standard approaches in personality research, such data are used in factor analysis to search for an underlying set of latent variables that can explain the structural covariation in the data. In a causal interpretation of latent variables (Borsboom, Mellenbergh, & van Heerden, 2003), responses to items such as "I like to go to parties" and "I have many friends" are viewed as being causally dependent on a latent variable (e.g., extraversion). For example, McCrae and Costa's (2008) interpretation of the relation between extraversion and its indicators is explicitly causal: "extraversion causes party-going behavior in

* Corresponding author.
E-mail address: giulio.costantini@unimib.it (G. Costantini).
[1] The first two authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jrp.2014.07.003
0092-6566/© 2014 Elsevier Inc. All rights reserved.

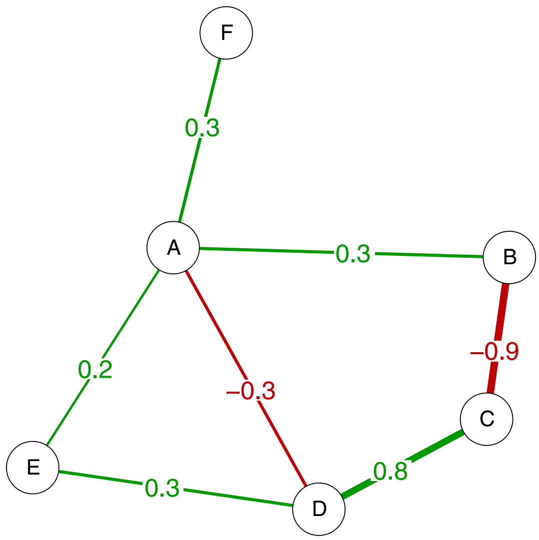

Fig. 1. A network with six nodes and seven edges. Positive edges are green and negative edges are red. The letters identify the nodes, the numbers represent weights associated to the edges. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

individuals'' (McCrae & Costa, 2008, p. 288). This approach has culminated in currently influential models such as the Five Factor Model of personality (McCrae & Costa, 2008), in which five dominant latent variables are ultimately held responsible for most of the structural covariation between responses to personality items (additional latent factors such as facets may cause some of the covariation).

Recently, however, this perspective has been challenged in the literature (Cramer et al., 2012a). In particular, it has been put forward that the default reliance on latent variable models in personality may be inappropriate, because it may well be that the bulk of the structural covariation in personality scales results from direct interactions between the variables measured through personality items. For instance, one may suppose that people who like to go to parties gain more friends because they meet more people, and people who have more friends get invited to good parties more often. In this way, one can achieve an explanation of the relevant pattern of covariation without having to posit latent variables.

Thus, in this scheme of thinking, one may suppose that, instead of reflecting the pervasive influence of personality factors, the structural covariance in personality is actually due to local interactions between the variables measured. In this way of thinking, personality resembles an ecosystem in which some characteristics and behaviors stimulate each other, while others have inhibitory relations. Under this assumption, the proper way to analyze personality data is not through the a priori imposition of a latent variable structure, but through the construction of a network that represents the most important relations between variables; this way, one may get a hold of the structure of the ecosystem of personality.

It is important to stress that not all personality scholars have embraced a causal view of latent factors. Some researchers for instance consider factors as the common elements shared by many observable variables and not as their causes (e.g., Ashton & Lee, 2005; Funder, 1991; Lee, 2012). Also from this different theoretical perspective, the heuristic value of network analysis remains important. Factor and network analysis differ, at the very least, in the fact that they direct the researcher's attention toward different aspects of personality. While factor analysis focuses almost exclusively on the elements shared among the indicators, whether or not interpreted causally, network analysis shifts the focus towards the direct relationships among the observable variables. We do not challenge the use of factor analysis as a statistical technique by itself: network analysis and factor analysis can in principle be combined (Cramer et al., 2012b; Steyer, 2012). However, a network perspective may foster important insights in the field that are unlikely to come by relying exclusively on a latent variable perspective.

The current section explains how a network structure can be estimated and visualized in R based on typical personality research data. We explain how networks are encoded in weights matrices, discuss the most important kinds of networks and show how to estimate these network.

1.1. Directed and undirected networks

There are different types of networks, which yield different kinds of information and are useful in different situations. In a *directed* network, relationships between nodes are asymmetrical. Research on directed networks has seen extensive developments in recent years since the work of Pearl (2000) and others on causal systems. Methodology based on directed networks is most useful if one is willing to accept that the network under consideration is *acyclic*, which means that there are no feedback loops in the system (if A influences B, then B cannot influence A). A directed network without feedback loops is called a Directed Acyclic Graph (DAG). In contrast, in an *undirected* network, all relationships are symmetrical. These networks are most useful in situations where (a) one cannot make the strong assumption that the data generating model is a DAG, (b) one suspects that some of the relations between elements in the network are reciprocal, and (c) one's research is of an exploratory character and is mainly oriented to visualizing the salient relations between nodes. Since the latter situation appears more realistic for personality research, the current paper focuses primarily on undirected networks.

1.2. Encoding a network in a weights matrix

The structure of a network depends on the relations between its elements. *Unweighted* networks represent only the presence or absence of the edges, while *weighted* networks encode additional information about the magnitude of the connections. When it is important to distinguish large from small connections—such as in personality—weighted networks are preferred. A weighted network can be encoded in a *weights matrix*, which is a square matrix in which each row and column indicate a node in the network. The elements of the matrix indicate the strength of connection between two nodes; a zero in row i and column j indicates that there is no edge between node i and node j. For example, the network of Fig. 1 can be represented with the following weights matrix:

	A	B	C	D	E	F
A	0	0.3	0	−0.3	0.2	0.3
B	0.3	0	−0.9	0	0	0
C	0	−0.9	0	0.8	0	0
D	−0.3	0	0.8	0	0.3	0
E	0.2	0	0	0.3	0	0
F	0.3	0	0	0	0	0

In this network there are positive connections, for instance between nodes A and B, and negative connections, for instance between nodes A and D. The zeroes in the matrix indicate that there are absent connections in the network, such as between nodes A and C. Furthermore, we may note that the matrix is symmetric and that the diagonal values are not used in the network.

The *qgraph* package (Epskamp et al., 2012) can be used to visualize such a weights matrix as a network:

```
mat <- matrix(c(
  0, 0.3, 0, −0.3, 0.2, 0.3,
  0.3, 0, −0.9, 0, 0, 0,
  0, −0.9, 0, 0.8, 0, 0,
  −0.3, 0, 0.8, 0, 0.3, 0,
  0.2, 0, 0, 0.3, 0, 0,
  0.3, 0, 0, 0, 0, 0), ncol = 6, nrow = 6,
  byrow = TRUE)
library(''qgraph'')
qgraph(mat, layout = ''spring'', edge.labels = TRUE,
  labels = LETTERS[1:6], fade = FALSE)
```

Here, the first argument in the `qgraph` function—the (`mat`) argument—calls the weights matrix to plot. The other arguments specify graphical layout.

1.3. Correlation networks, partial correlation networks, and LASSO networks

To illustrate network analysis on personality data we made public a dataset in which nine-hundred-sixty-four participants (704 female and 256 male, *M* age = 21.1, *SD* = 4.9, plus four participants who did not indicate gender and age) were administered the HEXACO-60 (Ashton & Lee, 2009). The HEXACO-60 is a short 60-items inventory that assesses six major dimensions of personality: honesty–humility, emotionality, extraversion, agreeableness vs. anger, conscientiousness and openness to experience (Ashton & Lee, 2007). Each of the major dimensions subsumes four facets, which can be computed as the average of two or three items. Participants indicated their agreement with each statement on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). An example of an item (of trait emotionality) is "When I suffer from a painful experience, I need someone to make me feel comfortable".

We can load the HEXACO dataset into R as follows:

```
Data <- read.csv(''HEXACOfacet.csv'')
```

The reader may use `str(Data)` to get an overview of the variables in the dataset. Exploratory factor analysis can be performed to inspect the structure of the dataset, using package *psych* (Revelle, 2013). The command `fa.parallel(Data)` executes parallel analysis, which suggests six factors.[2] The command `fa(r=Data, nfactors=6, rotate=''Varimax'')` can be used to extract six orthogonal factors. Factor loadings are reported in Table B.1 and reproduce the expected structure (Ashton & Lee, 2009). For each facet Table B.1 reports also the squared multiple correlation with all the other facets and the Hofmann's row-complexity index, which represents the number of latent variables needed to account for each manifest variable (Hofmann, 1978; Pettersson & Turkheimer, 2010) and is included in the output of function `fa`.

1.3.1. Correlation networks

We will construct networks by representing measured variables as nodes, connected by an edge if two variables interact with each other. To do this we can use a simple heuristic: node A is connected to node B if node A *is associated with* node B. A correlation matrix describes pairwise associations between the facets of the HEXACO and therefore can be used for estimating such a network structure. We can compute Pearson correlations on this dataset using the `cor` function:

```
cor(Data)
```

Notice that a correlation matrix is symmetric and that a value of zero indicates no connection. Thus, a correlation matrix, by default, has properties that allow it to be used as a weights matrix to encode an undirected network. Using this connection opens up the possibility to investigate correlation matrices visually as networks. To do so, we can use the *qgraph* package and ask it to plot the correlation matrix as a network; in the remainder, we will indicate this network as a *correlation network*. To facilitate interpretation, we color nodes according to the assignment of facets to traits as specified in the HEXACO manual:

```
groups <- factor(c(
  rep(''Honesty Humility'', 4),
  rep(''Emotionality'', 4),
  rep(''Extraversion'', 4),
  rep(''Agreeableness vs. Anger'', 4),
  rep(''Conscientiousness'', 4),
  rep(''Openness to experience'', 4)))
qgraph(cor(Data), layout = ''spring'', labels =
  colnames(Data),
  groups = groups)
```

Fig. 2A represents the correlation structure of the facets of the HEXACO dataset. Green lines represent positive correlations, while red lines represent negative correlations. The wider and more saturated an edge is drawn, the stronger the correlation. As the reader may expect, the figure shows that the correlations of facets within traits are generally higher than the correlations of facets between traits, which is likely to reflect the fact that in psychometric practice items are typically grouped and selected on the basis of convergent and discriminant validity (Campbell & Fiske, 1959).

In recent literature correlation networks have been applied to grasp complex co-variation patterns in personality data that would be harder to notice otherwise in, say, factor loading matrices. Epskamp et al. (2012) showed how *qgraph* can be used to visualize the correlational structure of a 240 node dataset (Dolan, Oort, Stoel, & Wicherts, 2009) in which the NEO-PI-R (Costa & McCrae, 1992; Hoekstra, De Fruyt, & Ormel, 2003) was used to assess the five factor model for personality (McCrae & Costa, 2008). Cramer et al. (2012a) further analyzed this network and showed that it did not correspond to a correlation network that should arise had the data been generated by the five factor model for personality. Ziegler, Booth, and Bensch (2013) constructed a correlation network on 113 personality facet scale scores from the NEO-PI-R, HEXACO, 6FPQ, 16PF, MPQ, and JPI and interpreted this network as a nomological network usable in scale development. Schlegel, Grandjean, and Scherer (2013) investigated the overlap of social and emotional effectiveness constructs and found the correlation network to display four meaningful components. Finally, Franić, Borsboom, Dolan, and Boomsma (2013) used correlation networks to show the similarity between genetic and environmental covariation between items of the NEO-FFI.

