Towards a psychology of individuals: The ergodicity information index and a bottom-up approach for finding generalizations

Hudson Golino[1], John Nesselroade[1], & Alexander P. Christensen[2]

[1] Department of Psychology, University of Virginia

[2] Department of Psychology and Human Development, Peabody College, Vanderbilt University

Author Note

Acknowledgements: The authors did not preregister this study. All data, code, and materials can be found on the Open Science Framework. Hudson Golino, Alexander P. Christensen. Simulations used in this study are available via a Shiny app: https://masked.shinyapps.io/eii_app/

The authors made the following contributions. Hudson Golino: Conceptualization, Formal Analysis, Methodology, Resources, Software, Validation, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing; John Nesselroade: Writing - Original Draft Preparation, Writing - Review & Editing; Alexander P. Christensen: Conceptualization, Formal Analysis, Methodology, Resources, Software, Validation, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Hudson Golino, 485 McCormick Road, Gilmer Hall, Room 102, Charlottesville, VA 22903. E-mail: hfg9s@virginia.edu

Abstract

In the last half of the 20th century, psychology and neuroscience have experienced a renewed interest in intraindividual variation. To date, there are few quantitative methods to evaluate whether a population (between-person) structure is likely to hold for individual people, often referred to as ergodicity. We introduce a new network information theoretic metric, the ergodicity information index (EII), that quantifies the amount of information lost by representing all individuals with a between-person structure. A Monte Carlo simulation demonstrated that EII can effectively delineate between ergodic and nonergodic systems. A bootstrap test is derived to statistically determine whether the empirical data is likely generated from an ergodic process. When a process is identified as nonergodic, then it's possible that a mixture of groups exist. To evaluate whether groups exist, we develop an information theoretic clustering method to detect groups. Finally, two empirical examples are presented using intensive longitudinal data from personality and neuroscience domains. Both datasets were found to be nonergodic, and meaningful groupings were identified in each dataset. Subsequent analysis showed that some of these groups are ergodic, meaning that the individuals can be represented with a single population structure without significant loss of information. Notably, in the neuroscience data, we could correctly identify two clusters of individuals (young vs. older adults) measured by a pattern separation task that were related to hippocampal connectivity to the default mode network.

Keywords: Network Psychometrics, Ergodicity, Multilayer Networks, Information Theory, Algorithm Complexity

Word count: X

Towards a psychology of individuals: The ergodicity information index and a bottom-up approach for finding generalizations

---

Introduction

In the last half of the 20th century, psychology has experienced a renewed interest in within-person (intraindividual) variation. The P-technique factor analysis set the statistical foundation for this focus (Cattell, Cattell, & Rhymer, 1947). The P-technique involves the periodic collection (e.g., hourly, daily) of responses to a battery of measures taken from a single person as opposed to across multiple people. The variation and covariation of these measures are within a person (intraindividual) as opposed to the more common between-person variation and covariation (interindividual) studied by differential psychology (Gonzales & Ferrer, 2014; Nesselroade & Molenaar, 1999). Although the P-technique did not quickly become a standard tool for exploring variability, the promise of studying intraindividual variation has attracted increasing attention (Bereiter, 1963; Gates & Molenaar, 2012; Gonzales & Ferrer, 2014; Molenaar, 2004; Nesselroade & Ford, 1985; West & Hepworth, 1991). Today, there is a considerable amount of research on intraindividual variability (Beck & Jackson, 2022; e.g., ; Diehl, Hooker, & Sliwinski, 2014; Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017; Gomes & Golino, 2015; Hultsch, Strauss, Hunter, & MacDonald, 2008; Ram, Gerstorf, Lindenberger, & Smith, 2011), which demonstrates the P-technique's impact on the field.

A primary implication on intraindividual variation is that the individual person becomes the main unit of analysis rather than the differences between people which has generally been the case for differential psychology (Molenaar, 2004). This re-emergence of intraindividual research is not surprising. The accessibility of both new technologies (e.g., smartphones) and analysis schemes (e.g., intensive longitudinal surveys) that capture intraindividual variability has enabled researchers unparalleled opportunity to study the person rather than people. Intraindividual and interindividual variability has often been cast in terms of the idiographic versus nomothetic debate (Gonzales & Ferrer, 2014; see e.g., Lamiell, 1998).[1] The recognition of these important thrusts has several consequences. There is an obvious need to formalize the definitions of both conceptions, and to do it in such a way that they can be readily distinguished at the conceptual level while also highlighting their relation to each other (Adolf, Schuurman, Borkenau, Borsboom, & Dolan, 2014; Gonzales & Ferrer, 2014; e.g., Molenaar, Huizenga, & Nesselroade, 2003; Nesselroade & Molenaar, 2010; Oertzen et al., 2020; Schmiedek, Lövdén, Oertzen, & Lindenberger, 2020; Wright & Zimmermann, 2019).

Accompanying the effort to identify their similarities, differences, and coordination, the two kinds of variation in empirical data must also be created and made operational (Molenaar, 2004). For example, to what extent can we expect to characterize the way individuals differ from each other with the same structures that characterizes how a person changes over time? This question harks back to the old debate in developmental psychology about the relative merits of cross-sectional versus longitudinal research designs and whether cross-sectional data can be relied on to furnish accurate information regarding change over time at the individual level.

More recently, the ergodic property has surfaced in this context as a condition under which one can expect within-person and between-person structures to match (Fisher, Medaglia, & Jeronimus, 2018). If a system's process is ergodic, between-person structures can be used to represent within-person structures (Gonzales & Ferrer, 2014; Molenaar et al., 2003) but, among developmental psychologists at least, there seems to be little reason to think that ergodicity will be a property of many (if any) developmental processes (e.g.,

[1] Readers interested in recent overview of the debate between intraindividual and interindividual variability can benefit from the recent overview presented by Hunter, Fisher, and Geier (2023), and the discussion by Richters (2021), Oertzen, Schmiedek, and Voelkle (2020) and Hamaker (2022) in psychological sciences, Peters (2019) and Domowitz and El-Gamal (2001) in economics, Medaglia, Ramanathan, Venkatesan, and Hillary (2011) in neurosciences, and Janczura and Weron (2015) in chemistry.

Molenaar, 2004). To date, there are few quantitative methods in psychology that can be used to evaluate whether a between-person structure is likely to hold for individual people—that is, whether the system possesses the ergodic property (Oertzen et al., 2020).

In other fields, such as chemical physics, the dynamic sublinear or superlinear increase in mean square displacement has been used as a means to distinguish ergodic from nonergodic processes in both large-scale (consisting of 1,000 time series with 1,000 time points) and small-scale (involving 10 time series with 1,000 time points) time series data, yielding promising results (Janczura & Weron, 2015). Econometrics has a rich history concerned with ergodicity (Peters, 2019) with Domowitz and El-Gamal (2001) developing algorithms that are tailored for testing the ergodicity of empirical time series data with 250, 500, or 1,000 time points. Biophysics has also analyzed ergodic properties of α-stable autoregressive fractionally integrated moving average (ARFIMA) processes and introduced a diagnostic tool for rapid evaluation, with practical applications in experimental data analysis, in univariate time series with 2,000 time points (Loch, Janczura, & Weron, 2016).

Among these applications, it's important to note that while these examples are highly relevant within their respective domains, they do not fully address the unique measurement challenges often encountered in psychology. Psychological research frequently involves data collection from items within scales, tests, or questionnaires designed based on specific factor-analytic theoretical models. These items are typically crafted to identify latent factors underlying observed responses. Therefore, it's imperative to consider the dimensionality structure of the variables, not just because of the nature of the instruments used for data collection in psychology but also due to how the scores from these instruments are intended to be employed, typically involving the summation of item responses. The work of Janczura and Weron (2015), Domowitz and El-Gamal (2001), and Loch et al. (2016) offers valuable tools for assessing ergodicity in time series data, but they are primarily designed for univariate time series and do not directly address the challenges posed by multivariate time series data with a dimensional structure similar to that commonly found in psychological research. Further, psychological studies typically involve the collection of intensive longitudinal data with much fewer time points that rarely exceeds 100. In contrast, fields like neuroscience often deal with fMRI and EEG time series data, which commonly comprised of hundreds and thousands of time points. This disparity in data characteristics emphasizes the need for tailored methodologies that can accommodate the intricacies of psychological research and its specific data structures.

Given the need to understand ergodic processes specific to psychological research (Fisher et al., 2018; Molenaar, 2004), the present research offers an information theoretic approach to evaluate the extent to which a system possesses ergodicity. From an information theoretic perspective, ergodicity can be framed in terms of the amount of information lost representing a set of measures as a single between-person structure (nomothetic structure) instead of as multiple within-person structures (within-person and idiographic structures).

Techniques to analyze complex systems with dynamic interactions between variables have a long history in statistical physics (e.g., Jaynes, 1957) and psychology (Boker, 2018; Cattell, 1965; Guttman, 1953), culminating in modern approaches such as network science (Epskamp & Fried, 2018; Epskamp, Waldorp, Mõttus, & Borsboom, 2018a). In networks, variables are represented as nodes (circles) and the edges (lines) between the nodes represent associations between variables. There are many ways to characterize complex networks (Newman, 2010), from the type of edge (e.g., directed or undirected, weighted or unweighted) to the complexity of the network (e.g., algorithmic complexity; Morzy, Kajdanowicz, & Kazienko, 2017; Zenil, Kiani, & Tegnér, 2018).

Using networks, we introduce a new metric – termed *ergodicity information index* (EII) – that can inform whether a set of variables should be represented as multiple individual networks (multiplex networks) or as a single, population network (aggregate network). The EII characterizes the relative algorithm complexity of the population structure with regard to multiple individual networks taking into consideration the number of underlying dimensions (e.g., communities, latent factors). Algorithm complexity of multiplex networks can be used to determine the optimal number of layers needed to represent a multiplex network and to detect structural and dynamical similarities among their layers (Santoro & Nicosia, 2020). The representation of intraindividual structures as a multiplex network and the quantification of their information relative to a single, population network structure are two central ideas in the development of our EII.

The paper is organized as follows. The first section briefly introduces how intraindividual and interindividual structures can be estimated in a single framework using dynamic exploratory graph analysis (Golino, Christensen, Moulder, Kim, & Boker, 2022), an approach that combines generalized local linear approximation (Boker, Deboek, Edler, & Keel, 2010) and exploratory graph analysis (Golino & Epskamp, 2017; Golino et al., 2020) to estimate dynamical factors in (intensive) longitudinal measures at different levels of analysis (individual, group, and/or population). In this section, we reframe the within- and between-person problem from a network perspective in which the intraindividual structures are represented as a multiplex network (i.e., a collection of individual networks), and the interindividual structure as a single, population network.

In the second section, we develop the EII, introducing information theoretic concepts such as algorithm complexity that are necessary to define its meaning and interpretation. Next, a Monte Carlo simulation study is implemented to investigate the accuracy of the EII to differentiate between within- and between-person structures. Afterwards, we introduce an information theoretic approach to clustering, which can be used if the system is determined to be nonergodic.

Finally, one simulated and two empirical examples from personality and neuroscience will be used to demonstrate how the EII can be used to determine whether a system possesses the ergodic property, and determine ergodic clusters (or groups) when the system is not. These new techniques are a step forward in the psychology of individuals, enabling the identification of generalizable constructs using a bottom-up approach (from individuals to groups of individuals with common network characteristics).

Representing intraindividual and interindividual structures as networks

To reframe the within- vs. between-person problem using a information theoretic network approach, we first need to show how networks can be estimated in (intensive) longitudinal data in a way that can generate both multiplex networks for the individuals (i.e., multiple individual, within-person networks) and a single population or between-person network. Recently, Golino et al. (2022) introduced a technique termed *dynamic exploratory graph analysis* (**DynEGA**), combining techniques from dynamical systems (i.e., time-delay embedding and generalized local linear approximation; Boker et al., 2010) and exploratory graph analysis (Golino & Epskamp, 2017; Golino et al., 2020) – a network psychometric approach for dimensionality assessment and reduction. The *DynEGA* technique can be used to estimate dynamical communities (e.g., latent factors) in (intensive) longitudinal measures at different levels of analysis (individual, group and/or population).

