European Journal of Personality, *Eur. J. Pers.* (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/per.1869

The Definition of Components and the Use of Formal Indexes are Key Steps for a Successful Application of Network Analysis in Personality Psychology

GIULIO COSTANTINI* and MARCO PERUGINI
Faculty of Psychology, University of Milan—Bicocca, Italy

Abstract: Conceiving personality as a network provides an interesting theoretical framework and a promising methodological perspective. The application of network analysis to personality psychology however is not straightforward, and some issues require careful consideration. We argue that the definition of components within networks cannot be limited to single items, and more work is needed to reflect the inherently hierarchical and indefinite nature of components. Additionally, we argue that formal empirical indexes must be clearly defined and consistently used to describe properties of personality networks. Copyright © 2012 John Wiley & Sons, Ltd.

Cramer and colleagues (2012) questioned the idea that latent dimensions of personality reflect underlying causes of behaviour (McCrae & Costa, 2008a) and proposed an alternative theoretical framework in which personality structure emerges from causal, homeostatic and logical relationships among different components. We agree on many basic aspects of this conception of personality as a network and think that it has a sound potential. We confess that, like Cramer and colleagues, we found epistemologically hollow claims such as ‘Extraversion causes party-going’, especially when Extraversion is measured including these same behaviours that is subsequently deemed to cause. If one takes the definition of causality seriously, there is no question that an *explanans* must be different from an *explanandum*, yet this basic principle is sometimes violated in personality research. From this perspective, a network analysis applied to personality structure opens up the possibility to investigate causes of behaviours, although we think that it is not the only possible approach. For example, there is work on the interplay between states and traits (e.g. Fleeson, 2001; Steyer, Schmitt, & Eid, 1999) that can provide a sound framework for causal analyses in which specific traits are associated to specific behaviours through specific mechanisms at work in specific situations (e.g. Baumert, Gollwitzer, Staubach, & Schmitt, 2011). We would be curious to know how Cramer and colleagues compare their network analysis with state–trait existing theoretical models, an issue that is only hinted at in the closing of their target article. Despite our general agreement with much of this contribution, we focus the rest of this commentary on two unresolved main issues.

1. *What is a component?* While in different applications of network analysis components may be readily defined as, for instance, individuals (Travers & Milgram, 1969), web pages (Albert, Jeong, & Barabási, 1999) or proteins (Vazquez, Flammini, Maritan, & Vespignani, 2003), the definition of components within personality psychology is more complex. From a theoretical point of view, Cramer and colleagues defined components as thoughts, feelings and acts that are associated with a unique causal system. From a measurement point of view, components are equated to single questionnaire items, as opposed to latent variables: Whereas latent variables are conceived as aggregations that emerge from the interconnections of different components, items are meant to reflect directly single basic units of cognitions, thoughts and acts. Even single items however constitute already aggregations of different phenomena that moreover are inextricably connected to the specific way in which they have been measured. One cannot not aggregate but, at best, can decide which level of aggregation is the most informative. This opens up at least three relevant issues. First, different units (components) may be useful for different purposes. Single items may be useful to provide a fine-grained understanding of the dynamic of a personality dimension, but they might also provide unreliable and unstable information. Aggregates of items (e.g. parcels, facets) may imply a loss in terms of definition but a gain in terms of reliability of the findings. Future work is needed to compare the relative merits of network representations at different levels of aggregation. Second, a unit or component does not exist outside of the specific measurement method that is used. Components can be items measured with self-reports or peer reports, behavioural indicators and so on. Each measurement method has advantages and disadvantages, and each provide only partial information about a personality structure. Future work is needed in which multiple methods are used to provide converging evidence of network analyses of a personality structure. Third, given the existence of different levels of aggregation of basic components, it becomes not obvious to what extent a network analysis can simply substitute a factor analysis or rather it represents a complementary but not alternative statistical tool. At some level, subcomponents (whatever they might be) will still need to be aggregated.

*Correspondence to: Giulio Costantini, Faculty of Psychology, University of Milan—Bicocca, Italy.
E-mail: g.costantini1@campus.unimib.it

There is the need therefore to articulate the interplay between the two statistical tools.