1.3.2. Partial correlation networks

Correlation networks are highly useful to visualize interesting patterns in the data that might otherwise be very hard to spot. However, they are not necessarily optimal for the application of network analysis if the goal is to extract the structure of a data generating network. The reason is that correlations between nodes in the network may be spurious, rather than being due to a genuine interaction between two nodes. For instance, spurious correlations may arise as the consequence of shared connections with a third node. Often, therefore, a network is constructed using the partial correlation matrix, which gives the association that is left between any two variables after conditioning on all other variables. The partial correlation coefficients are directly related to the inverse of the correlation matrix, also called the precision matrix (Lauritzen,

[2] The first seven eigenvalues are 3.52, 2.71, 2.27, 1.92, 1.73, 1.33, 0.86; the first seven eigenvalues extracted from random data are 1.29, 1.25, 1.22, 1.19, 1.16, 1.13, 1.11. Six factors explain the 42% of the common variance.

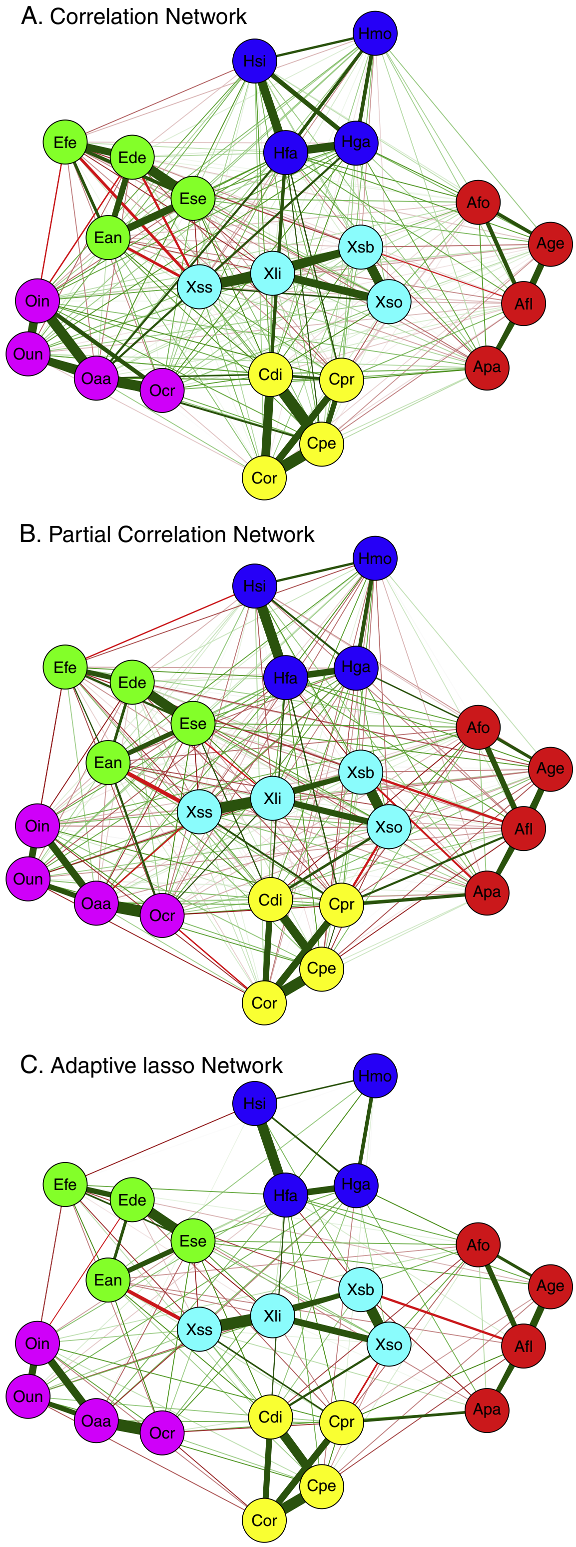

Fig. 2. Networks of the HEXACO-60. Nodes represent personality facets (a description of each facet is provided in Table A.1), green lines represent positive connections and red lines represent negative connections. Thicker lines represent stronger connections and thinner lines represent weaker connections. The node placement of all graphs is based on the adaptive LASSO network to facilitate comparison. The width and color are scaled to the strongest edge and not comparable between graphs; edge strengths in the correlation network are generally stronger than edge strengths in the partial correlation network. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1996; Pourahmadi, 2011). Networks constructed on this basis are called *partial correlation networks* or *concentration graphs* (Cox & Wermuth, 1993), and the statistical data generating structures that they encode are known as Markov random fields (Kindermann & Snell, 1980).

The partial correlation network can be obtained in *qgraph* by using the argument `graph = ‘‘concentration’’`:

```
qgraph(cor(Data), layout = ‘‘spring’’, labels =
  colnames(Data),
  groups = groups, graph = ‘‘concentration’’)
```

The partial correlation network is shown in Fig. 2B. We can see that nodes still cluster together; the partial correlations within traits are generally stronger than the partial correlations between traits. Comparing Fig. 2A and B we can see structure emerging in for example the Openess (purple) cluster: the creativity node (Ocr) is no longer directly connected by strong edges to the inquisitiveness (Oin) and unconventionality (Oun) nodes but now indirectly via the aesthetic appreciation (Oaa) node. Furthermore, we can see that the conscientiousness node prudence (Cpr) now has a more central role in the network and obtained relatively stronger connections with nodes of different traits: flexibility (Afl) and patience (Apa) of the agreeableness vs. anger trait and sociability (Xso) and Social self-esteem (Xss) of the extroversion trait.

1.3.3. Adaptive LASSO networks

In weighted networks, two nodes are connected if and only if the strength of connection between them is nonzero; a value of zero in the weights matrix encodes no connection between two nodes. Both the correlation and the partial correlation networks have been estimated based on an empirical sample and will therefore not result in exact zeroes. Thus, both networks will always be fully connected networks, possibly with arbitrarily small weights on many of the edges.

It has been argued that in social sciences everything is to some extent correlated with everything. This is akin to what Meehl and Lykken have called the *crud factor* or *ambient noise level* (Lykken, 1968, 1991; Meehl, 1990) and what may at least partly be responsible for the controversial general factor of personality (Musek, 2007). If a network model of pairwise interactions is assumed to underlie the data then all nodes that are indirectly connected will be correlated, mainly due to spurious connections. Therefore, even at the population level we can assume that most correlations in personality research will be nonzero, resulting in a fully connected correlation network.

While correlation networks of personality measures are likely to be fully connected in the population, partial correlation networks are not necessarily so. This is of specific interest since the absence of an edge in a partial correlation network entails that two nodes are conditionally independent given all other nodes in the network—they cannot directly interact. The model in which partial correlations are set to zero is called the Gaussian graphical model (GGM; Lauritzen, 1996) as it can be visualized as a network. An optimal GGM is both sparse (many absent edges) while maintaining a high likelihood. Finding such a model corresponds to checking which connections are absent in the population network. Default significance tests can be used for this purpose (Drton & Perlman, 2004). However, significance tests require an arbitrary choice of significance level; different choices yield different results, with more stringent significance levels resulting in sparser networks. If one ignores this issue, one has a multiple testing problem, whereas if one deals with it in standard ways (e.g., through a Bonferroni correction), one faces a loss of power.

A practical way to deal with the issue of arbitrary choices is to construct networks based on different choices and to see how stable the main results are; however, a more principled alternative is to use a LASSO penalty (Friedman, Hastie, & Tibshirani, 2008) in estimating the partial correlation networks. This causes small connections to automatically shrink to be exactly zero and results in a parsimonious network. If the data indeed arose from a sparse network with pairwise interactions, such a procedure will in fact converge on the generating network (Foygel & Drton, 2011).

The adaptive LASSO is a generalization of the LASSO that assigns different penalty weights for different coefficients (Zou, 2006) and outperforms the LASSO in the estimation of partial correlation networks, especially if the underlying network is sparse (Fan, Feng, & Wu, 2009; Krämer, Schäfer, & Boulesteix, 2009). The penalty weights can be chosen in a data-dependent manner, relying on the LASSO regression coefficients (Krämer et al., 2009). In simulation studies, the likelihood of false positives using this method resulted even smaller than that obtained with the LASSO penalization (Krämer et al., 2009), so if an edge is present in the adaptive LASSO network one can reasonably trust that there is a structural relation between the variables in question (of course, the network does not specify the exact nature of the relation, which may for instance be due to a direct causal effect, a logical relation pertaining to item content, a reciprocal effect, or the common effect of an unmodeled latent variable).

The adaptive LASSO is also convenient practically, as it is implemented in the R-package *parcor* (Krämer et al., 2009). Since the adaptive LASSO, as implemented in package *parcor*, relies on *k*-fold validation, `set.seed` can be used to ensure the exact replicability of the results, which might be slightly different otherwise. To estimate the network structure of the HEXACO dataset according to the adaptive LASSO, the following code can be used:

```
library(''parcor'')
library(''Matrix'')
set.seed(100)
adls <- adalasso.net(Data)
network <-
  as.matrix(forceSymmetric(adls$pcor.adalasso))
qgraph(network, layout = ''spring'', labels =
  colnames(Data), groups = groups)
```

The adaptive LASSO network is shown in Fig. 2C. One can see that, compared to the partial correlation network, the adaptive LASSO yields a more parsimonious graph (fewer connections) that encodes only the most important relations in the data; In this network 134 (48.6%) of the edges are identified as zero.

2. Analyzing the structure of personality networks

Once a network is estimated, several indices can be computed that convey information about network structure.[3] Two types of structure are important. First, one is typically interested in the *global* structure of the network: how large is it? Does it feature strong clusters? Does it reveal a specific type of structure, like a small-world (Watts & Strogatz, 1998)? Second, one may be interested in *local* patterns, i.e., one may want to know how nodes differ in various characteristics: which nodes are most central? Which nodes are specifically strongly connected? What is the shortest path from node A to node B? Here we discuss a limited selection of indices that we regard as relevant to personality research, focusing especially on centrality and clustering coefficients. More extensive reviews of network indices may be found in Boccaletti, Latora, Moreno, Chavez, and Hwang (2006), Butts (2008a), De Nooy et al. (2011), Kolaczyk (2009), and Newman (2010).

2.1. Descriptive statistics

Before the computation of centrality measures, a number of preparatory computations on the data are in order. The network is undirected, therefore the corresponding weights matrix is symmetric and each edge weight is represented twice, above and below the main diagonal. The function `upper.tri` can be used to extract the unique edge weights[4] and save them in a vector:

```
ew <- network[upper.tri(network)]
```

To compute the number of edges in the network, it is sufficient to define a logical vector that has value `TRUE` (=1) if the edge is different from zero and `FALSE` (=0) if the edge is exactly zero (i.e., absent). The sum of this vector gives the number of nonzero edges. With a similar procedure, it is possible to count the positive and the negative edges: it is sufficient to replace "`!=`" with ">" or "<".

```
sum(ew != 0) # the number of edges
sum(ew > 0) # the number of positive edges
sum(ew < 0) # the number of negative edges
```

The network has 142 edges, of which 100 are positive and 42 are negative. The function `t.test` can be used to compare the absolute weights of the positive vs. the negative edges:

```
t.test(abs (ew [ew > 0]), abs(ew [ew < 0]), var.equal
  = TRUE)
```

In our network, positive edges are generally associated to larger weights ($M = .11$, $SD = .09$) than the negative edges ($M = .06$, $SD = .04$), and the t-test indicates that this difference is significant, $t(140) = 3.13$, $p = .0022$.

2.2. Centrality measures

Not all nodes in a network are equally important in determining the network's structure and, if processes run on the network, in determining its dynamic characteristics (Kolaczyk, 2009). Centrality indices can be conceived of as operationalizations of a node's importance, which are based on the pattern of the connections in which the node of interest plays a role. In network analysis, centrality indices are used to model or predict several network processes, such as the amount of flow that traverses a node or the tolerance of the network to the removal of selected nodes (Borgatti, 2005; Crucitti, Latora, Marchiori, & Rapisarda, 2004; Jeong, Mason, Barabási, & Oltvai, 2001) and can constitute a guide for network interventions (Valente, 2012). Several indices of centrality have been proposed, based on different models of the processes that characterize the network and on a different conception of what makes a node important (Borgatti, 2005;

[3] The adaptive LASSO networks, the correlation and the partial correlation networks are characterized by the presence of both positive and negative edges. The importance of signed networks is apparent not only in the study of social phenomena, in which it is important to make a distinction between liking and disliking relationships (e.g., Leskovec, Huttenlocher, & Kleinberg, 2010), but also in the study of personality psychology (e.g., Costantini & Perugini, 2014). Some network indices have been generalized to the signed case (e.g., Costantini & Perugini, 2014; Kunegis, Lommatzsch, & Bauckhage, 2009), however most indices are designed to unsigned networks. For the computation of the latter kind of indices, we will consider the edge weights in absolute value.