Network models have been proposed in psychological research for decades (e.g.,; Boker, 2018; Cattell, 1965; Guttman, 1953). More recently, a number of developments in network modeling in psychology originated a new area termed *network psychometrics* (Epskamp, 2018) that relies mostly on the Gaussian graphical model (**GGM**; Lauritzen, 1996). Assuming multivariate normality, a GGM can be obtained by modeling the inverse of the variance-covariance matrix (and standardizing it to obtain partial correlations) in a way that non-zero elements are freely estimated (Epskamp et al., 2018a), generating a sparse model of the variance-covariance matrix (Epskamp, Rhemtulla, & Borsboom, 2017). As Epskamp et al. (2018a) note, inverting and standardizing the variance-covariance matrix won't lead to partial correlations that are exactly zero, meaning that the GGM is saturated. Regularization techniques, such as a variant of the *least absolute shrinkage and selection operator* (**LASSO**; Tibshirani, 1996) termed *graphical LASSO* (**GLASSO**; Friedman, Hastie, & Tibshirani, 2008), are generally used in network psychometrics (Epskamp & Fried, 2018; see; Epskamp et al., 2018a).

The GLASSO is a technique that is very fast to estimate both the model structure and the parameters of a sparse GGM (Epskamp, Waldorp, Mõttus, & Borsboom, 2018b). The most used network estimation approach in psychological research (termed *EBICglasso*) is implemented in the `qgraph` package (version 1.4.1; Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). The `EBICglasso` function (Epskamp et al., 2012) samples 100 logarithmically-spaced values of λ, following Foygel and Drton (2010). The ratio range of λ can be set by the user (defaults to 0.1). γ controls the severity of the model selection (defaults to 0.50). The tuning parameter (γ) is used control the preference for complex models using the extended Bayesian information criterion (EBIC; Chen & Chen, 2008). The *EBICglasso* has been shown to accurately retrieve the true network structure in simulation studies (Epskamp & Fried, 2018; Foygel & Drton, 2010; Williams & Rast, 2020; Williams, Rhemtulla, Wysocki, & Rast, 2019).

Network psychometrics on cross-sectional data has progressed into dimensionality assessment, where networks are estimated using the GGM or other network techniques (see Golino et al., 2020) and an algorithm for detecting communities in weighted networks (Walktrap; Pons & Latapy, 2005) is used to identify latent factors. Golino and Epskamp (2017) called this approach *exploratory graph analysis* (**EGA**). Simulation studies have found EGA to perform as well as the most accurate factor analytic method, parallel analysis, and produce the best large-sample properties of all the methods evaluated (Golino & Epskamp, 2017; Golino et al., 2020).

In (intensive) longitudinal data, EGA can be used, but instead of using the variance-covariance matrix of the raw data, it uses the variance-covariance of m-order derivatives. The resulting network structure (GGM of the m-order derivatives) conveys information on how variables are *changing* over time. As a result, the communities identified using the Walktrap algorithm reflect not simple, static factors, but dynamical factors of nodes that are fluctuating similarly as a function of time. Golino et al. (2022) called this technique the *dynamic exploratory graph analysis* (**DynEGA**), merging network psychometrics, dynamical systems modeling, and dimensionality assessment into a single framework that can estimate structures at the individual, group, and population levels.

DynEGA starts by transforming the time series of each variable $V = \{v_1, v_2, ..., v_N\}$ into a time delay embedding matrix $\mathbf{X}^{(n)}$, where n is the number of embedding dimensions. As pointed by Golino et al. (2022), a time delay embedding matrix is used to reconstruct the attractor of a dynamical system using a single sequence of observations (Takens, 1981; Whitney, 1936), preserving the phase-space dynamics of the system. In the time delay embedding matrix, each row is a phase-space vector (Rosenstein, Collins, & De Luca, 1993):

$$X = [X_1 \ X_2 \ ... \ X_M]', \tag{1}$$

where X_i is the state of the system at discrete time i and is given by:

$$X_i = \left[x_i \ x_{i+\tau} \ ... \ x_{i+(n-1)\tau}\right], \tag{2}$$

where τ is the number of observations to offset successive embeddings (i.e., lag or reconstruction delay) and n is the embedding dimension. The time-delay embedding matrix is a $M \times n$ matrix, where $M = N - (n-1)\tau$ and N is the number of time points.

Suppose that U_1 represents a column in a given dataset, depicting the time series of variable V1 from $t = 1$ to $t = 10$, with $U_1 = \{5, 6, 7, ..., 14\}$. It's important to note that the values in this example significantly exceed typical psychology data. Nonetheless, this example serves the purpose of illustrating the mechanics of time delay embedding. The transformation of the time series U_1 into a time delay embedding matrix with five embedding dimensions and $\tau = 1$ results in the following matrix:

$$\mathbf{X}^{(5)} = \begin{bmatrix} 5 & 6 & 7 & 8 & 9 \\ 6 & 7 & 8 & 9 & 10 \\ 7 & 8 & 9 & 10 & 11 \\ 8 & 9 & 10 & 11 & 12 \\ 9 & 10 & 11 & 12 & 13 \\ 10 & 11 & 12 & 13 & 14 \end{bmatrix} \tag{3}$$

Once all the time series (columns) in the dataset have undergone transformation into a time-delay embedding matrix $\mathbf{X}^{(n)}$, it becomes possible to estimate derivatives using the Generalized Local Linear Approximation (Boker et al., 2010; Deboeck, Montpetit, Bergeman, & Boker, 2009).

Once the time series of each variable collected in an (intensive) longitudinal study and transformed into a time-delay embedding matrix $\mathbf{X}^{(n)}$, derivatives can be estimated using generalized local linear approximation (Boker et al., 2010; **GLLA**; Deboeck et al., 2009). Derivatives can represent different aspects of change such as the rate of change or velocity at which the variable is changing over time (first-order derivatives) and the speed of the rate of change or acceleration (second-order derivatives).

Deboeck et al. (2009) and Boker et al. (2010) show how derivatives can be estimated in the GLLA framework:

$$\mathbf{Y} = \mathbf{XL}(\mathbf{L'L})^{\mathbf{-1}}, \tag{4}$$

where $\mathbf{Y}$ is a matrix of derivative estimates, $\mathbf{X}$ is a time delay embedding matrix (with n embedding dimensions; to simplify the notation, $\mathbf{X} = \mathbf{X}^{(n)}$), and $\mathbf{L}$ is a matrix with the weights expressing the relationship between the embedding matrix and the derivative estimates. The weight matrix $\mathbf{L}$ is a $n \times \alpha$ matrix, where n is the number of embedding dimensions and α is the (maximum) order of the derivative. Each column of the weight matrix is estimated as follows, considering the order of the derivatives going from zero to k, $\alpha = [0, 1, ..., k]$:

$$\mathbf{L}_\alpha = \frac{[\Delta_t(v - \bar{v})]^\alpha}{\alpha!} \tag{5}$$

where Δ_t is the time between successive observations in the time series, v is a vector from one to the number of embedded dimensions (i.e., $v = [1, 2, ...n]$), $\bar{v}$ is the mean of v, α is the order of the derivative of interest, and $\alpha!$ is the factorial of α.

After estimating the derivatives for all time series (i.e., all variables), EGA is used to estimate the multiplex networks (intraindividual or within-person structures). In this process, each matrix of derivatives will generate a different network and dimensionality structure (represented by clusters of nodes in the network) for each individual. The population (between-person or nomothetic) structure can be estimated by stacking the derivative matrix of each individual (i.e., row-binding the matrices) and applying EGA to the stacked matrix. In both cases (i.e., individual or population structures), the resulting clusters in the networks corresponds to variables that are changing together (Golino et al., 2022).

An example illustrates how this technique works: Suppose we ask three people to answer eight items of depression once a day for 100 days. After applying DynEGA to the data, we would obtain three important types of information (see Figure 1): the derivatives for the eight variables, the network structure for each person (intraindividual), and the network structure of the three people combined (interindividual). The left side of Figure 1 shows the time series of first-order derivatives while the network structure of each individual are at the center (top and bottom represents subject one and two, respectively). The interindividual network structure is depicted in the right side of Figure 1. Note that we are omitting the community membership information to make the plot easier to visualize (each node has the same color as their respective time series of first-order derivatives). As can be seen in the network structure of person one (top network at the center of Figure 1), the pair of nodes pink and green are not connected (i.e. there is no edge linking these pairs of nodes). However, the same pair of nodes are connected in the network structure of person two (middle network at the center of Figure 1) and person three (bottom network at the center of Figure 1), and in the between-individual network structure (or population network) depicted at the right side of Figure 1. Additionally, the network structure of persons one and two clearly shows that the eight depression items form two groups or communities, while the network structure of person three forms four groups of two nodes each. A question that must be answered now is: how much information is being lost by representing the structure of the three individuals as a single, interindividual network, relative to representing them as a multiplex network? Can the derivatives of person one, two, and three be stacked to estimate a single network or by doing so, is important information about each individual lost? The next section aims to answer these questions and to introduce the *ergodicity information index*.

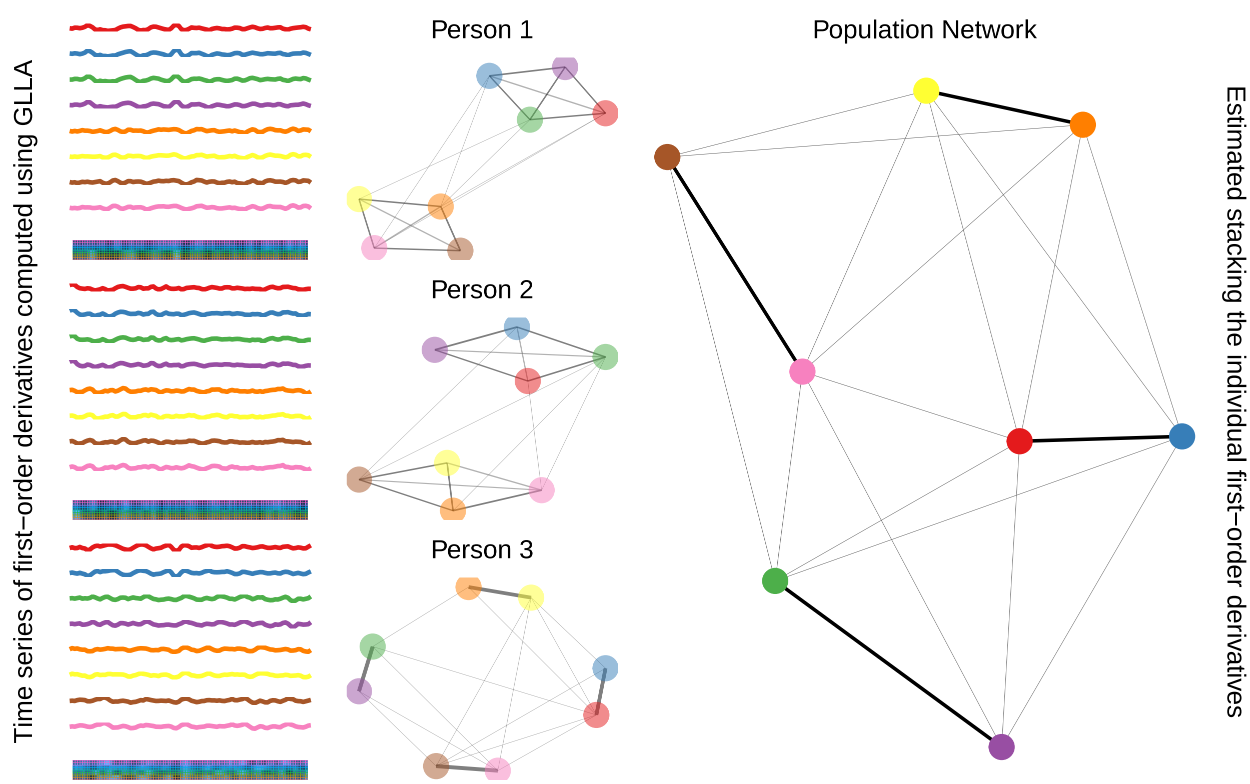


Figure 1. Network structures estimated using Dynamic Exploratory Graph Analysis.

