## Dynamic Exploratory Graph Analysis

Dynamic exploratory graph analysis (dynEGA; Golino, Christensen et al., 2020) is a dynamic dimension reduction approach based on network psychometrics. The approach works by first applying the generalized local linear approximation (GLLA; Boker et al., 2010; Deboeck, Montpetit, Bergeman, & Boker, 2009) to estimate how variables change over time. Using time-delayed embedding, derivatives are estimated corresponding to the rate of change (velocity; first derivative) and speed of the rate of change (acceleration; second derivative). From these derivatives, exploratory graph analysis (EGA; Golino & Epskamp, 2017; Golino, Shi et al., 2020) can be applied to individuals’, groups’, and sample’s data.

### Generalized Local Linear Approximation

A time-delay embedding matrix for each variable is constructed by taking the time series of a single variable, , and expanding it into a matrix with consecutive offset sequences ( or lag) of *n* embedding dimensions. Using an example of a time series with the length of , , and five embedding dimensions, we can construct a time-delay embedding matrix ():

where is the time-delay embedding matrix with five dimensions for variable . Time-delay embedding matrices are constructed for each variable in the dataset. After, derivatives can be estimated using GLLA. These derivatives are estimated as:

where is a matrix of derivative estimates, is the time-delay embedding matrix, and is a matrix a where is the number of embedding dimensions and is the (maximum) order of the derivative. Each column of the weight matrix, , is estimated with derivatives going from zero to , :

where is the time between consecutive observations in the time series, is a vector from one to the number of embedded dimensions (e.g., 5), and is the order of the derivative of interest. This procedure is carried out over each variable and results in a matrix of derivatives, .

To estimate the sample’s EGA structure, each participant’s derivative matrix, , can be “stacked” or appended row-wise. For group structures, stacks can be created that correspond to a grouping variable. Finally, each participant’s can be analyzed individual to obtain their structure.

### Exploratory Graph Analysis

EGA is a recently developed method to estimate the number of dimensions in multivariate data using undirected network models (Golino & Epskamp, 2017; Golino et al., 2020). EGA first applies a network estimation method followed by a community detection algorithm for weighted networks (Fortunato, 2010). EGA has been shown to be as accurate or more accurate than more traditional factor analytic methods such as parallel analysis (Christensen, Garrido, & Golino, 2021; Golino et al., 2020). For our study, Pearson’s correlations were used to compute the associations between derivatives of our variables.

**Network Estimation Method.** This study applied the graphical least absolute shrinkage and selection operator (GLASSO; Friedman, Haste, & Tibshirani, 2008, 2014), which estimates a Gaussian Graphical Model (GGM; Lauritzen, 1996) where nodes (circles) represent variables and edges (lines) represent the conditional dependence (or partial correlations) between nodes given all other nodes in the network. The least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996) of the GLASSO is a regularization technique that reduces parameter estimates with some estimates becoming exactly zero.

The LASSO uses a parameter called lambda (λ), which controls the sparsity of the network. Lower values of λ remove fewer edges, increasing the possibility of including spurious correlations, while larger values of λ remove more edges, increasing the possibility of removing relevant edges. When λ = 0, the estimates are equal to the ordinary least squares solution for the partial correlation matrix. In this study, the ratio of the minimum and maximum λ was set to 0.1.

The popular approach in the network psychometrics literature is to compute models across several values of λ (usually 100) and to select the model that minimizes the extended Bayesian information criterion (EBIC; Chen & Chen, 2008; Epskamp & Fried, 2018). The EBIC model selection uses a hyperparameter gamma (γ) to control how much it prefers simpler models (i.e., models with fewer edges; Foygel & Drton, 2010). Larger γ values lead to simpler models, while smaller γ values lead to denser models. When γ = 0, the EBIC is equal to the Bayesian information criterion. In this study, γ was set to 0.5. In network psychometrics literature, this approach has been termed *EBICglasso* and is applied using the *qgraph* package.

**Community Detection Algorithm.** The Louvain algorithm (also referred to as Multi-level; Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) is one of the most commonly applied in network science (Gates, Henry, Steinley, & Fair, 2016). The algorithm begins by randomly sorting nodes into communities with their neighbors and then uses modularity (Newman, 2006) to iteratively optimize its community partitions by exchanging nodes between communities and evaluating the change in modularity until it no longer improves. Then, the algorithm collapses the communities into latent nodes and identifies edge weights with other observed and latent nodes, which provides a multi-level structure (Gates et al., 2016). In this study, the algorithm was not used to identify hierarchical community structures in the network. The Louvain algorithm was implemented using the *igraph* package (Csardi & Nepusz, 2006) in R. It's important to note that the algorithm implemented in *igraph* is deterministic; however, other implementations are not (Gates et al., 2016).