[4] The function `upper.tri` extracts the elements above the main diagonal. One could equally consider those below the diagonal using the function `lower.tri`.

Borgatti & Everett, 2006). The following gives a succinct overview of the most often used centrality measures.[5]

2.2.1. Degree and strength

First, *degree centrality* is arguably the most common centrality index and it is defined as the number of connections incident to the node of interest (Freeman, 1978). The degree centrality of node C in Fig. 1 is 2 because it has two connections, with nodes B and D. Degree can be straightforwardly generalized to weighted networks by considering the sum of the weights of the connections (in absolute value), instead of their number. This generalization is called *strength* (Barrat, Barthelémy, Pastor-Satorras, & Vespignani, 2004; Newman, 2004). For instance, strength of node C in Fig. 1 is 1.7, which is the highest in the network. Degree and strength focus only on the paths of unitary length (Borgatti, 2005). A strength-central personality characteristic (e.g., an item, a facet or a trait) is one that can influence many other personality characteristics (or be influenced by them) directly, without considering the mediating role of other nodes.

2.2.2. Closeness and betweenness

Several other measures exist that, differently from degree centrality and the related indices, consider edges beyond those incident to the focal node. An important class of these indices rely on the concepts of *distance* and of *geodesics* (Brandes, 2001; Dijkstra, 1959). The distance between two nodes is defined as the length of the shortest path between them. Since, in typical applications in personality psychology, weights represent the importance of an edge, weights are first converted to lengths, usually by taking the inverse of the absolute weight (Brandes, 2008; Opsahl, Agneessens, & Skvoretz, 2010). The *geodesics* between two nodes are the paths that connect them that have the shortest distance. *Closeness centrality* (Freeman, 1978; Sabidussi, 1966) is defined as the inverse of the sum of the distances of the focal node from all the other nodes in the network.[6] In terms of network flow, closeness can be interpreted as the expected speed of arrival of something flowing through the network (Borgatti, 2005). A closeness-central personality characteristic is one that is likely to be quickly affected by changes in another personality characteristic, directly or through the changes in other personality features. Its influence can reach other personality features more quickly than the influence of those that are peripheral according to closeness, because of the short paths that connect itself and the other traits. In the network in Fig. 1, node D has the highest closeness. To compute the exact value of closeness, one should first compute the distances between D and all the other nodes: A (1/0.3), B (1/0.8 + 1/0.9), C (1/0.8), E (1/.3) and F (1/.3 + 1/.3). The sum of all the distances is 16.94 and the inverse, 0.059, is the closeness centrality of D.

Betweenness centrality is defined as the number of the geodesics between any two nodes that pass through the focal one. To account for the possibility of several geodesics between two nodes, if two geodesics exist, each one is counted as a half path and similarly for three or more (Brandes, 2001; Freeman, 1978). Betweenness centrality assumes that shortest paths are particularly important (Borgatti, 2005): if a node high in betweenness centrality is removed, the distances among other nodes will generally increase. Both closeness and betweenness centrality can be applied to weighted and directed networks, as long as the weights and/or the directions of the edges are taken into account when computing the shortest paths (e.g., Opsahl et al., 2010).

The betweenness centrality of node A in Fig. 1 is 4 and is the highest in the network. The four shortest paths that pass through A are those between F and the nodes B, C, D, and E. Betweenness centrality can also be extended to evaluate the centrality of edges instead of nodes, by considering the geodesics that pass through an edge: this generalization is called *edge betweenness centrality* (Brandes, 2008; Newman, 2004; Newman & Girvan, 2004). For instance, the edge-betweenness centrality of the edge (D, E) is 3 and the three shortest paths that pass through (D, E) are the one between D and E, the one between C and E (through D), and the between B and E (through C and D).

Betweenness-central personality characteristics and betweenness-central edges are particularly important for other personality characteristics to quickly influence each other. It is interesting to investigate the conditions in which some nodes become more or less central. For instance, a study that analyzed a network of moods showed that the mood "worrying" played a more central role for individuals high in neuroticism than for those with low neuroticism (Bringmann et al., 2013): the prominent role of worrying for neuroticism was recently confirmed by an experimental fMRI study (Servaas, Riese, Ormel, & Aleman, 2014).

Several other variants of the shortest-paths betweenness are discussed in Brandes (2008), some of which are implemented in package *sna* (Butts, 2008b). Generalizations of betweenness centrality that account for paths other than the shortest ones have been also proposed (Brandes & Fleischer, 2005; Freeman, Borgatti, & White, 1991; Newman, 2005). In addition, Opsahl et al. (2010) proposed generalizations of degree, closeness, and betweenness centralities by combining in the formula both the number and the weights of the edges. They introduced a tuning parameter that allows setting their relative importance: a higher value of the tuning parameter emphasizes the importance of the weights over the mere presence of the ties and vice versa. Another important family of centrality indices defines the centrality of a node as recursively dependent on the centralities of their neighbors. Among the most prominent of those indices are *eigenvector centrality* (Bonacich, 1972, 2007), *Bonacich power* (Bonacich, 1987) and *alpha centrality* (Bonacich & Lloyd, 2001).

2.3. Clustering coefficients

Besides centrality, other network properties have been investigated that are relevant also for personality networks. The local *clustering coefficient* is a node property defined as the number of connections among the neighbors of a focal node over the maximum possible number of such connections (Watts & Strogatz, 1998). If we define a triangle as a triple of nodes all connected to each other, the clustering coefficient can be equally defined as the number of triangles to which the focal node belongs, normalized by the maximum possible number of such triangles. The clustering coefficient is high for a node *i* if most of *i*'s neighbors are also connected to each other and it is important to assess the small-world property (Humphries & Gurney, 2008; Watts & Strogatz, 1998), as we detail below. Consider for instance the node D in Fig. 1, which has three neighbors, A C, and E. Of the three possible connections among its neighbors, only one is present (the one between A and E), therefore its clustering coefficient is 1/3.

The clustering coefficient can be also interpreted as a measure of how much a node is redundant (Latora, Nicosia, & Panzarasa, 2013; Newman, 2010): if most of a node's neighbors are also connected with each other, removing that node will not make it harder

[5] The functions to implement centrality indices, clustering coefficients and small-worldness are implemented in the R package *qgraph* (Epskamp et al., 2012). Some of the functions rely on procedures originally implemented in packages *igraph* (Csárdi & Nepusz, 2006), *sna* (Butts, 2008b), and *WGCNA* (Langfelder & Horvath, 2008, 2012). These packages are in our experience among the most useful for network analysis.

[6] The computation of closeness assumes that the network is connected (i.e., a path exists between any two nodes), otherwise, being the distance of disconnected nodes infinite, the index will result to zero for all the nodes. Variations of closeness centrality that address this issue have been proposed (e.g., Kolaczyk, 2009, p. 89; Opsahl et al., 2010, n. 1). Alternatively it can be computed only for the largest component of the network (Opsahl et al., 2010).

for its neighbors to reach or influence each other. A personality characteristic that has a high clustering coefficient is mainly connected to other personality features which are directly related to each other. In personality questionnaires the strongest connections are usually among nodes of the same subscale: in these cases, having a high clustering coefficient may coincide with having most connections with other nodes belonging to the same subscale, while having no large connection with nodes of other scales.

While in its original formulation the clustering coefficient can be applied only to unweighted networks (or to weighted networks, disregarding the information about weights), it has been recently generalized to consider positive edge weights (Saramäki, Kivelä, Onnela, Kaski, & Kertész, 2007). The first of such generalizations was proposed by Barrat et al. (2004) and has been already discussed in the context of personality psychology and psychopathology (Borsboom & Cramer, 2013). Onnela, Saramäki, Kertész, and Kaski (2005) proposed a generalization that is based on the geometric averages of edge weights of each triangle centered on the focal node. A different generalization has been proposed in the context of gene co-expression network analysis by Zhang and Horvath, which is particularly suited for networks based on correlations (Kalna & Higham, 2007; Zhang & Horvath, 2005). All of these generalizations coincide with the unweighted clustering coefficient when edge weights become binary (Saramäki et al., 2007). Recently three formulations of clustering, the unweighted clustering coefficient (Watts & Strogatz, 1998), the index proposed by Onnela and colleagues (2005) and the one proposed by Zhang and Horvath (2005) have been generalized to signed networks and the properties of such indices have been discussed in the context of personality networks (Costantini & Perugini, 2014).

Transitivity (or global clustering coefficient) is a concept closely connected to clustering coefficient that considers the tendency for two nodes that share a neighbor to be connected themselves for the entire network, instead than for the neighborhood of each node separately. It is defined as three times the number of triangles, over the number of connected triples in the network, where a connected triple is a node with two edges that connect it to an unordered pair of other nodes (Newman, 2003). Differently from the local clustering coefficient, transitivity is a property of the network and not of the single nodes. For instance, the network in Fig. 1 has one triangle (A, D, E) and 12 connected triples, therefore its transitivity is $(3^{*}1)/12 = 1/4$. Transitivity has been extended by Opsahl and Panzarasa (2009) to take into account edge weights and directions, and by Kunegis and collaborators to signed networks (Kunegis et al., 2009).

2.4. Small worlds

The transitivity and clustering coefficient can be used to assess the network *small-world property*. The small-world property was initially observed in social networks as the tendency for any two people to be connected by a very short chain of acquaintances (Milgram, 1967). The small-world property is formally defined as the tendency of a network to have both a high clustering coefficient and a short average path length (Watts & Strogatz, 1998). Small-world networks are therefore characterized by both the presence of dense local connections among the nodes and of links that connect portions of the network otherwise far away from each other. An index of *small-worldness* for unweighted and undirected networks has been proposed as the ratio of transitivity to the average distance between two nodes. Both transitivity and path length are standardized before the computation of small-worldness, by comparing them to the corresponding values obtained in equivalent random networks (with the same N and the same degree distribution). Alternatively, the index can be computed using the average of local clustering coefficients instead of transitivity. A network with a small-worldness value higher than three can be considered as having the small-world property, while a small-worldness between one and three is considered a borderline value (Humphries & Gurney, 2008). Because the assessment of small-worldness relies on shortest paths between all the pairs of nodes, it can be computed only for a connected network or the giant component of a disconnected network.

Table 1
Correlation of node centralities, row-complexity and squared multiple correlation (SMC).