The Ergodicity Information Index

The individual networks and population network in Figure 1 can be compared in terms of their algorithm complexity. Algorithm complexity can be used to analyze complex objects in an unbiased manner using mathematical principles (Zenil et al., 2018), and is based on the work of Kolmogorov (1968), Martin-Löf (1966), Solomonoff (1964), and others. As Zenil et al. (2018) and Morzy et al. (2017) show, the algorithm (or Kolmogorov) complexity of a string s is formally defined as:

$$K_s = min(|P|\,, T(P) = s)$$

where P is a program producing the string s when running on a universal Turing machine T, and $|P|$ is the number of bits required to represent P (i.e., the length of P). A Turing machine is a formal model of a general-purpose computer that can be programmed to reproduce any computable object, such as a string (Zenil et al., 2018). The Kolmogorov complexity of a string is defined, in other words, as the length of the shortest possible program that can produce that string as its output (Santoro & Nicosia, 2020).

Figure 2 helps to illustrate the idea of Kolmogorov (or algorithm) complexity. The two networks with six nodes have different Kolmogorov complexity because the program needed to described them differs in length. The "program" defining *network one* has the following code:

1. *red* is connected to *blue*, *green* and *purple*
2. *blue* is connected to *red* and *green*
3. *green* is connected to *red* and *blue*
4. *purple* is connected to *red*, *orange*, and *yellow*
5. *orange* is connected to *purple* and *yellow*
6. *yellow* is connected to *orange* and *purple*

While the "program" defining *network two* has the following code:

1. All pairs of nodes are connected

Network one has a higher Kolmogorov complexity because the program needed to define it is long, while network two has a lower Kolmogorov complexity because the program needed to define it is short. In sum, in information theory complexity is similar to description length. The longer the description length, the higher the complexity. This concept is also related to randomness and structure, chaos and regularity (Downey & Hirschfeldt, 2010; Shen, Uspensky, & Vereshchagin, 2022; Velupillai, 2011).

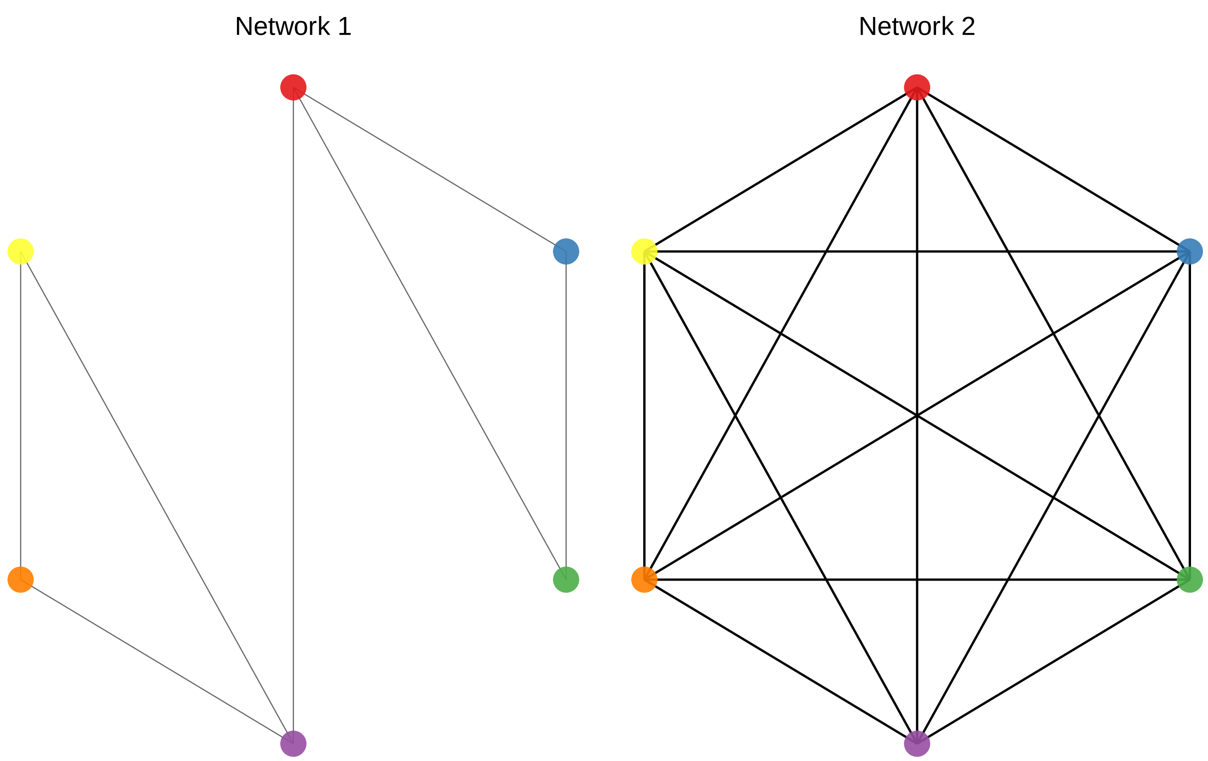


Figure 2. Two networks with six nodes and different algorithm complexity.

Kolmogorov complexity has a drawback of being intractable (Zenil et al., 2018) since there is no way to estimate the number of possible programs that could produce the string s (Kolmogorov, 1968). But it can be approximated by using compression algorithms (Morzy et al., 2017; Santoro & Nicosia, 2020) in which the compressed string s is an estimate of K_s.

To obtain an estimation of Kolmogorov complexity in networks, the most common approach is to compute the size of the compressed weighted edge list (Santoro & Nicosia, 2020). For single, unique networks, this is straightforward. Taking Figure 2 as an example, we start by getting the weighted edge list of network one, and then we transform it to a string vector (row-wise). Then, a compression algorithm is used, and the length of the compressed string vector of the weighted edge list is an approximation to Kolmogorov complexity. A more robust estimation of Kolmogorov complexity (Santoro & Nicosia, 2020) involves shuffling the weighted edge list several times and computing the mean length of the compressed string vector. The string vector of the weighted edge list of network one (Figure 2) is:

2, 2, 1, 1, 6, 0, 1, 5, 0, 1, 3, 0, 4, 5, 0, 6, 5, 0.5, 2, 6, 0, 5, 5, 1, 6, 5, 0.5, 5, 1, 0, 3, 2, 0.5, 3, 1, 0.5, 1, 6, 0, 2, 6, 0, 4, 3, 0, 2, 4, 0, 4, 1, 0.5, 2, 5, 0, 5, 5, 1, 6, 6, 1, 3, 4, 0, 3, 3, 1, 2, 1, 0.5, 6, 1, 0, 2, 2, 1, 3, 2, 0.5, 3, 2, 0.5, 2, 3, 0, 1, 2, 0, 3, 5, 0, 5, 4, 0.5, 4, 5, 0, 2, 6, 0, 2, 4, 0, 6, 2, 0, 1, 5, 0.

The compressed vector string is:

78 9c 65 90 db 0d 40 21 08 43 57 71 00 63 b4 02 fb 8f 76 93 62 7d e4 26 7c 34 15 38 15 d4 82 5a 06 2b 6a e9 14 2e 31 29 4c 4e a4 68 ce a1 6c 77 d6 78 1e d3 e8 9c c7 f2 66 7a cd 2f d4 5e 62 42 81 3a 9d 71 50 fe 43 85 f2 99 40 09 c0 99 0b a5 80 ec 37 0e ce fe a9 0f 43 db 36 d1 56 d7 3e 02 ae f4 a6 b3 e0 3e dd 07 a8 f7 38 7c

and has a length of 103 bits.

Shuffling the weighted edge list of network one and network two (Figure 2) 1,000 times, and compressing the shuffled weighted edge list for each network generates a distribution seen on Figure 3. The mean length of the compressed weighted edge list of the networks is their estimated Kolmogorov complexity.

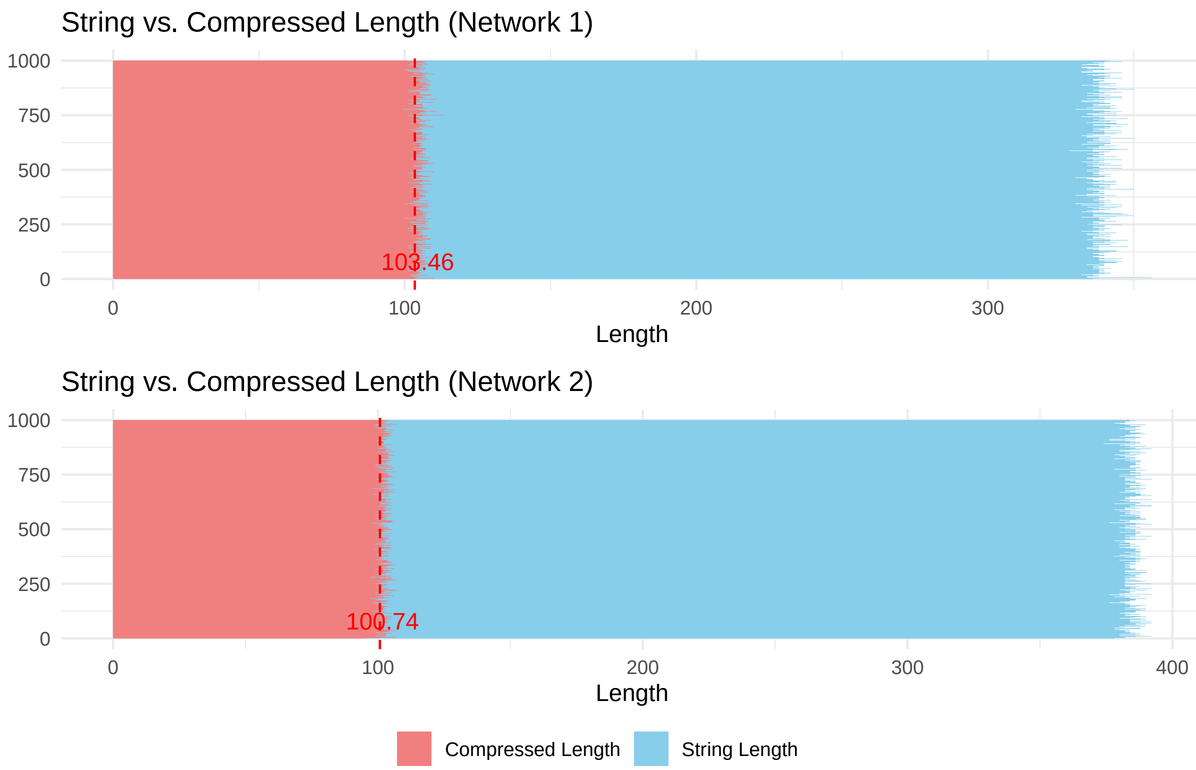
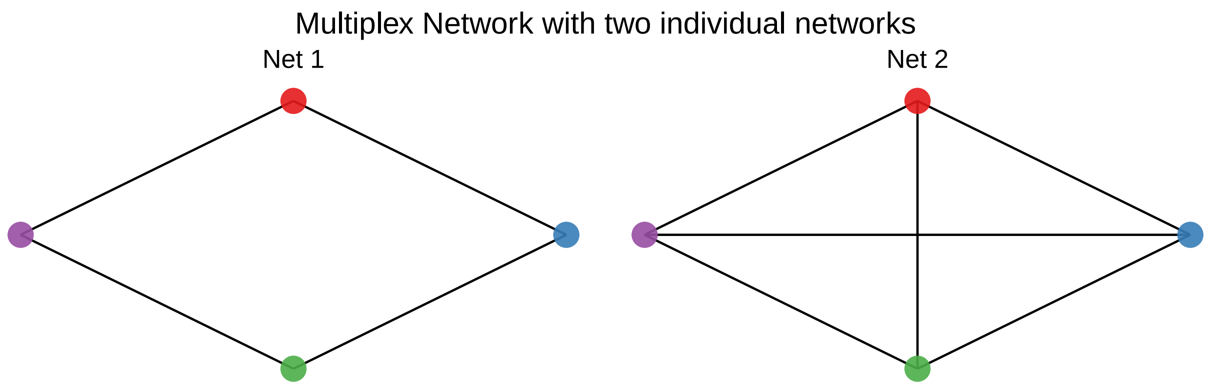


Figure 3. Distribution of the length of the weighted edge list and their compressed lengths for one thousand random shuffles of the edge list for networks one and two.