To check for unidimensionality, we used the Leading Eigenvalue algorithm (Newman, 2006) approach, which was applied to the zero-order correlation matrix (Christensen et al., 2021). The algorithm is based on spectral properties of the network using the eigenvector of the first eigenvalue of the modularity matrix to determine optimal community structures. The algorithm begins by computing the first eigenvector of the modularity matrix and then splits the network into two communities with the aim of improving modularity. This process iteratively unfolds until there is no longer improvement in modularity. If the algorithm returns one dimension, then the result unidimensional; otherwise, the standard EGA procedure is followed (Christensen et al., 2021). The Leading Eigenvalue algorithm was implemented using the *igraph* package.

## Loadings Comparison Test

For our study, we set out to determine whether between-person and within-person personality structures were more like a factor or network model. Although the between-person structure may be closer to one model over the other, all within-person structures need not be closer to that same model.

To evaluate whether the data at both the between-person and within-person level were structured more like a factor or network model, we employed the Loadings Comparison Test (LCT; Christensen & Golino, 2021). The LCT uses a neural network ensemble to predict whether data are generated from a factor or network model. The inputs of the neural networks are factor and network loadings (see Christensen & Golino, in press for details) from the same data using the same number of dimensions. The neural networks were trained on 480,000 datasets generated from a factor or small-world network model. They were then validated on different datasets generated from a factor, small-world network, and random network model (Christensen & Golino, 2021). The LCT demonstrated robust generalizability and was consistent with real-world examples of data structured more like a factor and network model.

The LCT makes three predictions: empirical, bootstrap, and proportion. The empirical prediction is based solely on the empirical data. The bootstrap and proportion predictions use a bootstrap approach by using the empirical covariance structure to generate replicate samples (e.g., 100) from a multivariate normal distribution. The bootstrap prediction makes a single prediction on the average neural network input parameters, while the proportion prediction makes a prediction on each replicate sample. The proportion prediction allows a more nuanced interpretation of the number of samples that were predicted to be either a factor or network model, with values .50 being favored towards that model. The proportion prediction was the most robust in the empirical and simulation validation performed by Christensen and Golino (2021) and was therefore the approach we applied in this study.

## Data Analysis

dynEGA and LCT were applied using the *EGAnet* package (version 0.9.9; Golino & Christensen, 2021) in R (version 4.0.5; R Core Team, 2021). dynEGA and associated results were visualized using the *GGally* (version 2.1.1; Schloerke et al., 2021), *ggplot2* (version 3.3.3; Wickham, 2020), and *qgraph* (version 1.6.9; Epskamp et al., 2012) packages in R.

# References

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*, P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>

Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. (2010). Generalized local linear approximation of derivatives from time series. In S. M. Chow, E. Ferrer, & F. Hsieh (Eds.), *Statistical methods for modeling* *human dynamics: An interdisciplinary dialogue* (pp. 161–178). New York, NY: Routledge.

Christensen, A. P., Garrido, L. E., & Golino, H. (2021). Comparing community detection algorithms in psychological data: A Monte Carlo simulation. *PsyArXiv*. <https://doi.org/10.31234/osf.io/hz89e>

Christensen, A. P., & Golino, H. (2021). Factor or network model? Predictions from neural networks. *Journal of Behavioral Data Science*, *1*(1), 85–126. <https://doi.org/10.35566/jbds/v1n1/p5>

Christensen, A. P., & Golino, H. (in press). On the equivalency of factor and network loadings. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01500-6>

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*, 1–9. Retrieved from <https://pdfs.semanticscholar.org/1d27/44b83519657f5f2610698a8ddd177ced4f5c.pdf>

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*, 1–18. <https://doi.org/10.18637/jss.v048.i04>

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

Deboeck, P. R., Montpetit, M. A., Bergeman, C., & Boker, S. M. (2009). Using derivative estimates to describe intraindividual variability at multiple time scales. *Psychological Methods*, *14*(4), 367–386. <https://doi.org/http://dx.doi.org/10.1037/a0016622>

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *3–5*, 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>

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

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

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

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A Monte Carlo evaluation of weighted community detection algorithms. *Frontiers in Neuroinformatics*, *10*, 45. <https://doi.org/10.3389/fninf.2016.00045>

Golino, H., & Christensen, A. P. (2021). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics. Retrieved from <https://cran.r-project.org/package=EGAnet>

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

Golino, H., Shi, D., Garrido, L. E., Christensen, A. P., Nieto, M. D., Sadana, R., . . . Martinez-Molina, A. (2020b). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292–230. <https://doi.org/10.1037/met0000255>

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

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*, 8577–8582. <https://doi.org/10.1073/pnas.0601602103>

R Core Team (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from [https://www.R-project.org](https://www.r-project.org/)

Schloerke, Cook, Larmarange, Briatte, & Marbach (2021). GGally: Extention to 'ggplot2'. Retrieved from <https://cran.r-project.org/package=GGally>

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>

Wickham, H. (2020). *ggplot2: Elegant graphics for data analysis.* New York, NY: Springer. Retrieved from [https://ggplot2-book.org](https://ggplot2-book.org/)