	1	2	3	4	5
1. Betweenness	1	.61**	.72***	.32	.54**
2. Closeness	.61**	1	.75***	.15	.69***
3. Strength	.70***	.82***	1	.47*	.75***
4. Complexity	.41*	.28	.43*	1	.11
5. SMC	.56**	.73***	.79***	.12	1

Note. Pearson correlations are reported below the diagonal, Spearman correlations are reported above the diagonal. Complexity = Hofmann's row-complexity index. SMC = squared multiple correlation.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

2.5. Application to the HEXACO data

2.5.1. Centrality analyses

The function `centrality_auto` allows to quickly compute several centrality indices. It requires the weights matrix as input. The function automatically detects the type of network and can handle both unweighted and weighted networks, and both directed and undirected networks. For a weighted and undirected network, the function gives as output the node strength, the weighted betweenness and the weighted closeness centralities. The edge betweenness centrality is also computed.

```
centrality <- centrality_auto(network)
nc <- centrality$node.centrality
ebc <- centrality$edge.betweenness.centrality
```

The centrality values are computed and stored in variable `centrality`. Node centralities are then saved in the variable `nc` while edge betweenness centralities are saved in the variable `ebc`. The values of centrality for each node are reported in Table A.1. The command `centralityPlot(network)` can be used to plot the centrality indices in a convenient way, that allows to quickly compare them. Table 1 reports the correlations among the three indices of node centrality together with Hofmann's (1978) row-complexity and the squared multiple correlation of each facet with all the others. All the indices of centrality have positive significant correlations with each other. Strength centrality and, to a lower extent, betweenness centrality, seem to be favored by row-complexity: sharing variance with more than one factor allows a facet to play a more central role. This results suggest that, in this network, facets tend to be central to the whole network and not only to their purported parent traits. All centrality indices, especially strength and closeness, correlate with the squared multiple correlations: The more variance a facet shares with other facets, the stronger are its connections and the more central results the corresponding node.[7]

The three indices of centrality converge in indicating that node Cpr (prudence) is among the four most central nodes in this network. Cpr is also the more closeness central node and owes its high

[7] Despite being substantial, the correlations of centrality indices with row-complexity and squared multiple correlations do not suggest that the indices fully overlap. Moreover, the relations can vary substantially and it is possible to imagine situations in which the relations are absent or even in the opposite direction.

centrality to the very short paths that connect it to other traits. For instance, facets Apa (patience), Xso (sociability), and Xss (social self-esteem) are even closer to Cpr than other conscientiousness facets are.[8] This suggests that in the personality network it is very easy that a change in some portion of the network will eventually make a person either more reckless or more prudent. On the other hand, if a person becomes more reckless or more prudent, we can expect important changes in the overall network. This result, although it should be considered as preliminary, is in line with studies that investigated the evolution of conscientiousness. Impulse-control, a facet of conscientiousness that is very similar to prudence (Cpr), shows the most marked variation through the individual development compared to other conscientiousness facets (Jackson et al., 2009). It is possible that this is the case also because changes in other personality traits are expected to affect prudence more quickly than other facets, as revealed by its high closeness.

Hfa (fairness) is the most betweenness-central and strength-central node, but it is not particularly closeness-central (it is ranked 10th in closeness centrality). Fig. 3 highlights the edges lying on the shortest paths that travel through node Hfa, in a convenient layout (the code for producing this figure is in the Supplemental materials). The high betweenness centrality of Hfa is due the role that Hfa plays in transmitting the influence of other honesty–humility facets to different traits, and vice versa. The edge between nodes Hsi (sincerity) and Hfa is also the most betweenness-central in the whole network: most of the shortest paths between Hsi and other personality traits travel through this edge and therefore through Hfa. These results suggest that, if it was possible to reduce the possibility for fairness (Hfa) to vary, the influence of the other honesty–humility facets would propagate less easily to the rest of personality facets and vice versa. Such hypotheses could be tested for instance by comparing the personality networks of individuals that typically face situations in which their fairness is allowed to become active to the networks of individuals that usually face situations in which their fairness cannot be activated (Tett & Guterman, 2000). The characteristics of situations for instance could be assessed by using valid instruments such as the Riverside Situational Q-sort (Sherman, Nave, & Funder, 2010), which includes items such as "It is possible for P to deceive someone", or "Situation raises moral or ethical issues" that would be relevant for this case.

2.5.2. Clustering coefficients

Many indices of clustering coefficient can be easily computed using function `clustcoef_auto`. The function requires the same input as `centrality_auto` and is similarly programmed to recognize the kind of data given as input and to choose an appropriate network representation for the data. By applying the function, we can immediately collect the results:

```
clustcoef <- clustcoef_auto(network)
```

The command `clusteringPlot(network, signed = TRUE)` can be used to plot the clustering coefficients in a convenient layout. Table 2 reports the correlation among several clustering coefficients. The unsigned indices are computed using the absolute values of the weights. In the following analyses we will use the signed version of the Zhang's clustering coefficient (Costantini & Perugini, 2014; Zhang & Horvath, 2005), which resulted more resistant to random variations in the network (see Section 2.5.6).

2.5.3. Combining clustering coefficients and centrality

The signed clustering coefficient can be interpreted as an index of a node's redundancy in a node's neighborhood (Costantini & Perugini, 2014): the importance of the unique causal role of highly clustered nodes is strongly reduced by the presence of strong connections among their neighbors. In general, it is interesting to inspect whether there is a relation between centrality indices and clustering coefficients: in our experience, we found that the centrality indices were often inflated by the high clustering in correlation networks. However this might be not true for networks defined with adaptive LASSO, which promotes sparsity (Krämer et al., 2009).

The following plots can be used to visualize both the centrality and the clustering coefficient of each node. The code reported here is for betweenness centrality, but it is easy to extend it to other indices by just replacing "`Betweenness`" with the index of interest. First the plot is created and then the node labels are added in the right positions, using the command text. Command `abline` can be used to trace lines in the plot. A horizontal line is created to visually identify the median value of betweenness and a vertical line to identify the median value of the clustering coefficient.

```
plot(clustcoef$signed_clustZhang,
  nc$Betweenness, col = ''white'')
text(clustcoef$signed_clustZhang,
  nc$Betweenness, rownames(nc))
abline(h = median(nc$Betweenness), col = ''grey'')
abline(v = median(clustcoef$signed_clustZhang),
  col = ''grey'')
```

The resulting plots are shown in Fig. 4. It is apparent that the most central nodes do not have a particularly high clustering coefficient in this case and this is especially true for nodes Hfa and Cpr, which are among the most central in this network. The clustering coefficient correlates negatively with closeness centrality ($r = -.67$, $p < .001$), with strength ($r = -.82$, $p < .001$), and with betweenness centrality ($r = -.50$, $p = .013$).

One node, Hmo (modesty), emerges as both particularly high in clustering coefficient and low in all the centrality measures. Modesty correlates almost exclusively with other honesty–humility facets and has the lowest multiple correlation with all the other variables in our dataset and this is likely to have determined its peripherality. A closer exam of its connections reveals that Hmo has seven neighbors, the three other facets of honesty–humility (His, Hfa, and Hga), facets anxiety and fearfulness of emotionality (Ean), facet social boldness of extraversion (Xsb) and facet prudence of conscientiousness (Cpr), the connections with fearfulness, social boldness and prudence having very small weights. Moreover many of its neighbors are connected with each other. Even if the edges incident in node Hmo were blocked, its neighbors would be nonetheless connected to each other directly or by a short path. Modesty therefore does not seem to play a very important unique role in the overall personality network.

2.5.4. Transitivity and small-world-ness

The function `smallworldness` computes the small-worldness index (Humphries & Gurney, 2008). First the function converts

[8] As an anonymous reviewer pointed out, one could wonder how can the length of the path between Cpr and other conscientiousness facets be longer than the path between Cpr and other nodes, given that Cpr's strongest correlations are those with the other conscientiousness facets. This happens because we did not consider the network defined by the zero-order correlations, but the adaptive LASSO penalized network of partial correlations (Krämer et al., 2009). As an example, consider the shortest path between Cpr and Cdi (diligence), which is slightly longer (8.80) than the shortest path between Cpr and Apa (patience; 6.82). Although the correlation among Cpr and Cdi is stronger ($r = .26$) than the correlation between Cpr and Apa ($r = .22$), in the adaptive LASSO network, the direct connection between Cpr and Cdi is smaller ($pr = .04$) than the one with Apa ($pr = .15$). While the shortest path between Cpr and Apa travels through their direct connection, the shortest path between Cpr and Cdi travels through node Cor (organization): prudence seems to influence (or to be influenced by) diligence especially through changes in orderliness, but this path of influence is longer than the direct path between Cpr and Apa.

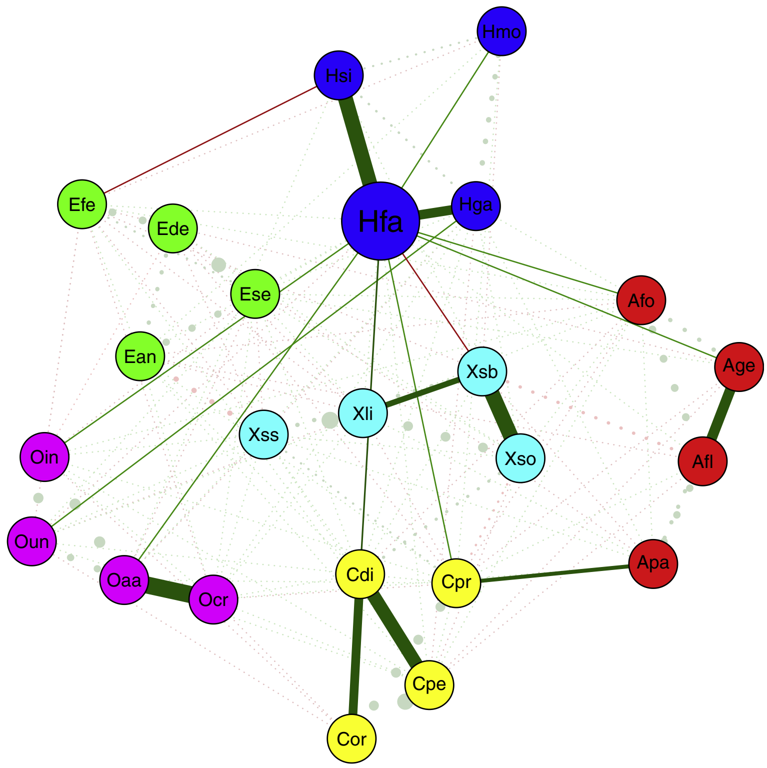

Fig. 3. Shortest paths that pass through node Hfa (fairness). The edges belonging to the shortest-paths are full, while the other edges are dashed.

the network to an unweighted one, which considers only the presence or the absence of an edge. Then the average path length and the global transitivity of the network are computed and the same indices are calculated on B = 1000 random networks, with the same degree distribution of the focal network. The resulting values are entered in the computation of the small-worldness index. The output includes the small-worldness index, the transitivity of the network, and its average path length. It also returns summaries of the same indices computed on the random networks: the mean value and the .005 and .995 quantiles of the distribution. Function `set.seed` can be used to ensure the exact replicability of the results. The function requires the network as input and it is optionally possible to set the values of three parameters, `B`, `up` and `lo`, which are respectively the number of random networks and the upper and lower probabilities for the computation of the quantiles.

```
set.seed(100)
smallworldness(network)
```

The small-worldness value for our network is 1.01. An inspection of the values of transitivity and of average path length shows that they are not significantly different from those emerged from similar random networks. Therefore we may conclude that this personality network does not show a clear small-world topology.

2.5.5. Emerging insights

In this section, we showed how it is possible to perform a network analysis on a real personality dataset. We identified the most central nodes and edges, discussed centrality in the light of clustering coefficient and investigated some basic topological properties of the network, such as the small-world property. Two nodes resulted particularly central in the network and were the facet prudence of conscientiousness (Cpr) and the facet fairness of honesty–humility (Hfa).

Our network did not show the small-world property. The absence of a strong transitivity means that the connection of two nodes with a common neighbor does not increase the probability of a connection between themselves. The absence of a particularly short path length implies that it is not generally possible for any node to influence any other node using a short path. This result is not in line with the small-worldness property that emerged in the DSM-IV network reported by Borsboom, Cramer, Schmittmann, Epskamp, and Waldorp (2011). It has been hypothesized that the small-world property might be at the basis of

Table 2
Correlation among indices of local clustering coefficient.