For multiplex networks, the Kolmogorov complexity requires a strategy to encode all individual graphs into a single network. Santoro and Nicosia (2020) proposed the use of a prime-weight encoding matrix Ω that assigns a distinct prime number ($p^{[\alpha]}$) to each individual network (i.e., each of the A layers of the multiplex networks) and sets each element Ω_{ij} equal to the product of the primes associated to the layers where an edge between node i and j exists:

$$\Omega_{ij} = \begin{cases} \prod_{\alpha:\alpha_{ij}^{[\alpha]}=1} p^{[\alpha]} \\ 0 \text{ if } \alpha_{ij}^{[\alpha]} = 0 \ \forall \ \alpha = 1, \ldots, A \end{cases} \tag{6}$$

For the multiplex network depicted in Figure 4, the two layers has the following unique edges:



	N1	N2	N3	N4
N1	1	0.5	0	0.5
N2	0.5	1	0.5	0
N3	0	0.5	1	0.5
N4	0.5	0	0.5	1

	N1	N2	N3	N4
N1	1	0.5	0.5	0.5
N2	0.5	1	0.5	0.5
N3	0.5	0.5	1	0.5
N4	0.5	0.5	0.5	1

Figure 4. A Multiplex Network with two layers: net 1 and net 2.

$$\begin{array}{c c} & \begin{array}{cccc} N1 & N2 & N3 & N4 \end{array} \\ \begin{array}{c} N1 \\ N2 \\ N3 \\ N4 \end{array} & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \tag{7}$$

These unique edges represent the *adjacency* matrix of the multiplex network, in which an edge between node i and j has a value of one if this edge is present in at least one layer (i.e., network) of the multiplex network. Therefore, all pairs of variables with a value of one in this adjacency matrix are used in the next steps of the computation of the Kolmogorov complexity.

Since there are only two layers in the multiplex network depicted in Figure 4, only two prime numbers are used: 2 and 3. In canonical prime association, prime numbers are associated to layers in increasing order of their total number of edges. Because network one in Figure 4 has less edges than network two, the first prime number (2) is associated to this network, and network two is associated to the second prime number (3). Considering the pair of variables in the matrix of the common edges from the multiplex network (the adjacency matrix showed above), the prime-weight transformation for network one (Figure 4) results in the following prime-weight matrix:

$$\mathbf{PW1} = \begin{bmatrix} 2^1 & 2^1 & 2^0 & 2^1 \\ 2^1 & 2^1 & 2^1 & 2^0 \\ 2^0 & 2^1 & 2^1 & 2^1 \\ 2^1 & 2^0 & 2^1 & 2^1 \end{bmatrix} \tag{8}$$

And for network two, the prime-weight transformation results in the following matrix:

$$\mathbf{PW2} = \begin{bmatrix} 3^1 & 3^1 & 3^1 & 3^1 \\ 3^1 & 3^1 & 3^1 & 3^1 \\ 3^1 & 3^1 & 3^1 & 3^1 \\ 3^1 & 3^1 & 3^1 & 3^1 \end{bmatrix} \tag{9}$$

After applying the prime-weight transformation to each layer of the multiplex network (using the pair of items in the adjacency matrix of the common edges), Ω_{ij} from equation 6 is computed by multiplying all prime-weight transformed edges i, j in $PW1$ (8) and $PW2$ (9):

$$\mathbf{\Omega} = \begin{bmatrix} 6 & 6 & 3 & 6 \\ 6 & 6 & 6 & 3 \\ 3 & 6 & 6 & 6 \\ 6 & 3 & 6 & 6 \end{bmatrix} \tag{10}$$

Matrix 10 shows the *unweighted* prime-weight encoding matrix Ω of the multiplex network depicted in Figure 4. The weighted prime-weight encoding follows the same approach, but uses the *weights* of the layers of the multiplex network instead of the adjacency matrix. The edges are still the same edges used in the example above (edges that are present in at least one layer of the multiplex network). In the current example, the weights of the layers for network one and two in Figure 4 are only 0.5 or 0. The corresponding prime-weight transformation for each network are therefore:

$$\mathbf{PW1} = \begin{bmatrix} 2^{0.5} & 2^{0.5} & 2^{0} & 2^{0.5} \\ 2^{0.5} & 2^{0.5} & 2^{0.5} & 2^{0} \\ 2^{0} & 2^{0.5} & 2^{0.5} & 2^{0.5} \\ 2^{0.5} & 2^{0} & 2^{0.5} & 2^{0.5} \end{bmatrix} \tag{11}$$

$$\mathbf{PW2} = \begin{bmatrix} 3^{0.5} & 3^{0.5} & 3^{0.5} & 3^{0.5} \\ 3^{0.5} & 3^{0.5} & 3^{0.5} & 3^{0.5} \\ 3^{0.5} & 3^{0.5} & 3^{0.5} & 3^{0.5} \\ 3^{0.5} & 3^{0.5} & 3^{0.5} & 3^{0.5} \end{bmatrix} \tag{12}$$

Resulting in the following prime-weight encoding matrix Ω:

$$\mathbf{\Omega} = \begin{bmatrix} 2.449 & 2.449 & 1.732 & 2.449 \\ 2.449 & 2.449 & 2.449 & 1.732 \\ 1.732 & 2.449 & 2.449 & 2.449 \\ 2.449 & 1.732 & 2.449 & 2.449 \end{bmatrix} \tag{13}$$

Matrix 13 shows the prime-weight encoding matrix Ω of the multiplex network depicted in Figure 4, using the weight matrix of the common edges of all layers of the multiplex network. This is the *weighted* prime-weight transformation.

How does a prime-weight encoding of a multiplex network looks like? Figure 5 shows the network representation of the weighted prime-weight encoding matrix for individual networks estimated using dynamic exploratory graph analysis and presented on Figure 1. The prime-weight encoding matrix preserves full information about the placement of all edges of the multiplex network (Santoro & Nicosia, 2020).

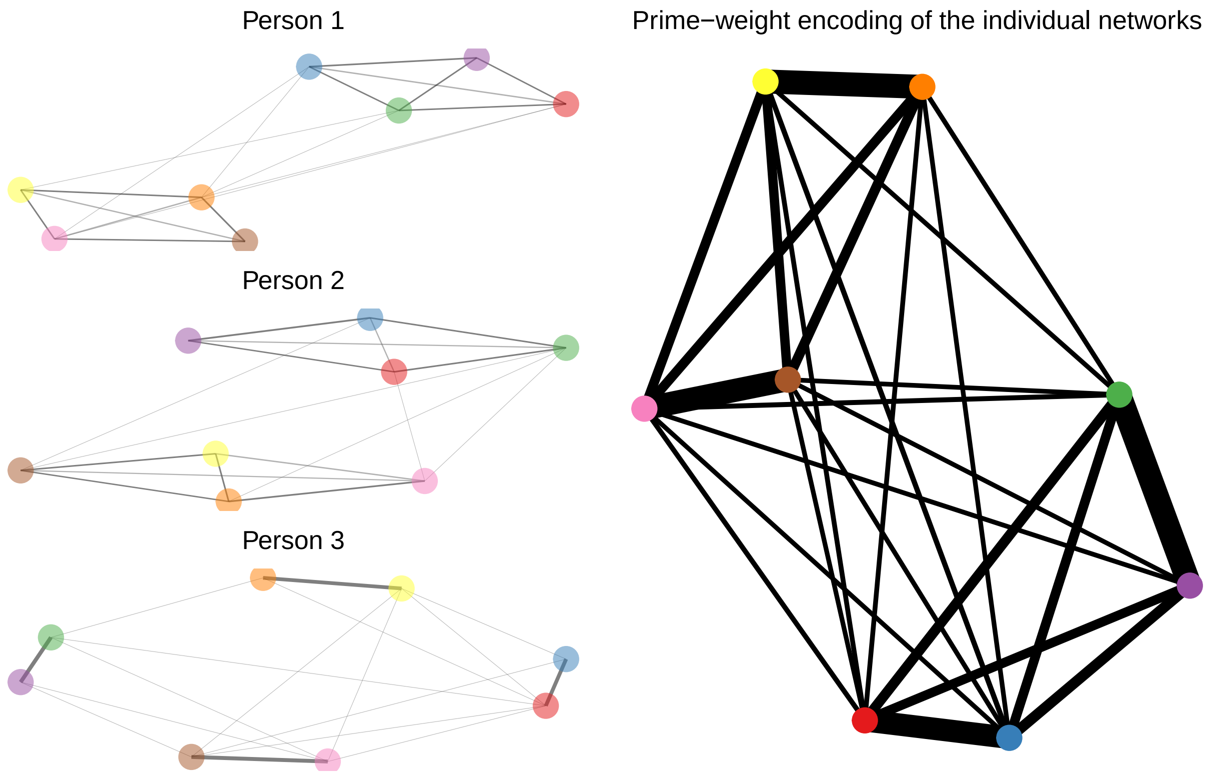


Figure 5. Network representation of the weighted prime-weight encoding matrix for individual networks estimated using dynamic exploratory graph analysis

The prime-weight encoding of a multiplex network enables the estimation of the Kolmogorov complexity for all layers (networks) simultaneously. The prime-weight encoding uses prime numbers to uniquely "tag" each network. By using the prime numbers as "tags" for each network in a increasing order (canonical prime encoding) in which each prime number is associated with the layers with the lower number of edges, the prime numbers act as unique IDs to encode which edges came from which individual networks. This enables a joint representation of all networks in the multiplex network. It also opens up the possibility of comparing these individual networks of the multiplex network with population structures, as estimated in the dynamic exploratory graph analysis technique.

Santoro and Nicosia (2020) proposed a new metric for quantifying the algorithm complexity of multiplex networks that can be computed as the ratio of the (approximate) Kolmogorov complexity of the prime-weight matrix Ω of a multiplex network with A layers and the Kolmogorov complexity of an aggregated network combining all layers:

$$\mathcal{C}(\mathcal{M}) = \left[\frac{\left(\frac{K_\Omega}{K_W}\right)}{\log(L_\chi)}\right] \tag{14}$$

In which K_Ω is the Kolmogorov complexity of the prime-weight encoding matrix of the individual networks and K_W is the Kolmogorov complexity of the single-layer weighted matrix W, obtained by considering the aggregate binary matrix multiplied by the largest entry of Ω. Lastly, $L\chi$ is the number of distinct edges in multiplex network (Santoro & Nicosia, 2020).

In the current paper, we propose a similar strategy to quantify the algorithm complexity of the networks estimated using DynEGA. The multiplex networks are all individual networks estimated using the derivatives computed via GLLA in the DynEGA technique. Instead of comparing the algorithm complexity of Ω with an weight aggregation of the multiplex networks, it is more informative to compare it with the population network (i.e., the network estimated stacking the derivatives estimated using GLLA for all individuals). Additionally, in psychology, it is also important to consider the number of latent factors underlying the intensive-longitudinal data. Therefore, our *ergodicity information index* can be computed as:

$$\xi = \sqrt{F_P + 1}^{\left[\frac{\left(\frac{K_\Omega}{K_{P*}}\right)}{\log(L_\chi)}\right]} \tag{15}$$

where $\sqrt{F_P}$ is the square-root of the number of factors estimated in the population structure using DynEGA, K_Ω is the algorithm complexity of the prime-weight encoding matrix of the individual networks (that composes the multiplex network χ), K_{P*} is the algorithm complexity of the prime-weight transformation of the population network (i.e., each element in the population network, P_{ij}, is transformed such that $P_{ij}* = 2^{P_{ij}}$),and $L\chi$ is the number of distinct edges across the networks that make up the multiplex network (i.e., non-zero edges). One is added to the number of factors estimated in the population network to deal with unidimensional population structures.

The EII (ξ) computes the amount of information lost representing a set of measures as a single interindividual structure (nomothetic structure) instead of representing the measures as multiple individual structures (within-person or intraindividual structures). Larger values of the EII indicate that the intraindividual networks encode a relatively larger amount of information with respect to the population network.