2. *Network properties*. A researcher properly using factor analysis or similar statistical models would not rely on a graphical inspection to draw substantial conclusions but would consider instead specific quantitative indicators (e.g. loadings). We do not see why it should be any different for a network analysis. Take for instance the concept of centrality. In the target paper, the graphical position of the items within networks is often straightforwardly interpreted as centrality. However, even though the *qgraph* algorithm (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) allows for an intuitive visual representation of correlation matrices, using the position of nodes within graphs as an index of centrality may be misleading. Position of nodes depends also on the overall graph (e.g. being at the centre of one factor is different than being at the centre of two), and it can be a by-product of a two-dimensional representation of complex relations (Fruchterman & Reingold, 1991). Centrality is not a feature of a graphical representation, but it is better characterized as a class of properties that do not always covary. As an example, in Figure 1, we report a graph generated with simulated data[1] relative to 201 variables. Four clusters of 50 variables each are at the periphery of the graph, while variable 'X' is at the very centre. Note that variable X is highly correlated with only one variable for each of the four clusters (Y1 to Y4) and nothing at all with the remaining 196 variables. We inspected some formal indexes of centrality (Freeman, 1978) by means of R package *tnet* (Opsahl, Agneessens, & Skvoretz, 2010) to verify in which sense variable X is central. Values of weighted degree, betweenness and closeness are reported in Table 1. If we look at closeness and betweenness centrality, variable X is central. However, if we look at degree centrality, X is the most peripheral variable. Moreover, the four variables connected with X (Y1 to Y4) rank second in terms of betweenness centrality far above all other 196 variables, yet they are visually represented nowhere near the centre of the graph. This admittedly extreme example helps to highlight two issues. First, a network analysis should go beyond the inspection of a graph and focus instead on specific quantitative indicators. Second, which one of these indicators is most informative and in respect to what property must be clearly articulated as they can provide entirely different and non-converging information.

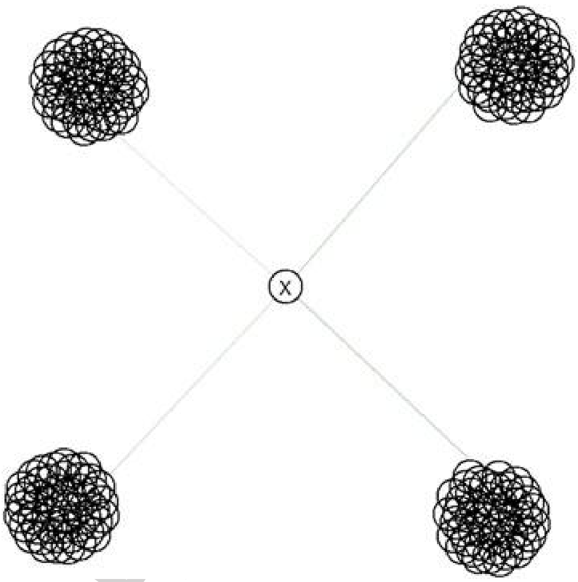

Figure 1. Graph built with qgraph using simulated data.

Table 1. Indexes of centrality of nodes in Figure 1

	Weighted degree	Weighted betweenness	Weighted closeness
Variable X	2.66	15 000	~~0.0026~~
Variables connected to X (Y1 to Y4, *n* = 4)	*M* = 40.84, *SD* = 0.56	7399†	*M* = ~~0.0018~~ ~~(~~*SD* = 0.00010~~)~~
Remaining variables (*n* = 196)	*M* = 40.85, *SD* = 0.89	0†	*M* = ~~0.0017~~ ~~(~~*SD* = 0.00007~~)~~

Note: Centrality indexes were computed with R package *tnet*, using functions degree_w, betweenness_w and closeness_w.
†The value is exactly the same for all variables.

In conclusion, we are excited by the potential of network analysis and think that it can push the personality field towards important theoretical and methodological advancements, but we also think that substantial work remains to be carried out before it can move beyond being a metaphor to become an actual research tool.

[1]The R code and the correlation matrix are available upon request.

DOI: 10.1002/per