	1	2	3	4	5	6	7
1. Watts and Strogatz (1998)	1	.25	.65***	.51*	.90***	.57**	.94***
2. Watts and Strogatz, signed (Costantini & Perugini, 2014)	.26	1	.28	.45*	.29	.76***	.25
3. Zhang and Horvath (2005)	.49*	.30	1	.89***	.50*	.59**	.71***
4. Zhang and Horvath, signed (Costantini & Perugini, 2014)	.34	.33	.94***	1	.37	.79***	.53**
5. Onnela et al. (2005)	.89***	.25	.37	.24	1	.55**	.84***
6. Onnela et al., signed (Costantini & Perugini, 2014)	.61**	.76**	.59**	.64**	.66***	1	.53**
7. Barrat et al. (2004)	.94***	.30	.57**	.37	.87***	.60**	1

Note. Pearson correlations are reported below the diagonal, Spearman correlations are reported above the diagonal.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

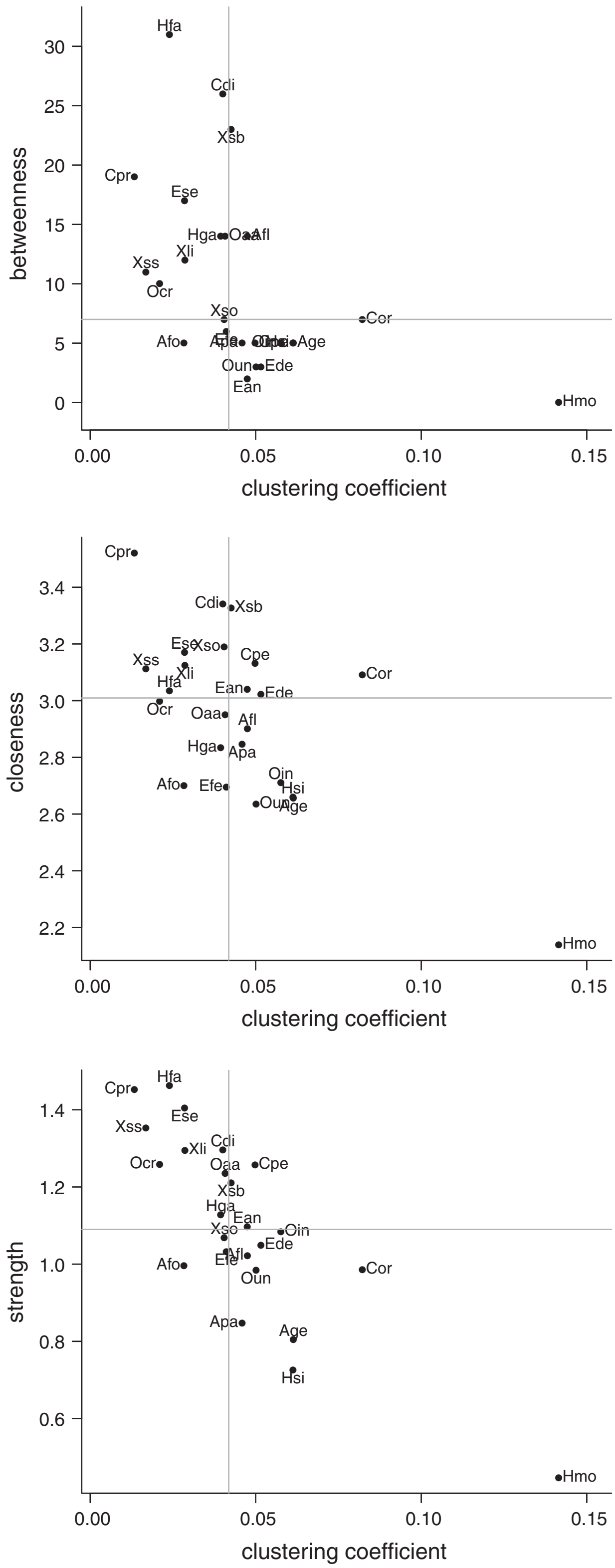

Fig. 4. Centrality and clustering coefficient. The horizontal and the vertical lines represent the median values of centrality and clustering coefficient respectively. The closeness values are multiplied by 1000.

phenomena connected to the comorbidity that arise in psychopathology (Cramer, Waldorp, van der Maas, & Borsboom, 2010); this also may simply not be a property of normal personality. This difference could reflect the fact that different personality characteristics represent distinct systems, while psychopathology systems seem to be more integrated. This result may be also attributable to the strategies that were used for defining this network and the DSM-IV network and may have been influenced by the particular personality scales under study. Future research may be directed towards the question of what network structure characterizes normal vs. abnormal personality.

2.5.6. Stability of results

The adaptive LASSO chooses the LASSO penalty parameter based on *k*-fold crossvalidation, subdividing the dataset in *k* (10 by default) random samples. Because of this, under different random seeds slightly different network structures will be obtained. To investigate the stability of the results discussed in this section, we repeated the network estimation procedure 900 times under different random seeds and recomputed the strength, closeness and betweenness centrality measures and the signed versions of the clustering coefficients proposed by Zhang and by Onnela. The codes to replicate these findings can be found in the Supplementary materials.

Visually the resulting graphs looked remarkably similar and only differed in the weakest edges in the graph. Fig. 5 shows a histogram of the amount of nonzero connections present in each of the replications; the median amount of estimated edges was 138. Fig. 6 shows the estimated centrality and clustering coefficients for both the graph used in the analyses (colored line) and the 900 replications (vague gray lines). It can be seen that overall the measures are stable across different replications. Among the three centrality measures, more stable results were obtained for closeness and strength than for betweenness. Between the clustering coefficients we can see that Zhang's clustering coefficient is much more stable than Onnela's; in Onnela's clustering coefficient especially the Hmo node shows divergent behavior. This behavior is due to the number small of connections of Hmo obtained in each replication, ranging from 3 to 11 ($M = 3.96$, $SD = 0.64$). Onnela's clustering coefficient is scaled to the number of connections

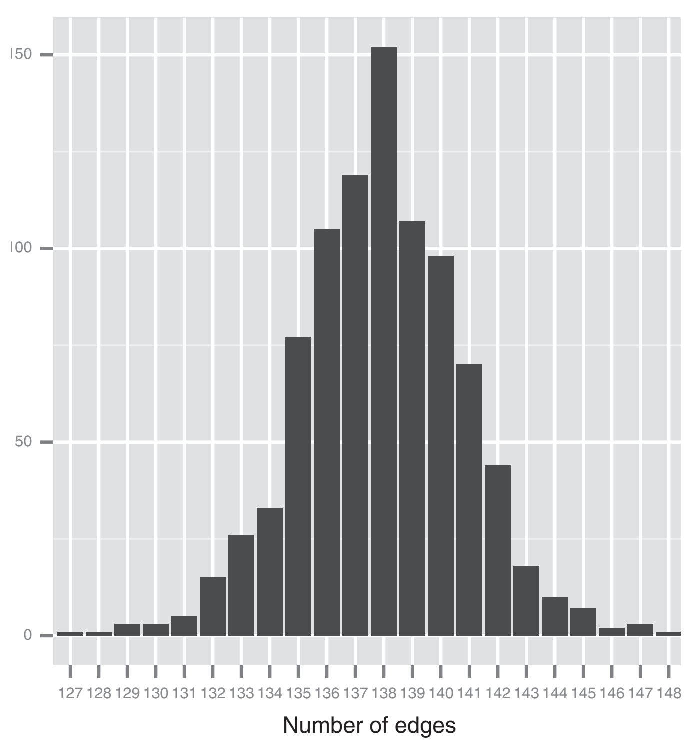

Fig. 5. Histogram of the number of edges estimated in 900 replications of the adaptive LASSO.

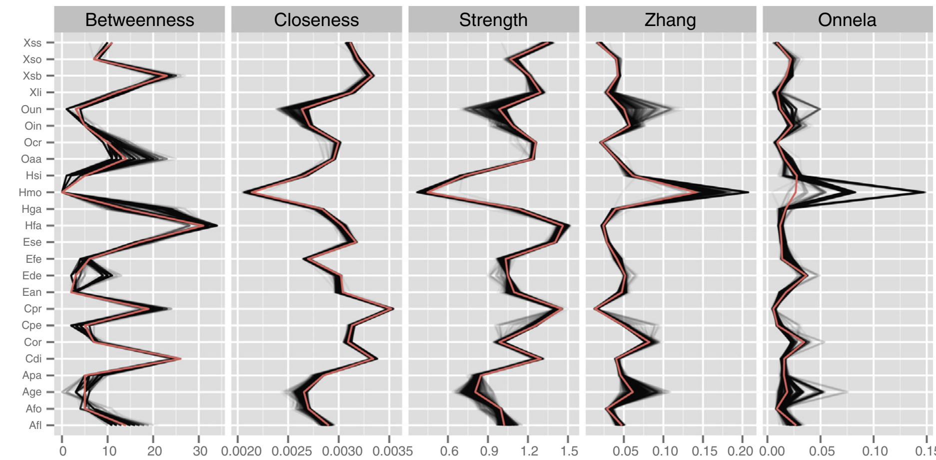

Fig. 6. Estimated centrality and clustering coefficients under 900 replications of the adaptive LASSO. The colored line represents the results discussed in the paper. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

regardless of weight. Therefore the relatively small difference in connections can have a large impact on this clustering coefficient.

From these results, we advise that Zhang's clustering coefficient should be preferred over Onnela's clustering coefficient in adaptive LASSO networks. Furthermore, we advise the reader to replicate these measures under different random seeds and to check for the stability of the results before substantively interpreting them.

3. Simulating personality networks

In addition to the analysis of empirical data, network modeling offers extensive possibilities in the area of theory development. This is because, in contrast to purely data analytic models like factor analysis, networks are naturally coupled to dynamics (e.g., see Kolaczyk, 2009): they can evolve, grow, and change over time, with direct consequences for their dynamic behavior. This makes it possible to start thinking about questions like: How do personality networks form in development? Do they grow and, if so, how, do they change in structure over time? Do different people have different network structures, and how would such differences relate to growth and dynamics?

Because networks have been so extensively studied in other fields, one can use existing analytical insights on the relevant processes (e.g., Grimmett, 2010; Kolaczyk, 2009; Newman, 2008). When applicable, existing analytical approaches can be very powerful. However, in order to use such analytical approaches, one often has to consider assumptions that are unlikely to be met in personality (e.g., many theorems require one to assume that nodes are exchangeable save for their position in the network, or work only for unweighted networks). In such cases, specifically tailored simulation methodology can be an extremely versatile tool to study the behavior of networks. This can both enlighten one's data analytic results (e.g., by checking how a given dynamical process would pan on a network extracted from data; e.g., see Borsboom et al., 2011) and help in theory development (e.g., by working out what a hypothesized network would imply theoretically).

In particular, simulation work can be used to design some hypothetical data and see how these data "behave" in appropriate analyses. Here, designing data refers to simulating data according to some pre-specified rules. The obvious strength of testing analytical procedures or concepts with simulated data is that the mechanisms by which the data arose are known—a luxury researchers almost never have when working with real data. Therefore, it is possible to see if the focal theoretical concept can, in principle, result in the expected kind of observed data or co-exist with other concepts, or whether the analytical procedure of interest can yield accurate conclusions. Obviously, designed data can provide no empirical proof for a theoretical concept—but they can guide thinking and this is almost as good.

One can attempt to simulate personality network data to exactly the same two ends. For example, some relevant questions can be the following. Is it, in principle, possible to take network principles and generate data that look similar to what personality psychologists commonly work with? And if so, how do available network analyses tools behave in these data? This section describes only a possible way of simulating personality data from the network perspective. In particular, we demonstrate how the coalescence of observable variables into traits can be simulated. This simulation is very simplistic and just serves to produce data: it does not attempt to provide a dynamic model of real-world processes.