The ratio $\frac{K_\Omega}{K_{P*}}$ is normalized by the number of distinct edges in the multiplex network (set of individual networks) because a set of networks with larger number of edges is expected to have a higher complexity. Similar to Santoro and Nicosia (2020), we implemented a canonical prime association to compute ξ. In canonical prime association, prime numbers are associated to layers in increasing order of their total number of edges.

In terms of estimation, the Kolmogorov complexity is estimated for each network (i.e., prime-weight transformation of the population network $P*$ and the prime-weight encoding of the multiplex networks Ω) as the length of the compression of the string formed by their edge list, using the *Gzip* compression algorithm available in the `memCompress` function of base R (R Core Team, 2017). In theory, Kolmogorov complexity can also be computed using other representations of the network such as a string composed by the rows of the Laplacian matrix, degree list, degree distribution, or weights list (Morzy et al., 2017). In this paper, we focus on the algorithm complexity estimated in the *unweighted* and *weighted* prime-weight encodings (i.e., in the weighted edge list of the unweighted prime encoding and the weighted edge list of the weighted prime encoding, respectively), and the *edge list* of the prime encoding only (with no weights). Since Kolmogorov complexity is affected by the order of elements in the edge list, the final estimation of algorithm complexity is based on the mean of 5,000 computations of K over an edge list randomly ordered. One advantage of using Kolmogorov complexity to estimate the complexity of networks is that it is less dependent on the network representation than other metrics of complexity such as entropy-based metrics (Morzy et al., 2017).

The use of the EII implies a different type of ergodicity that we call *super-weak ergodicity*. In a strict definition of ergodicity, there are two central requirements: homogeneity of all participants and stationarity for all time points (Voelkle, Brose, Schmiedek, & Lindenberger, 2014). A softer type of ergodicity termed *weak ergodicity*, requires only that the marginal distributions for all participants and for all time points be identical (Oertzen et al., 2020). The *super-weak ergodicity*, on the other side, doesn't require stationary for all participants (i.e., the same covariance matrix for all subjects), homogeneity for all time points (i.e., the same covariance matrix across time), or an equal marginal distributions for all participants and time points. It requires a much weaker condition: the algorithm complexity of the population (or between-person) network be similar (but not equal) to the algorithm complexity of the prime-weight encoded network of all individuals.

Suppose we have four people that are assessed using an eight-item questionnaire for 100 days. Persons one and two are more similar than persons three and four, although none of them are exactly equal to one another. DynEGA is used to compute the first-order derivatives for each variable and to generate a network for each individual, and two between-person or population networks: one for persons one and two, and one for persons three and four. If we calculate the correlation of the derivatives for each variable, none of them are exactly equal to the others. Figure 6 shows the heatmap of the correlation matrices (calculated using the first-order derivatives of each variable) and the resulting network structure for each person and each population with node colors representing the estimated latent factors.

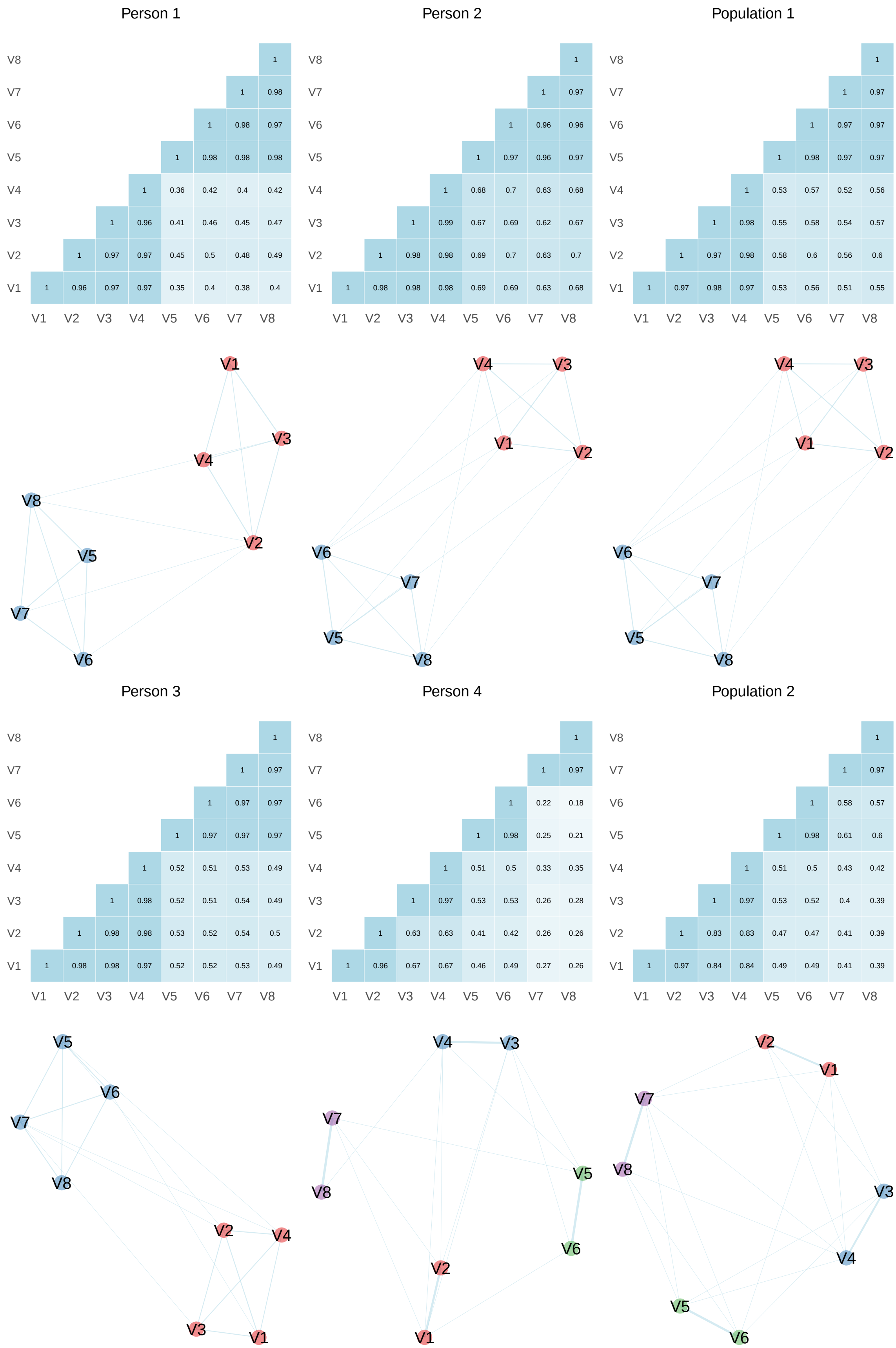


Figure 6. Heatmap of the correlation matrices and the resulting network structure for each person and each population with node colors representing the estimated latent factors.

Two factors were identified in individuals one, two and three, and four factors were identified in individual four. The population network for persons one and two indeed shows two factors, while the population network for persons three and four shows four factors. Calculating the EII (ξ), we obtain $\xi = 1.24$ for individuals one and two and $\xi = 1.35$ for individuals three and four. The ergodicity information index was computed using the Kolmogorov complexity of the weighted prime-weight encoding of the networks as discussed above.

What EII (ξ) shows is that more information is lost by representing the eight measures as a single structure for individuals three and four (bottom of Figure 6, population 2) than for individuals one and two (top of Figure 6, population 1). Another way to interpret the results above is that the intraindividual networks for individuals three and four encode a relatively larger amount of information with respect to the population network, compared to individuals one and two and their population structure.

Simulation Design

To verify the suitability of the EII to identify if intensive longitudinal measures should be represented as a set of within-person structures (multiple individual networks) or as a single, between-person or population structure (only one network), a Monte Carlo simulation is implemented. Four data conditions were systematically manipulated: sample size (50 and 100), number of variables per factor (4 and 6), number of factors (2, 3), error (0.125, 0.25, and 0.50 – each one squared in the measurement error covariance matrix), and dynamic factor loadings (0.4, 0.6, and 0.8).

Two separate set of conditions were used in the simulation. In the first, **all individuals had the same number of factors** (*Eq* condition). Therefore, representing these individuals using a single population structure is reasonable. In the second set of conditions, **half of the subjects had the same number of factors (and variables per**

factor) as in the "Eq" condition, but the other half had a different configuration, with more or less factors (*NotEq* condition). In the *NotEq* condition, when the primary group consisted of two factors, each comprising four items, the secondary (distinct) group featured four factors, each with two items. In the scenario where the main group comprised two factors, each with six items, the second group was characterized by three factors, each containing four items. When the primary group involved three factors, each consisting of four items, the secondary group exhibited four factors, each encompassing three items. In the case of the main group having three factors, each with six items, the secondary group displayed two factors, with nine items within each factor.

The sample sizes were selected to reflect moderate (N = 50) and high (N = 100) samples, consistent with many empirical papers using an intensive longitudinal measurement design that typically don't use data from more than 100 individuals (Liu, Zhou, Palumbo, & Wang, 2016; Schmiedek, Lövdén, & Lindenberger, 2020). In terms of the number of variables per factor, three are the minimum required for factor identification (Anderson, 1958). In the present simulation, the number of items were selected to reflect adequate (4) and slightly overindentified (6) factors (Velicer, 1976; Widaman, 1993). However, in one condition was not possible to have an adequate factor identification for the *not equal* data structure, since the primary group of simulated data consisted of two factors, each comprising four items, but the secondary (distinct) group featured four factors, each with two items only.

The measurement error covariance matrix had 0.125^2, 0.25^2 and 0.5^2 in the diagonal, representing low, moderate, and high errors. This strategy allows for the impact of the other variables systematically manipulated in the simulation to be better understood. The dynamic factor loadings of the *DAFS* model were selected to reflect small (0.4), moderately high (0.6), and high (0.8) loadings. Two conditions were held constant: the matrix with autoregressive (0.8) and cross-regressive coefficients (0), and the covariance matrix for the random shock (off-diagonal = 0.18; diagonal = 0.36). The values of the autoregressive and the cross-regressive coefficients and the random shock matrix were selected following Zhang, Hamaker, and Nesselroade (2008).

Together, 216 conditions are tested, with all possible combinations of sample size (N = 50 or 100), variables per factor of the main group (NVAR = 4 or 6), number of factors of the main group (NFAC = 2 or 3), variables per group of the secondary group for the "NotEq" condition (NVAR2 = 2, 3, 4, or 9 depending on NVAR and NFAC), number of factors of the secondary group for the "NotEq" condition (NFAC2 = 2, 3, or 4, depending on NVAR and NFAC), loadings (LOAD = 0.4, 0.6, or 0.8), EII method (weighted, unweighted, or edge list), and error (Standard-deviation of the measurement error matrix: 0.125, 0.25, 0.5).

Data Generation

In the Monte Carlo simulation, 500 data matrices were generated for each combination of variables (number of factors, number of variables per factor, factor loading, and sample size) according to the DAFS model. First, the matrix of random shock vectors $\mathbf{v}_t$ was generated following a multivariate normal distribution with mean zeros and $q \times q$ covariance matrix $\mathbf{D}$ (off-diagonal values = 0.18; diagonal values = 0.36), where q is the number of factors and t is the number of time points plus 1,000 (used as the burn-in estimates for the Markov chain). Second, the factor scores are calculated and the first 1,000 estimates are removed (burn-in phase). Third, the measurement error matrix is generated following a multivariate normal distribution with mean zeros and $p \times p$ covariance matrix Q, where p is the total number of variables (number of variables per factor times F). Finally, the observed variables $\mathbf{Obs}_t$ at time t ($t = 1, 2, ..., N$) are calculated using the following equation 16.

$$\mathbf{Obs}_t = \mathbf{\Lambda F}_t + \mathbf{e}_t, \tag{16}$$

where $\mathbf{\Lambda}$ is the factor loading matrix ($p \times q$), $\mathbf{F}_t$ is a $q \times 1$ vector of factors at time t, and $\mathbf{e}_t$ is a $p \times 1$ vector with measurement errors following a multivariate normal distribution with mean zeros and covariance matrix Q (Nesselroade, McArdle, Aggen, & Meyers, 2002;

Zhang et al., 2008).

The factor scores, $\mathbf{F}_t$, are calculated as follows:

$$\mathbf{F}_t = \sum_{l=1}^{L} \mathbf{B}_l \mathbf{F}_{t-l} + \mathbf{v}_t \tag{17}$$

where $\mathbf{B}_l$ is a $q \times q$ matrix of autoregressive and cross-regressive coefficients, $\mathbf{F}_{t-l}$ is a vector of factor score l occasions prior to occasion t and $\mathbf{v}_t$ is a random shock vector (or innovation vector) following a multivariate normal distribution with mean zeros and $q \times q$ covariance matrix $\mathbf{D}$ (Nesselroade et al., 2002; Zhang et al., 2008). In the DAFS model, $\mathbf{\Lambda}$, $\mathbf{B}_l$, $\mathbf{Q}$ and $\mathbf{D}$ are invariant over time.

Data following the DAFS model can be simulated using the `simDFM` function of the *EGAnet* package (version 2.0.3; Golino & Christensen, 2019).

After all individual data is simulated using the DAFS model (as described above), they are combined into a single data file (row bind). Each data file with all the individual data corresponds to a replicate i (from 1 to 500) of a condition j (from 1 to 216). The DynEGA method is used to estimate a single network structure per individual and a population structure for all individuals (replicate i, condition j). Because the data from each individual comes from the same population model (the DAFS model) with the same number of items per factor, same number of factors, and same factor loadings (as defined by the condition j), the corresponding weights for each individual network will be similar, but not the same. Therefore, for the "Eq" condition, all individuals will have the same structure in terms of number of factors and number of variables per factor, but the estimated network weights vary from individual to individual, to the random variation of the data generation process.

In the "NotEq" condition, the difference is that there are two data generation mechanisms (generating data for two distinct groups of people), differing in terms of the number of factors and items per factor, but with the total number of items constant in both groups. So, in a specific group, individuals have equal number of factors and number of variables per factor, but their network weights estimated via DynEGA varies do to the random variation of the data generation process.

Data Analysis

In our current simulation, we assess the performance of three ergodicity information index methods, each computed using distinct prime-weight encoding strategies. The *unweighted* method utilizes prime-weight encoding derived from the adjacency matrix of individual networks (i.e., layers within the multiplex network). The *weighted* method employs prime-weight encoding computed from the network weights, while the *edge list* method utilizes only the edge list of the prime-weight encoding.

Within the simulation, we generated data for two sets of individuals under the direct autoregressive factor score model. In the first set, all individuals shared an equal number of factors, known as the "Eq" condition, enabling the representation of these individuals with a single population structure. In the second set of conditions, half of the subjects matched the "Eq" condition, while the other half exhibited a different configuration with varying factors, resulting in the "NotEq" condition.

We compared the three ergodicity information index methods (*unweighted*, *weighted*, and *edge list*) based on their *hit rate* or accuracy. A score of one was assigned each time the ergodicity information index yielded a lower value for the "Eq" condition (all individuals sharing the same dimensionality structure) compared to the "NotEq" condition (two subgroups with different dimensionality structures). This score indicates the accuracy of the ergodicity information index in identifying the set of individuals for which representing them with a single population structure is reasonable, as opposed to a set where such representation leads to a loss of information. The conditions used in the simulation were compared in terms of the mean accuracy of the ergodicity information index methods (*unweighted*, *weighted*, *edge list*). The partial eta-squared effect size is also computed for each condition and their combinations, in order to identify which impacts the accuracy of the ergodicity information index the most.

We also examine the average ergodicity information index values for both the "Eq" and "NotEq" data conditions. Additionally, we assess the mean difference between the ergodicity information index values estimated in the "NotEq" data condition and those in the "Eq" condition. Positive differences indicate instances where the new index is accurate, while negative differences signal situations where the index proves to be inaccurate.

Results

In our study, we explored the performance of three ergodicity information index methods: *edge.list*, *unweighted*, and *weighted*. The accuracy, also referred to as the "hit rate," varied slightly across these methods. The highest accuracy (or hit rate) was achieved by *unweighted* (83.48%), followed by *weighted* (81.39%), and by *edge.list* (82.86%). Table 1 provides an overview of the effect sizes (partial eta-squared) for each method and their impact on accuracy. Notably, two conditions exhibited a moderate effect size (loadings and number of variables), while one condition showed a high effect size (error).

Ergodicity Information Performance Across Error Levels

The accuracy of the ergodicity information index methods displayed noteworthy trends across different error levels. For a relatively low error of 0.125, the accuracy was high (*unweighted* = 100%, *weighted* = 99.98%, *edge.list* = 99.98%).

At an error of 0.25, the accuracy decreases, but the ergodicity information index methods continued to perform well (*unweighted* = 91.29%, *weighted* = 89.16%, *edge.list* = 90.40%), but when subjected to a more substantial error of 0.50, the accuracy experienced a significant decline (*unweighted* = 59.13%, *weighted* = 55.05%, *edge.list* = 58.20%).

Table 1
Non-zero effect sizes (partial-eta-squared) for the three ergodicity information index methods per condition tested.

	Weighted EII	Unweighted EII	Edge List EII
LOAD	0.11	0.10	0.09
Error	0.32	0.29	0.29
NFAC	0.04	0.03	0.04
NFAC2	0.03	0.03	0.03
NVAR	0.08	0.08	0.08
LOAD:Error	0.06	0.06	0.05
Error:NFAC	0.02	0.02	0.02
Error:NFAC2	0.02	0.02	0.02
Error:NVAR	0.04	0.04	0.04
LOAD:Error:NFAC	0.01	0.02	0.02
LOAD:Error:NFAC2	0.02	0.02	0.02
LOAD:Error:NVAR	0.04	0.02	0.03

Influence of Loadings on Accuracy

We observed a moderate effect of the magnitude of the loadings in the accuracy of ergodicity information index. For a loading of 0.4, the mean accuracy was 69.57% for *unweighted*, 66.47% for *weighted*, and 69.44% for *edge.list*. As the loading increased to 0.6, the accuracy improved: 88.86% for *unweighted*, 87.31% for *weighted*, and 87.97% for *edge.list*. For the highest loading magnitude of 0.8, the accuracy continued to rise: 92% for *unweighted*, 90.41% for *weighted*, and 91.17% for *edge.list*.

Influence of Number of Variables on Accuracy

We observed a moderate effect of the number of variables in the reference group in the accuracy of ergodicity information index. For four variables per factor, the mean accuracy was 76.79% for *unweighted*, 74% for *weighted*, and 75.83% for *edge.list*. With six variables per factor, the accuracy increased to: 90.16% for *unweighted*, 88.79% for *weighted*, and 89.89% for *edge.list*.

These findings shed light on the performance of ergodicity information index methods under varying conditions, including error levels, loadings, and number of variables per factor on accuracy.

General Trends

We will primarily focus on the unweighted ergodicity information index (EII) when discussing the general trends. This choice is driven by two key factors. First, the *unweighted* EII demonstrated the highest accuracy across all tested conditions, albeit only marginally surpassing the *weighted* and *edge list* methods. Secondly, generating figures for all three methods would be redundant, as they exhibit a consistent pattern across conditions.

In Figure 7, the mean value of the ergodicity information index (unweighted) for each condition is illustrated. Two main trends become evident. Firstly, the mean EII value for the "Eq" data condition (where all individuals share the same dimensionality structure) is consistently lower than that computed in the "NotEq" condition (characterized by two subgroups with different dimensionality structures). This pattern signifies that the EII correctly identifies the set of individuals for whom representing them with a single population structure is reasonable, resulting in lower information loss. Conversely, in scenarios where such representation leads to a significant loss of information, the EII values are higher.

The EII fails to exhibit this pattern in specific conditions, including those with high error (0.5) and low loadings (0.4), as well as conditions with moderate error (0.25) and low loadings (0.4), and those with high error (0.5) and low, moderate, and high loadings in conditions featuring two factors and four items per factor in the reference group, as well as four factors and two items in the secondary group.

These patterns are also evident in Figure 8 that shows the mean accuracy of the ergodicity information index (unweighted) per condition (left) and the mean EII difference between the "NotEq" and the "Eq" data conditions (right). In the left side of Figure 8, in

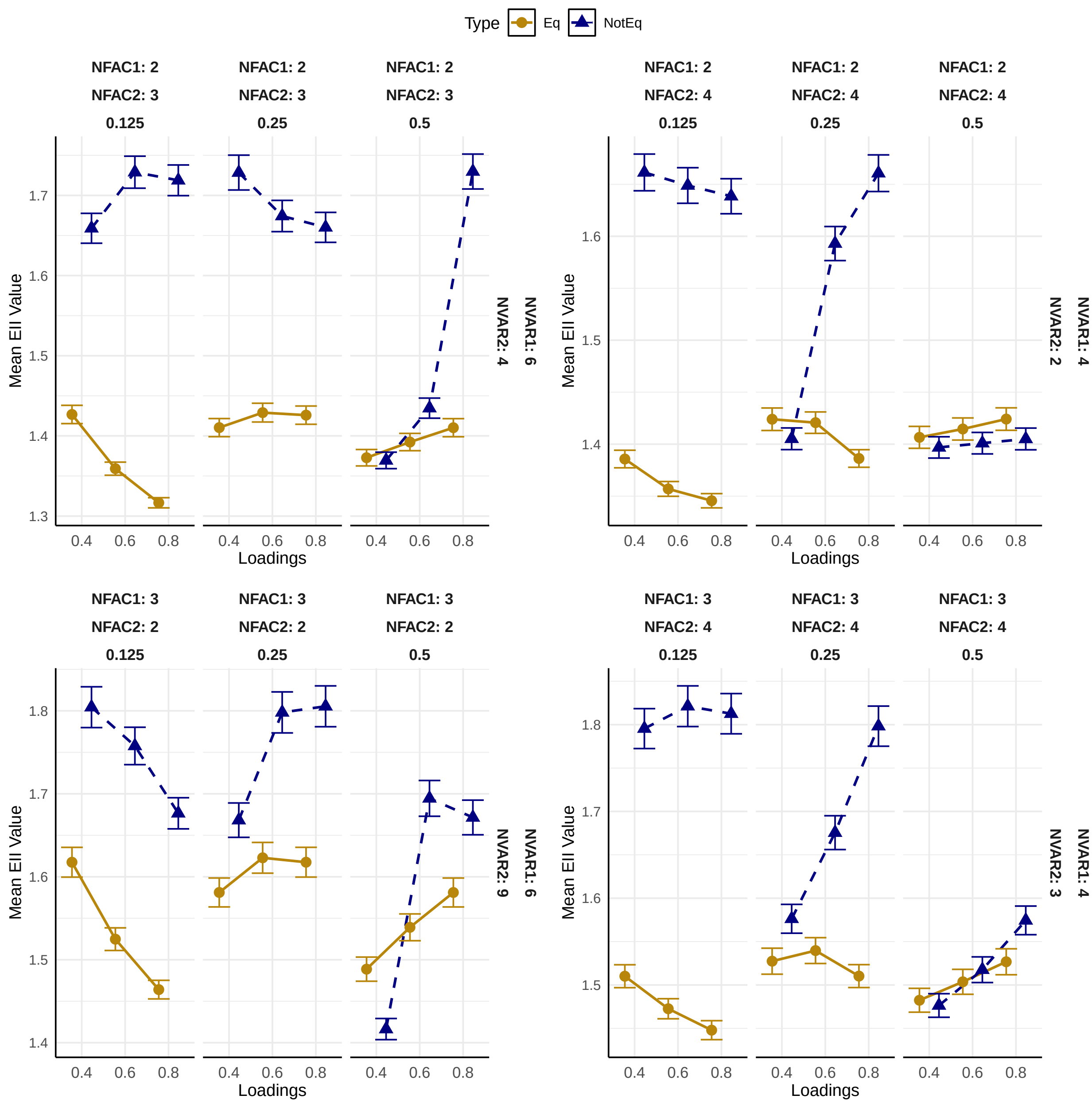


Figure 7. Mean value of the ergodicity information index (unweighted) per condition.

conditions featuring two factors and four items per factor in the reference group, as well as four factors and two items in the secondary group, the accuracy or hit rate is nearly perfect for conditions with small error, increases from small loadings to high loadings in conditions with moderate error, but is very low for conditions with high error. In the other conditions used in the simulation, the accuracy of EII is nearly perfect (100%) for small and moderate errors, and increases with the increase in the factor loadings for conditions with a higher error.