3.1. One possible way to start

We can start off with creating k nodes (vector y). For the purpose at hand, we assume that the nodes are unrelated at the outset and that their clustering results from direct causal connections among them. Therefore, the initial value of each node is drawn separately from standard normal distribution with a mean of 0 and standard deviation of 1. Let k be 30:

```
y <- rnorm(30)
```

Before we let the nodes connect to each other, we need to specify the weights (matrix w), which reflect the amount of influence from one node to another that is entailed by each connection. We can also draw the weights randomly from a standard normal distribution, with a mean of m and standard deviation of s:

```
w <- matrix(ncol=k, rnorm(k*k, m, s))
```

However, if both the node scores and their interconnections are completely random, then this will most likely result in a chaotic network structure. This is not characteristic of real personality data, as was shown above. Instead, we have to assume that some nodes have more influence on each other than others, which makes the scores of these nodes more similar and leads to the structuredness of personality network. Here, we propose that when a node connects to other nodes (targets), its influence is inversely proportional to the distance between it and the target nodes. For this to work, we need a network structure, which specifies the distances between the nodes. A really simple way to obtain this is to imagine that all nodes are positioned on a line such that the distances between them and other nodes increase monotonically in both directions. Accordingly, a matrix of distances d can be created as follows:

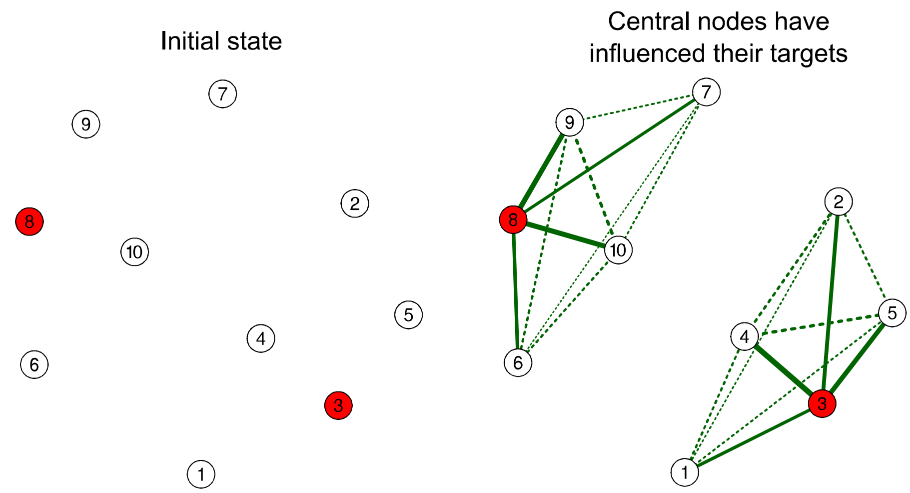

Fig. 7. A network of 10 nodes. At the initial stage (left panel), no influences have been spread around and therefore nodes are uncorrelated. At a later stage (right panel), two central nodes (red) have sent direct influences (solid lines) to nodes close to them. Dashed lines represent indirect connections. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

```
d <- as.matrix(dist(1:k))
```

Of course, we do not have to assume that the network architecture is exactly the same for each and every agent. By adding some noise to the distance matrix, we can distort it so that the distances between nodes become uneven and nodes swap their places; if this noise differs across agents, the resulting networks structures will vary as well. Let *n* be the average amount of noise added to each value in the distance matrix. We note that this may be an interesting parameter to vary as it allows us to see how much consistency in network architectures is needed for any common structure (e.g., factor-like clusters at the level of cross-agent differences) to emerge.

```
d <- d * matrix(abs(rnorm(k*k, 1, n)), k)
```

We can normalize the distance matrix and use this to inversely weigh the weight matrix:

```
d <- d / max(d)
w <- w / d
```

We can also employ the concept of centrality and assume that not all nodes are powerful enough to influence others: only central nodes may have this privilege. Obviously, the number and selection of both central and their target nodes can vary across agents. If we then think of the network dynamics as a step-by-step process (for the ease of understanding), then the connections may happen as follows: each central node goes through each of its target nodes and updates its score by adding a little bit of itself (as specified by the respective weight in the weight matrix) to the targets current score. If the central node is v_i and a target nodes is v_j, then the updating process for this target node could be written as: $v_j = v_j + w_{ij} * v_i$.

In R this can be achieved by the following code (`centrals` is the vector pointing to central nodes and `n.targets` is the number of targets each central node has):

```
for(i in centrals) {
  t <- sample(k-1, n.targets)
  y[-i][t] <- y[-i][t] + w[,-i][i,t] * y[i]
  }
```

In fact, this is all that may be necessary for creating a simple personality network of a single agent. The underlying idea of this simulation is depicted in Fig. 7. Of course, this simulation does not create data reflecting anything close to a self-organizing system that human personality most likely is. However, it may be helpful for thinking of how the system may be interconnected. The commands are above are wrapped into the function `simulator`. Running this function *N* times, we can simulate data for *N* agents:

```
agents <- replicate(N, simulator(...))
```

3.2. The emergence of factors as we know them

It could be expected that by being influenced by a common node the levels of the respective target nodes become correlated and thereby a trait-like cluster appears; in factor analysis (FA) or principal component analysis (PCA), the central node would appear as having the highest factor loading. Note that in this case the central node essentially serves the role of the latent variable in factor analysis, only that it is not really latent as it is one of the observed variables. If this idea works in the simulation—and it is really so trivial that it must work—then it suggests an interesting theoretical possibility: perhaps one of the indicators (e.g., item or facet) of a personality factor is the cause of other indicators rather than there being an underlying direct cause for all of them (that is, there may be an underlying cause for the central node but then its effect on nodes other than the central one is indirect, mediated by the central node). Of course, if the central node does not happen to be observed because, for example, the relevant item(s) or facet were not included in the questionnaire, a trait may still appear and then there is indeed an unobserved direct common cause for all of the measured variables.

Given that real scales may not have a single item or facet clearly having the highest loading in FA or PCA, it is likely that they reflect multiple central nodes. If the multiple nodes can influence each other (regardless of whether they belong to a common or different purported traits), they will tend to become correlated and so will their target nodes, resulting in a unidimensional-like scale (see also below).

3.3. A simplest possible simulation

To illustrate the principles by which networks can produce the appearance of statistical factors in the data, we run `simulator` with 10 nodes, specifying the fifth node as central and allowing it to influence all other nodes with strengths that are drawn from a normal distribution with a mean of .3 and standard deviation of .1. Let the noise coefficient to distort weight matrices of individual agents be .3. These parameter values are of course completely arbitrary. Let 5000 agents be simulated and subsequent analyses be carried out on this "sample". The relevant code is:

```
agents <- replicate(5000,
  simulator(k=10,
  m=.3, s=.1, n=.3, centrals=5,
  n.targets=9))
```

Subjecting the resulting data (i.e., `agents`) to PCA [`principal` from the *psych* package] results in a one-component solution that accounts for about 35–55% of variance in the ten variables that had initially been uncorrelated. The fifth variable has the highest correlation with the component and the further away from it the smaller the loadings generally become. Centrality analysis based on `qgraph` shows that the fifth node tends to have the highest betweenness and closeness centralities. Fitting a unidimensional reflective confirmatory factor analysis (CFA) model (one latent trait defined by the ten variables without residual correlations allowed) on the data tends to yield good model fit. CFA models can be fitted with the `cfa` function from *lavaan* package.

Obviously, there may be more than one central node responsible for a trait-like cluster as it is quite unlikely that the whole network is driven by a single central node. If they can influence (i.e., are close to) each other, they become correlated and so become their target nodes. As a result, a single trait-like cluster emerges. For example, there may be, say, two interrelated central nodes among those that coalesce into Neuroticism: anxiety and low mood (nodes 5 and 6 in the below code):

```
agents <- replicate(5000,
  simulator(k=10,
  m=.3, s=.1, n=.3, centrals=c(5,6),
  n.targets=9))
```

Or, one can place the central nodes more apart and see what happens then. If the number of target nodes is reduced, it is likely that the nodes form two clusters and that therefore two factors/components emerge:

```
agents <- replicate(5000,
  simulator(k=10, m=.3, s=.1, n=.3, centrals=c(3,8),
  n.targets=4))
```

3.4. A slightly more complex simulation

Using the same `simulator` function, we can also simulate the coalescence of nodes into multiple trait-like clusters. This can be done by placing a number of central nodes apart from each other, as said above. Another, much less contrived way to obtain data that looks realistic in terms of their correlational structure is to allow the clusters emerge naturally within the set of nodes, without any prespecified constraints.

As one possible scenario, any number of nodes can be central and influence any other number of nodes and these parameters can freely vary across agents; note only that the influence wanes with distance, as above. This setup is likely to result in data wherein each node has the strongest correlation with its immediate neighbors, whereas the correlations with other nodes wane with increasing distance. In other words, every node is somewhat correlated with every other node, but the correlations become increasingly higher as the distance between the nodes decreases. Using terms perhaps more familiar for personality psychologists, this corresponds to what can be called the hierarchical structure with (a) general factor(s) at the top and increasingly narrower factors below it. Such structure is evident in data, on which one can impose factor solutions with different numbers of factors (De Raad et al., 2014; Markon, Krueger, & Watson, 2005; Soto & John, 2014). In network terms, such traits correspond to areas of network with arbitrarily drawn borders. In fact, one can draw borders around an area of any size and location and call it a trait.

One way to obtain such data is the following. Note that in this simulation the average connection strength is also allowed to vary across agents, in addition to the network architectures being idiosyncratically distorted by the noise coefficient *n*. This is just to demonstrate the presence of this option.

```
agents <- replicate(5000,
simulator(k=30, n=.25, m=rnorm(1,.005,.001), s=.001,
centrals=sample(k, sample(k,1)),
  n.targets=sample(k-1,1)))
```

On the resulting data, for example, one can fit models with various numbers of components or factors extracted, starting from one and moving up to, say, ten. Curiously, all these different solutions are likely to yield "interpretable" loading patterns in the sense that nodes closer to each other in the network will always be more likely to belong to the same factors or components. What varies as a function of the number of factors or components extracted, is merely the size of the chunk of the network included in each factor or component. This simulation may give us one possible hint on what underlies the commonly observed hierarchical patterns of associations in personality ratings (Markon et al., 2005).

3.5. Extensions to more complicated cases

This section demonstrated only one way of simulating personality network data; there are likely to be other approaches that start from very different conceptual mechanisms and may or may not end up with similar results. Likewise, the demonstrated simulations were conceptually very simple and only addressed the coalescence of nodes into trait-like clusters. To the extent that the network perspective correctly reflects human personality, however, such networks are likely to function as dynamic systems that grow, obtain relative stability and interact with environment. Such networks can also be simulated using R (Mõttus et al., unpublished results), but this is beyond the scope of this section.

4. Discussion

Network approaches offer a rich trove of novel insights into the organization, emergence, and dynamics of personality. By integrating theoretical considerations (Cramer et al., 2010), simulation models (Mõttus et al., unpublished results; Van der Maas et al., 2006), and flexible yet user-friendly data-analytic techniques (Epskamp et al., 2012), network approaches have potential to achieve a tighter fit between theory and data analysis than has previously been achieved in personality research. At the present time, the basic machinery for generating, analyzing, and simulating networks is in place. Importantly, the R platform offers an impressive array of packages and techniques for the researcher to combine, and most of the important analyses are currently implemented. We hope that, in the present paper, we have successfully communicated the most important concepts and strategies that characterize the approach, and have done so in such a way that personality researchers may benefit from using network modeling in the analysis of their own theories and datasets.