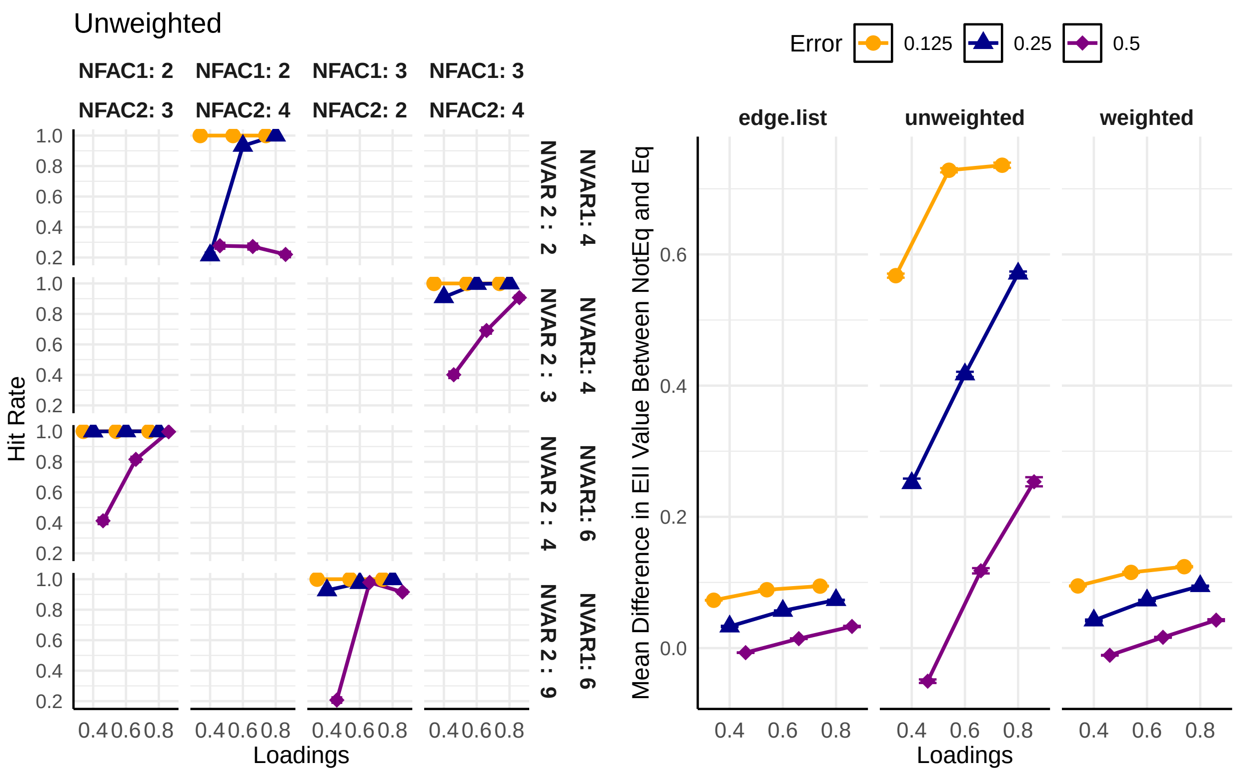


Figure 8. Mean accuracy of the ergodicity information index (unweighted) per condition (left) and the mean EII difference between the NotEq and the Eq data conditions (right).

The right side of Figure 8 illustrates the mean difference in EII between the "NotEq" and "Eq" data conditions. Positive differences indicate cases where the new index is accurate, while negative differences suggest situations where the index proves to be inaccurate. Across all conditions tested, the mean difference is negative only for small factor loadings (0.4). The difference is most pronounced in conditions with minimal error, gradually diminishing as the error rate increases.

An intriguing observation is that the *unweighted* EII exhibits the highest overall mean difference. This suggests that not only does this method slightly outperform the *weighted* and *edge list* methods in terms of accuracy, but it also demonstrates the greatest differentiation between conditions. Consequently, it should be the preferred method for applied research.

A Test for Ergodicity

The EII provides a relative metric for the information lost when representing the sample as a between-person, population structure relative to within-person, individual structures. Determining whether the amount of information lost is substantial requires understanding the information loss relative to when within-person structures are similar to the between-person structure.

Starting with an ergodic process, the expectation is that all individuals will have a systematic relationship with the population. Destroying this relationship should result in significant loss of information between the population and individual structures. With this premise, the shared edges between the population network and each individual's network should be meaningfully, and not randomly, related. In information theory terms, the information contained in the population network should be related to an individual's network beyond what can be expected from a random process.

Our EII test is based on this conjecture. The test shuffles a random subset of edges that exist in the population that is equal to the number of shared edges it has with an individual. This shuffling allows for an equivalent amount of information that is scrambled yet still shared between the population and replicate individual's network. The edges that are unique to the individual's network (i.e., present in the individual's network but not the population network) are then added to the replicate individual network. This addition of the individual's unique edges ensures that the unique information of the individual remains constant. The result is a replicate individual that contains the same total number of edges as the actual individual but its shared information with the population has been scrambled. This procedure is repeated for each individual in the sample and EII is computed.

To create a sampling distribution, this process is repeated for X number of times (e.g., 1000). This sampling distribution represents the expected information between the population and individuals when a random process underlies their shared relationships. The empirical EII value is then compared against this distribution. If the underlying process is ergodic, then this procedure is expected to destroy the relationship between the population and individuals, resulting in a sampling distribution that represents information loss due to representing the system as a random process. Said differently, the shared information between the population and the actual individual networks are sufficiently meaningful and this meaningfulness gets destroyed when the information is scrambled. In contrast, an empirical EII that is non-significant or significantly greater suggests that the information shared between the population and individuals is no different than random—that is, the pattern of relationships in the individuals are no different than can be expected when the shared information between the population and each individual is random. In other words, the individuals cannot be adequately represented with the same dynamics as the population.

A Bottom-up Approach to Find Generalizations

If the system is nonergodic, what then? It's plausible that ergodicity may exist at some level of the system such as sub-groups. Identifying sub-groups could reveal ergodic systems that exist in the overall system. Heterogeneity in sample composition limits the ability our ability to establish generalizability (Gates & Molenaar, 2012; Richters, 2021). Psychological processes may manifest themselves differently from individual-to-individual, which in turn affects the extent to which a single, population-level model can generalize to all individuals (Molenaar & Nesselroade, 2012). If you're trying to determine the average number of bedrooms in single family dwellings in the U.S., then a truly representative sampling of single family dwellings is highly desirable. But if you're trying to determine the nature of the onset and progression of depression, a representative sample that includes a variety of paths of onset and progression is not helpful and may aggregate over paths to a point where the average representation does not generalize any single person in the sample.

A "bottom-up" approach to generalizability has been argued to be "more appropriate to…emphasize first understanding individuals well and then identifying similarities across persons, thus accruing generalizability gradually than initially fitting models to heterogeneous samples in order to claim generalizability" (p. 15, Nesselroade & Molenaar, 2016). Following this line of reasoning, we propose an approach that aims to quantify the pairwise similarity of individual network structures to search for sub-groups or generalizable characteristics using an information theoretic metric.

The Von Neumann entropy of the spectral properties of a network can provide insights into both the topological features (i.e., connectivity between nodes) but also the community structure (Chauhan, Girvan, & Ott, 2009) and temporal dynamics (Almendral & Dıaz-Guilera, 2007) of a network. Recent work has taken advantage of these properties to determine whether individual networks in a multiplex network can be aggregated into groups (De Domenico, Nicosia, Arenas, & Latora, 2015). De Domenico et al. (2015) proposed a multiplex network reduction approach by computing Von Neumman entropy of two networks and computing their Jensen-Shannon Distance (JSD). A similar approach has been applied in the Bayesian context of comparing networks in psychology (Williams, Rast, Pericchi, & Mulder, 2020). Von Neumann entropy of a network can be computed as follows (De Domenico et al., 2015):

$$h_A = -\mathrm{Tr}[\mathcal{L}_G \log_2 \mathcal{L}_G],$$

where $\mathcal{L}_G = c \times (D - A)$ is the combinatorial Laplacian rescaled by c or one over the sum of the weights in the network. D is a matrix with the strength of each node (i.e., sum of each node's connections) on its respective diagonal and A is the network. $\mathcal{L}_G$ is a density matrix that is then used to compute Von Neumman entropy:

$$h_A = -\sum_{i=1}^{N} \lambda_i \log_2(\lambda_i),$$

where λ are the eigenvalues of $\mathcal{L}_G$. Using Von Neumann entropy of the network, we can compute the Jensen-Shannon Divergence between two networks, which is the symmetric version of the Kullback-Leibler Divergence (De Domenico et al., 2015):

$$\mathcal{D}_{JS}(\rho||\sigma) = h(\mu) - \frac{1}{2}[h(\rho) + h(\sigma)],$$

where ρ and σ are $\mathcal{L}_G$ of each network being compared and $\mu = \frac{1}{2}(\rho + \sigma)$. Taking the square root of $\mathcal{D}_{JS}$ produces a $[0, 1]$ bound metric often referred to as JSD.

De Domenico et al. (2015) used an entropy-based quality function and Ward's method for agglomerative hierarchical clustering using the JSD to determine whether individual layers (i.e., networks) in a multiplex network can be aggregated. Our approach follows allow similar lines: We use Ward's agglomerative hierarchical clustering (Ward, 1963) on the pairwise JSD values between all individuals' networks in the sample. The trees are then cut through all possible cuts (2 through N), obtaining $N - 1$ sets of possible clusters. Afterward, a similarity matrix (1 - JSD) is obtained and used to compute modularity for each set of clusters. Modularity is commonly used as a criterion in many community detection algorithms in network science (Newman, 2006). The modularity metric quantifies the extent to which within-cluster similarity is maximized and between-cluster similarity is minimized. The difference between each consecutive cluster's modularity is computed and the cluster solution with the highest modularity is selected.

Simulated and Empirical Examples

To demonstrate the statistical test for EII and the information theory clustering, we applied them to a simulated example and two empirical examples. In the simulated example, three groups were generated using the DAFS method above. The first group had 2 factors with 12 variables per factor; the second group had 3 factors with 8 variables per factor; the third group had 4 factors with 6 variables per factor. All groups had 25 individuals with 50 time points each, loadings were randomly drawn from a uniform distribution between 0.50 and 0.70, autoregressive parameter set to 0.80, cross-regressive parameter set to 0.00, variance shock set to 0.36, covariance shock set to 0.18, and error set to 0.375.

These groups were first combined and the bootstrap EII test was applied. This test determined that the data were nonergodic, $EII = 2.824, p = 0.020, M_{EII} = 2.815$ ($SD = 0.004$) (Figure 9). After determining the full sample was nonergodic, the dynamic EGA results were passed to the information theory clustering approach. This approach correctly identified the three groups that made up the sample (Figure 9).

Once separated into three groups identified by the clustering approach, the EII test was applied again. For all groups, the empirical EII was significantly less than their respective sampling distributions (Group 1: $EII = 1.729, p = 0.002, M_{EII} = 1.760$ ($SD = 0.003$); Group 2: $EII = 2.01, p = 0.002, M_{EII} = 2.125$ ($SD = 0.005$); Group 3: $EII = 2.322, p = 0.002, M_{EII} = 2.524$ ($SD = 0.007$)), suggesting that they were ergodic.