In the present paper, we have applied network modeling to an illustrative dataset, with several intriguing results that may warrant further investigation. However, we do stress that many of our results are preliminary in nature. The primary reason for this is that current personality questionnaires are built according to psychometric methodology that is tightly coupled to factor analysis and classical test theory (Borsboom, 2005). This makes their behavior predictable from pure design specifications, which in turn limits their evidential value. That is, if one makes the a priori decision to have, say, 10 items per subscale, and selects items on the basis of their conformity to such a structure, many of the correlations found in subsequent research are simply built into the questionnaire. Therefore, it is hardly possible to tell to what extent results reflect a genuine structure, or are an artifact of the way personality tests are constructed. Trait perspectives are not immune to this problem, as in some cases the factors of personality may simply appear from questionnaire data because they have been carefully placed there. Future research should investigate potential solutions to this issue, for instance by considering variable sets consisting of ratings on the familiar personality-descriptive adjectives of a language, as in lexical studies (e.g., Ashton & Lee, 2005; De Raad et al., 2014; Goldberg, 1990; Saucier et al., 2014), and by comparing the characteristics of such networks to networks that emerge from questionnaire data.

An interesting question is whether all individuals are scalable on all items, as current methodology presumes. It is entirely possible, if not overwhelmingly likely, that certain items assess variables that simply do not apply to a given individual. Current psychometric methods have never come to grip with the "n.a." answer category, and in practice researchers simply force all individuals to answer all items. In networks, it is easier to deal with the n.a.-phenomenon, as nodes deemed to be inapplicable to a given person could simply be omitted from that person's network. This would yield personality networks that may differ in both structure and in size across individuals, an idea that resonates well with the notion that different people's personalities might in fact be also understood in terms of distinct theoretical structures (Borsboom et al., 2003; Cervone, 2005; Lykken, 1991). The application of experience sampling methodology and of other ways to gather information on dynamical processes personality may also offer an inroad into this issue (Bringmann et al., 2013; Fleeson, 2001; Hamaker, Dolan, & Molenaar, 2005).

The notion that network structures may differ over individuals, and that these differences may in fact be the key for understanding both idiosyncrasies and communalities in behavior, was illustrated in the simulation work reported in the present paper. Future research might be profitably oriented to questions such as (a) what kind of structural differences in networks could be expected based on substantive theory, (b) how such differences relate to well-established findings in personality research, and (c) which network growth processes are theoretically supported by developmental perspectives. Of course, ultimately, such theoretical models would have to be related back to empirical data of the kind discussed in the data-analysis part of this paper; therefore, a final highly important question is to derive testable implications from such perspectives. This includes the investigation of how we can experimentally or quasi-experimentally distinguish between explanations based on latent variables, and explanations based on network theory.

Ideally, these future developments are coupled with parallel developments in statistical and technical respects. Several important extensions of network models are called for. First, in this work we focused on the adaptive lasso, which is an effective method to extract a network from empirical data that has been profitably used in other fields (Krämer et al., 2009). However network analysis is a field in rapid evolution and alternative methods are being developed and studied. Among these, we consider particularly promising the graphical lasso (Friedman et al., 2008), for which adaptations exist that take into account the presence of latent variables in the network (Chandrasekaran, Parrilo, & Willsky, 2012; Yuan, 2012). Alternative methods based on Bayesian approaches have also been proposed and implemented (Mohammadi & Wit, 2014). Further research is needed to systematically compare these and other methods in the complex scenarios that are usually encountered in personality psychology. Second, as noted in this paper, many network analytics were originally designed for unweighted networks. Although some of the relevant analyses have now been extended to the weighted case (see Boccaletti et al., 2006; Costantini & Perugini, 2014; Opsahl et al., 2010), several other techniques still await such generalization. One important such set of techniques, which were also illustrated in the

Table A.1
Centrality indices.

Node	Dimension	Facet	Betweenness	Closeness	Strength
Hsi	Honesty–humility	Sincerity	5	2.66	0.73
Hfa	Honesty–humility	Fairness	**31**	3.03	**1.46**
Hga	Honesty–humility	Greed-avoidance	14	2.83	1.13
Hmo	Honesty–humility	Modesty	0	2.14	0.45
Efe	Emotionality	Fearfulness	6	2.70	1.03
Ean	Emotionality	Anxiety	2	3.04	1.10
Ede	Emotionality	Dependence	3	3.02	1.05
Ese	Emotionality	Sentimentality	17	3.17	**1.40**
Xss	Extraversion	Social self-esteem	11	3.11	**1.35**
Xsb	Extraversion	Social boldness	**23**	**3.33**	1.21
Xso	Extraversion	Sociability	7	**3.19**	1.07
Xli	Extraversion	Liveliness	12	3.12	1.29
Afo	Agreeableness vs. anger	Forgiveness	5	2.70	1.00
Age	Agreeableness vs. anger	Gentleness	5	2.66	0.80
Afl	Agreeableness vs. anger	Flexibility	14	2.90	1.02
Apa	Agreeableness vs. anger	Patience	5	2.85	0.85
Cor	Conscientiousness	Organization	7	3.09	0.99
Cdi	Conscientiousness	Diligence	**26**	**3.34**	1.30
Cpe	Conscientiousness	Perfectionism	5	3.13	1.26
Cpr	Conscientiousness	Prudence	**19**	**3.52**	**1.45**
Oaa	Openness to experience	Aesthetic appreciation	14	2.95	1.24
Oin	Openness to experience	Inquisitiveness	5	2.71	1.08
Ocr	Openness to experience	Creativity	10	3.00	1.26
Oun	Openness to experience	Unconventionality	3	2.63	0.98

Note. The four most central nodes according to each index are reported in bold. The closeness values are multiplied by 1000.

present work, deals with the determination of network structure. Both the theoretical definition of global structures, such as in terms of small-worlds, scale-free networks (Barabási & Bonabeau, 2003), and random networks, and the practical determination of these structures (e.g., through coefficients such as small-worldness or through fitting functions on the degree distribution) are based on unweighted networks. It would be highly useful if these notions, and the accompanying techniques, would be extended to the weighted network case. Another technical improvement that should be within reach is how to deal with data that likely reflect mixtures of distinct networks (as in the second simulation in the current paper). In the case of time series data, such approaches have already been formulated through the application of mixture modeling (Bringmann et al., 2013); however, statistical techniques suited to this problem may also be developed for the case of cross-sectional data. The issue is important in terms of modeling idiosyncrasies in behavior, but may also be key in terms of relating normal personality to psychopathology (Cramer et al., 2010). Naturally, this includes the question of how we should think about the relation between normal personality and personality disorders.

Acknowledgments

This work was supported by Fondazione Cariplo research Grant "Dottorato ad alta formazione in Psicologia Sperimentale e Neuroscienze Cognitive" (Advanced education doctorate in experimental psychology and cognitive neurosciences), Grant Number 2010-1432 (awarded to Giulio Costantini) and by NWO "research talent" Grant Number 406-11-066 (awarded to Sacha Epskamp).

Appendix A

See Table A.1.

Table B.1
Factor loadings. Factors are labeled according to their highest loadings.

	E	C	O	X	H	A	Uniq.	Compl.	Smc
Hsi	−.05	.11	.11	.05	**.60**	−.05	.61	1.17	.26
Hfa	.14	.22	.15	−.04	**.63**	.19	.48	1.69	.39
Hga	.11	−.01	.24	.03	**.54**	.14	.62	1.65	.29
Hmo	.04	−.01	.05	−.05	**.44**	.07	.79	1.12	.16
Efe	**.48**	.03	−.16	−.22	−.07	−.04	.69	1.72	.27
Ean	**.55**	.17	.08	−.12	.11	−.11	.63	1.54	.30
Ede	**.66**	−.01	−.11	−.08	−.01	−.03	.55	1.10	.34
Ese	**.68**	.07	.02	.10	.13	.08	.50	1.18	.36
Xss	−.36	.18	.06	**.53**	−.08	.00	.54	2.14	.38
Xsb	−.05	.08	.07	**.63**	−.02	−.25	.52	1.40	.36
Xso	.17	−.02	.03	**.65**	.06	.01	.55	1.17	.33
Xli	−.11	.06	.02	**.67**	.00	.12	.52	1.13	.37
Afo	.09	−.09	.04	.13	.16	**.43**	.75	1.68	.20
Age	.09	−.06	−.02	.04	.13	**.54**	.68	1.21	.23
Afl	−.06	−.02	−.01	−.10	.06	**.67**	.53	1.08	.29
Apa	−.11	.10	.14	−.01	.09	**.49**	.71	1.45	.22
Cor	.01	**.73**	−.07	.06	.01	.00	.46	1.03	.37
Cdi	.19	**.58**	.19	.21	.18	−.03	.51	1.99	.41
Cpe	.08	**.70**	.18	.05	.06	−.08	.46	1.22	.41
Cpr	−.21	**.52**	.12	−.12	.15	.12	.62	1.87	.32
Oaa	−.04	.17	**.71**	−.04	.15	.04	.44	1.23	.42
Oin	−.25	.09	**.59**	.04	.15	−.01	.56	1.55	.35
Ocr	.15	.01	**.62**	.14	.01	.08	.56	1.26	.32
Oun	−.07	.01	**.57**	.10	.11	−.08	.65	1.22	.29

Note. E = loading on emotionality, C = loading on conscientiousness, O = loading on openness to experience, X = loading on extraversion, H = loading on honesty–humility, A = loading on agreeableness vs. anger. Smc = squared multiple correlation of each facet with all the others. Uniq. = uniqueness. Compl. = Hofmann's row-complexity index (1978).

Appendix B

See Table B.1.

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jrp.2014.07.003.

References

Ashton, M., & Lee, K. (2005). A defence of the lexical approach to the study of personality structure. *European Journal of Personality, 19*, 5–24. http://dx.doi.org/10.1002/per.541.
Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*(2), 150–166. http://dx.doi.org/10.1177/1088868306294907.
Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*(4), 340–345. http://dx.doi.org/10.1080/00223890902935878.
Barabási, A. L., & Bonabeau, E. (2003). Scale-free networks. *Scientific American, 288*, 60–69.
Barrat, A., Barthelémy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America, 101*(11), 3747–3752.
Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4–5), 175–308. http://dx.doi.org/10.1016/j.physrep.2005.10.009.
Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology, 2*, 113–120. http://dx.doi.org/10.1080/0022250X.1972.9989806.
Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology, 92*(5), 1170–1182. http://dx.doi.org/10.1086/228631.
Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks, 29*(4), 555–564. http://dx.doi.org/10.1016/j.socnet.2007.04.002.
Bonacich, P., & Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks, 23*(3), 191–201. http://dx.doi.org/10.1016/S0378-8733(01)00038-7.
Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. http://dx.doi.org/10.1016/j.socnet.2004.11.008.
Borgatti, S. P., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks, 28*(4), 466–484. http://dx.doi.org/10.1016/j.socnet.2005.11.005.
Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge: Cambridge University Press.
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. http://dx.doi.org/10.1146/annurev-clinpsy-050212-185608.
Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE, 6*(11), e27407. http://dx.doi.org/10.1371/journal.pone.0027407.
Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*(2), 203–219. http://dx.doi.org/10.1037/0033-295X.110.2.203.
Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology, 25*(2), 163–177. http://dx.doi.org/10.1080/0022250X.2001.9990249.
Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks, 30*(2), 136–145. http://dx.doi.org/10.1016/j.socnet.2007.11.001.
Brandes, U., & Fleischer, D. (2005). Centrality measures based on current flow. In V. Diekert & B. Durand (Eds.). *STACS 2005* (Vol. 3404, pp. 533–544). Berlin: Springer. http://dx.doi.org/10.1007/978-3-540-31856-9_44.
Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., et al. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*(4), e60188. http://dx.doi.org/10.1371/journal.pone.0060188.
Butts, C. T. (2008a). Social network analysis: A methodological introduction. *Asian Journal of Social Psychology, 11*(1), 13–41. http://dx.doi.org/10.1111/j.1467-839X.2007.00241.x.
Butts, C. T. (2008b). Social network analysis with sna. *Journal of Statistical Software, 24*(6), 1–51.
Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. http://dx.doi.org/10.1037/h0046016.
Cervone, D. (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology, 56*, 423–452. http://dx.doi.org/10.1146/annurev.psych.56.091103.070133.
Chandrasekaran, V., Parrilo, P. A., & Willsky, A. S. (2012). Latent variable graphical model selection via convex optimization. *The Annals of Statistics, 40*(4), 1935–1967. http://dx.doi.org/10.1214/11-AOS949.

Costa, P. T., & McCrae, R. R. (1992). *Professional manual: Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Costantini, G., & Perugini, M. (2014). Generalization of clustering coefficients to signed correlation networks. *PLoS ONE, 9*(2), e88669. http://dx.doi.org/10.1371/journal.pone.0088669.

Cox, D. R., & Wermuth, N. (1993). Linear dependencies represented by chain graphs. *Statistical Science, 8*(3), 204–218.

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., et al. (2012a). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality, 26*(4), 414–431. http://dx.doi.org/10.1002/per.1866.

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., et al. (2012b). Measurable like temperature or mereological like flocking? On the nature of personality traits. *European Journal of Personality, 26*(4), 451–459. http://dx.doi.org/10.1002/per.1879.

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences, 33*(2–3), 137–193. http://dx.doi.org/10.1017/S0140525X09991567.

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2004). Error and attack tolerance of complex networks. *Physica A, 340*(1–3), 388–394. http://dx.doi.org/10.1016/j.physa.2004.04.031.

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems, 1695*(5).

De Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with Pajek* (2nd ed.). Cambridge: Cambridge University Press.

De Raad, B., Barhelds, D. P. H., Timmerman, M. E., De Roover, K., Mlačić, B., & Church, A. T. (2014). Towards a pan-cultural personality structure: Input from 11 psycholexical studies. *European Journal of Personality*. http://dx.doi.org/10.1002/per.1953.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*(1), 269–271. http://dx.doi.org/10.1007/BF01386390.

Dolan, C. V., Oort, F. J., Stoel, R. D., & Wicherts, J. M. (2009). Testing measurement invariance in the target rotated multigroup exploratory factor model. *Structural Equation Modeling, 16*(2), 20. http://dx.doi.org/10.1080/10705510902751416.

Drton, M., & Perlman, M. D. (2004). Model selection for Gaussian concentration graphs. *Biometrika, 91*(3), 591–602. http://dx.doi.org/10.1093/biomet/91.3.591.

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*(4), 1–18.

Fan, J., Feng, Y., & Wu, Y. (2009). Network exploration via the adaptive LASSO and SCAD penalties. *The Annals of Applied Statistics, 3*(2), 521–541. http://dx.doi.org/10.1214/08-AOAS215.

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*(6), 1011–1027. http://dx.doi.org/10.1037//0022-3514.80.6.1011.

Foygel, R., & Drton, M. (2011). *Bayesian model choice and information criteria in sparse generalized linear models*. arXiv Preprint arXiv:1112.5635.

Franić, S., Borsboom, D., Dolan, C. V., & Boomsma, D. I. (2013). The Big Five personality traits: Psychological entities or statistical constructs? *Behavior Genetics*. http://dx.doi.org/10.1007/s10519-013-9625-7.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239. http://dx.doi.org/10.1016/0378-8733(78)90021-7.

Freeman, L. C., Borgatti, S. P., & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks, 13*(2), 141–154. http://dx.doi.org/10.1016/0378-8733(91)90017-N.

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

Funder, D. C. (1991). Global traits: A neo-allportian approach to personality. *Psychological Science, 2*(1), 31–39. http://dx.doi.org/10.1111/j.1467-9280.1991.tb00093.x.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229. http://dx.doi.org/10.1037/0022-3514.59.6.1216.

Grimmett, G. (2010). *Probability on graphs: Random processes on graphs and lattices*. Cambridge: Cambridge University Press.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research, 40*(2), 207–233. http://dx.doi.org/10.1207/s15327906mbr4002_3.

Hoekstra, H. A., De Fruyt, F., & Ormel, J. (2003). *NEO-PI-R/NEO-FFI: Big Five personality inventory-manual*. Lisse: Swetz & Zeitlinger.

Hofmann, R. J. (1978). Complexity and simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research, 13*(2), 247–250. http://dx.doi.org/10.1207/s15327906mbr1302_9.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE, 3*(4), e0002051. http://dx.doi.org/10.1371/journal.pone.0002051.

Jackson, J. J., Bogg, T., Walton, K. E., Wood, D., Harms, P. D., Lodi-Smith, J., et al. (2009). Not all conscientiousness scales change alike: A multimethod, multisample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology, 96*(2), 446–459. http://dx.doi.org/10.1037/a0014156.

Jeong, H., Mason, S. P., Barabási, A. L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature, 411*(6833), 41–42. http://dx.doi.org/10.1038/35075138.

Kalna, G., & Higham, D. J. (2007). A clustering coefficient for weighted networks, with application to gene expression data. *AI Communications, 20*(4), 263–271.

Kindermann, R., & Snell, J. (1980). *Markov random fields and their applications*. Providence: American Mathematical Society. http://dx.doi.org/10.1090/conm/001.

Kolaczyk, E. D. (2009). *Statistical analysis of network data: Methods and models*. New York: Springer. http://dx.doi.org/10.1007/978-0-387-88146-1.

Krämer, N., Schäfer, J., & Boulesteix, A.-L. (2009). Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC Bioinformatics, 10*, 384. http://dx.doi.org/10.1186/1471-2105-10-384.

Kunegis, J., Lommatzsch, A., & Bauckhage, C. (2009). The Slashdot Zoo: Mining a social network with negative edges. In *Proceedings of the 18th international conference on world wide web* (pp. 741–750). http://dx.doi.org/10.1145/1526709.1526809.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics, 9*, 559. http://dx.doi.org/10.1186/1471-2105-9-559.

Langfelder, P., & Horvath, S. (2012). Fast R functions for robust correlations and hierarchical clustering. *Journal of Statistical Software, 46*(11), 1–17.

Latora, V., Nicosia, V., & Panzarasa, P. (2013). Social cohesion, structural holes, and a tale of two measures. *Journal of Statistical Physics, 151*(3–4), 745–764. http://dx.doi.org/10.1007/s10955-013-0722-z.

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

Lee, J. (2012). Correlation and causation in the study of personality. *European Journal of Personality, 26*(4), 372–390. http://dx.doi.org/10.1002/per.1863.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Signed networks in social media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1361–1370. http://dx.doi.org/10.1145/1753326.1753532.

Lykken, D. T. (1968). Statistical significance of psychological research. *Psychological Bulletin, 70*(3, Pt.1), 151–159. http://dx.doi.org/10.1037/h0026141.

Lykken, D. (1991). What's wrong with psychology, anyway? In D. Cicchetti & W. M. Grove (Eds.). *Thinking clearly about psychology* (Vol. 1). Minneapolis: University of Minnesota Press.

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*(1), 139–157. http://dx.doi.org/10.1037/0022-3514.88.1.139.

McCrae, R. R., & Costa, P. T. J. (2008). Empirical and theoretical status of the five-factor model of personality traits. In G. Boyle, G. Matthews, & D. Saklofske (Eds.). *Sage handbook of personality theory and assessment* (Vol. 1, pp. 273–294). Los Angeles: Sage. http://dx.doi.org/10.4135/9781849200462.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*(1), 195–244. http://dx.doi.org/10.2466/PR0.66.1.195-244.

Milgram, S. (1967). The small world problem. *Psychology Today, 1*(1), 61–67.

Mohammadi, A., Wit, E. C., 2014. *Bayesian structure learning in sparse Gaussian graphical models.* arXiv preprint arXiv:1210.5371v6.

Mõttus, R., Penke, L., Murray, A. L., Booth, T., & Allerhand, M. (unpublished results). *Personality differences without common-cause latent factors are possible and can explain key findings in personality psychology*.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*(6), 1213–1233. http://dx.doi.org/10.1016/j.jrp.2007.02.003.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45*(2), 167–256. http://dx.doi.org/10.1137/S003614450342480.

Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E, 70*(5), 056131. http://dx.doi.org/10.1103/PhysRevE.70.056131.

Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks, 27*(1), 39–54. http://dx.doi.org/10.1016/j.socnet.2004.11.009.

Newman, M. E. J. (2008). The physics of networks. *Physics Today, 61*(11), 33–38. http://dx.doi.org/10.1063/1.3027989.

Newman, M. E. J. (2010). *Networks: An introduction*. New York: Oxford University Press.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*, 026113. http://dx.doi.org/10.1103/PhysRevE.69.026113.

Onnela, J.-P., Saramäki, J., Kertész, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E, 71*(6), 065103. http://dx.doi.org/10.1103/PhysRevE.71.065103.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245–251. http://dx.doi.org/10.1016/j.socnet.2010.03.006.

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks, 31*(2), 155–163. http://dx.doi.org/10.1016/j.socnet.2009.02.002.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge: MIT Press.

Pettersson, E., & Turkheimer, E. (2010). Item selection, evaluation, and simple structure in personality data. *Journal of Research in Personality, 44*(4), 407–420. http://dx.doi.org/10.1016/j.jrp.2010.03.002.

Pourahmadi, M. (2011). Covariance estimation: The GLM and regularization perspectives. *Statistical Science, 26*(3), 369–387. http://dx.doi.org/10.1214/11-sts358.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
Revelle, W. (2013). *Psych: Procedures for personality and psychological research. R package version 1.4.1*.
Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*(4), 581–603. http://dx.doi.org/10.1007/BF02289527.
Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K., & Kertész, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E, 75*(2), 027105. http://dx.doi.org/10.1103/PhysRevE.75.027105.
Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L., et al. (2014). A basic bivariate structure of personality attributes evident across nine languages. *Journal of Personality, 82*(1), 1–14. http://dx.doi.org/10.1111/jopy.12028.
Schlegel, K., Grandjean, D., & Scherer, K. R. (2013). Constructs of social and emotional effectiveness: Different labels, same content? *Journal of Research in Personality, 47*(4), 249–253. http://dx.doi.org/10.1016/j.jrp.2013.02.005.
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*(1), 43–53. http://dx.doi.org/10.1016/j.newideapsych.2011.02.007.
Servaas, M. N., Riese, H., Ormel, J., & Aleman, A. (2014). The neural correlates of worry in association with individual differences in neuroticism. *Human Brain Mapping*. http://dx.doi.org/10.1002/hbm.22476.
Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology, 99*(2), 330–343. http://dx.doi.org/10.1037/a0019796.
Soto, C. J., & John, O. P. (2014). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality, 82*(3), 182–199. http://dx.doi.org/10.1111/jopy.12044.
Steyer, R. (2012). Does network theory contradict trait theory? *European Journal of Personality, 26*(4), 447–448. http://dx.doi.org/10.1002/per.1877.
Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality, 34*(4), 397–423. http://dx.doi.org/10.1006/jrpe.2000.2292.
Valente, T. W. (2012). Network interventions. *Science, 337*(6090), 49–53. http://dx.doi.org/10.1126/science.1217330.
van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*(4), 842–861. http://dx.doi.org/10.1037/0033-295X.113.4.842.
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440–442. http://dx.doi.org/10.1038/30918.
Yuan, M. (2012). Discussion: Latent variable graphical model selection via convex optimization. *The Annals of Statistics, 40*(4), 1968–1972. http://dx.doi.org/10.1214/12-AOS979.
Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology, 4*(1). http://dx.doi.org/10.2202/1544-6115.1128.
Ziegler, M., Booth, T., & Bensch, D. (2013). Getting entangled in the nomological net. Thoughts on validity and conceptual overlap. *European Journal of Psychological Assessment, 29*, 157–161. http://dx.doi.org/10.1027/1015-5759/a000173.
Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association, 101*(476), 1418–1429. http://dx.doi.org/10.1198/016214506000000735.