Next, two empirical data examples that are commonly represented with aggregate structures, personality and brain networks, were examined for whether they were ergodic and if not, whether there were any clusters. For personality, we used an empirical example taken from an intensive longitudinal experience sampling study examining Big Five personality measured by the Big Five Inventory-2 (Beck & Jackson, 2022; Soto & John, 2017). There were 199 participants who completed between 1 and 158 time points. To ensure optimal data

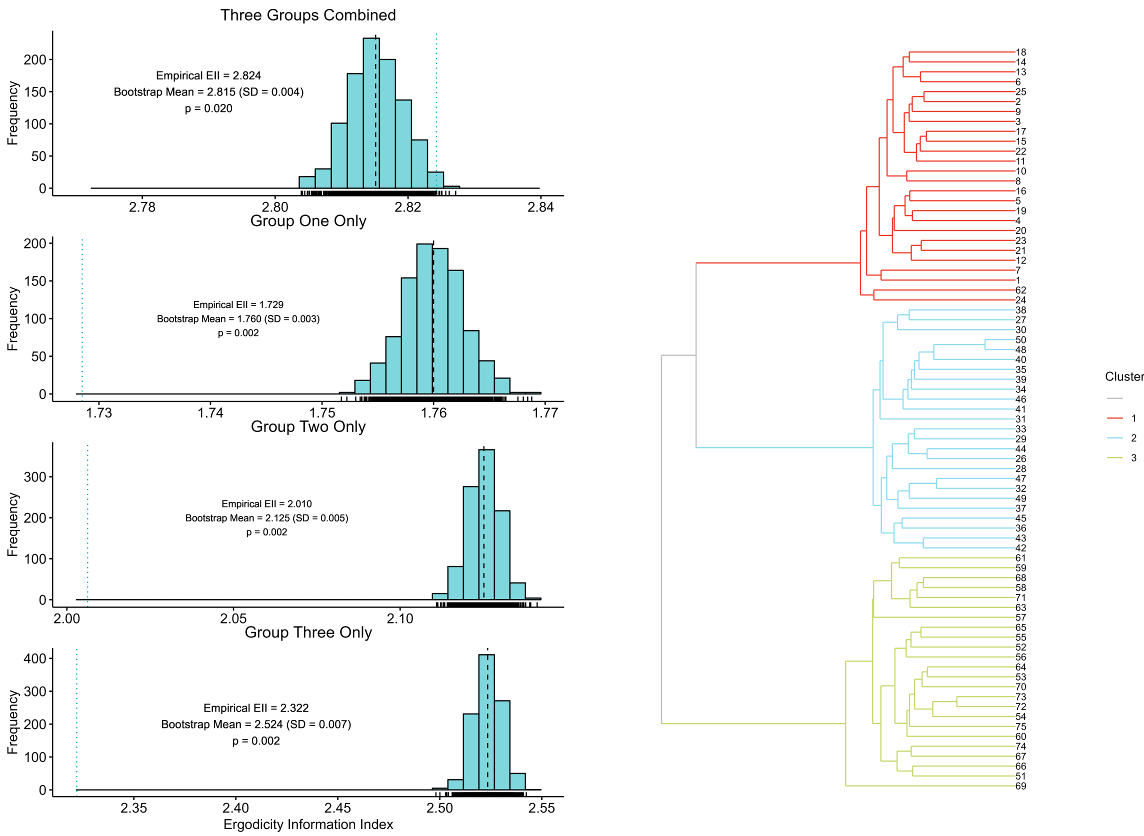


Figure 9. The results of the bootstrap EII test (left) and the resulting clusters from the information theory clustering (right). The top left bootstrap EII test is the result when collapsed across groups, the middle left represents the test for group one only, and the bottom left represents the test for group two only. On the right, the dendrogram depicts the splits of the two groups identified by the information theory clustering approach.

quality, we only included participants who completed at least 20 time points and had network densities of at least 0.15 (i.e., at least 15% of all possible connections present). These criteria narrowed the final sample to 119 participants.

The personality data was found to be nonergodic, $EII = 3.688, p = 0.006, M_{EII} = 3.681$ $(SD = 0.002)$. When followed up with the information theory clustering, 2 clusters were identified (cluster 1 and cluster 2-7, respectively; Figure 10). To determine whether these clusters were ergodic, the EII test was applied to groups separately. Cluster 1 $(N = 44)$ was found to be ergodic, $EII = 3.994, p = 0.002, M_{EII} = 4.033$ $(SD = 0.002)$; cluster 2 $(N = 75)$, however, was nonergodic, $EII = 3.616, p = 0.002, M_{EII} = 3.605$ $(SD = 0.002)$. Because cluster 2 was

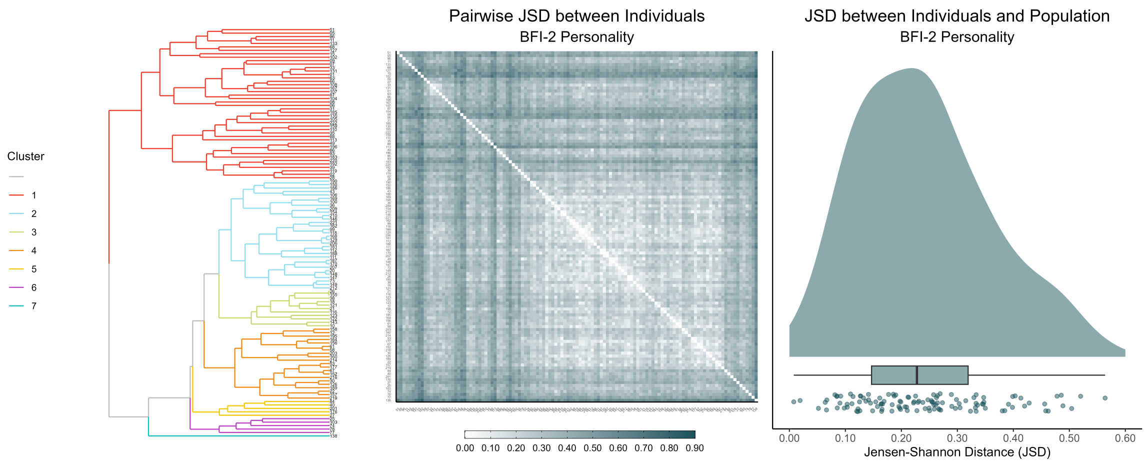


Figure 10. The left figure depicts the hierarchical clustering solution identified by the information theory method. The middle figure depicts the JSD between each individual that was used in the clustering approach (smaller values equal more similar). The right figure depicts the JSD of each individual to the population structure.

nonergodic, the information clustering was followed up on that subgroup only. Six clusters were identified, ranging in size from 1 to 32, suggesting that there were smaller groups and perhaps an outlier remaining (Figure 10). Our procedure stopped at this point; however, in an applied setting, it would be recommended to follow through on clusters until ergodic subgroups can be identified.

Figure 11 shows the population structure with three individuals that represent the lowest JSD (closest to the population), median JSD, and highest JSD (furthest from the population).

For the brain data example, we obtained open-source resting-state data from a sample of younger ($N = 34$) and older ($N = 28$) adults ($N = 62$ in total; retrieved from: https://doi.org/10.18112/openneuro.ds003871.v1.0.2). The original study examined episodic memory differences measured by a pattern separation task that were related to hippocampal connectivity to the default mode network (Wahlheim, Christensen, Reagh, & Cassidy, 2022). Given that this sample had defined groups of younger and older adults that demonstrated significant differences in their connectivity related to memory, it serves as an example where there are a priori groups and ergodicity is not expected to hold.

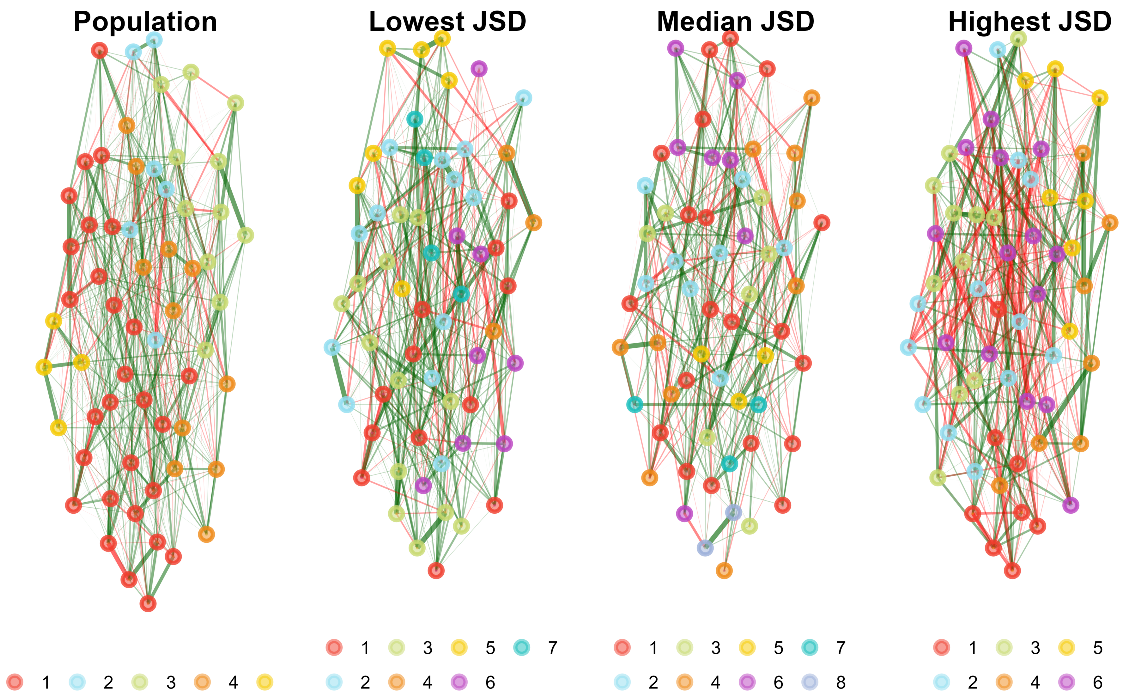


Figure 11. Examples of the individual personality network structures relative to the population structure. The population network is on left followed by the most similar, median similarity, and least similar individual networks.

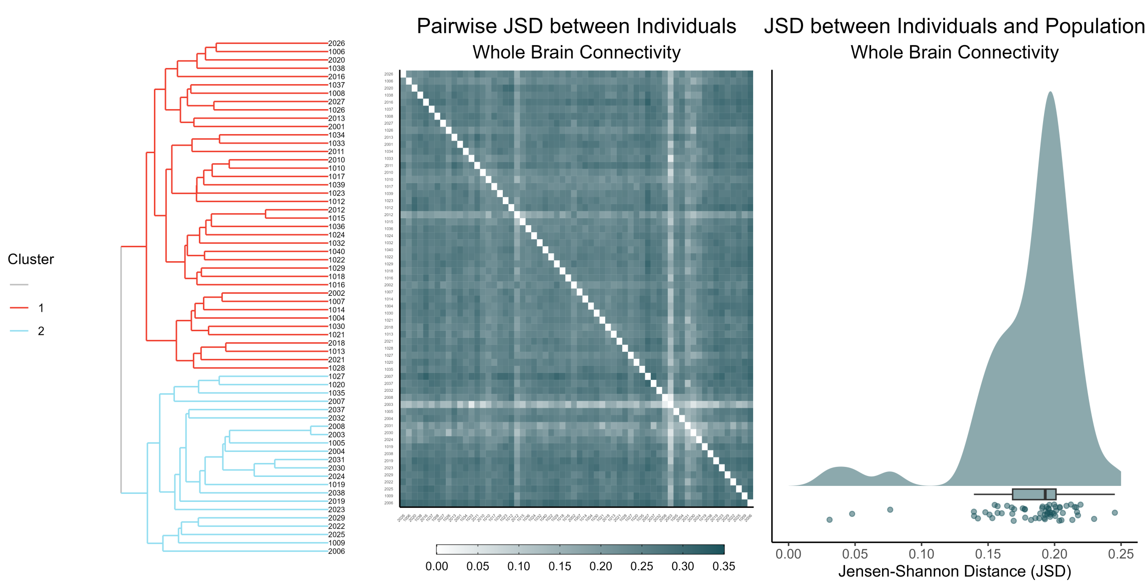


Figure 12. The left figure depicts the hierarchical clustering solution identified by the information theory method. The middle figure depicts the JSD between each individual that was used in the clustering approach (smaller values equal more similar). The right figure depicts the JSD of each individual to the population structure.

The brain data was found to be nonergodic, $EII = 7.945, p = 0.923, M_{EII} = 7.946$ $(SD = 0.010)$. Because there were two a priori groups of younger and older adults, we separated the groups and performed the EII test on them separately. Both the younger ($EII = 6.106, p = 0.002, M_{EII} = 6.182$ $(SD = 0.007)$) and older ($EII = 5.387, p = 0.002, M_{EII} = 5.607$ $(SD = 0.010)$) adults brain connectivity were found to be ergodic suggesting that their population structure can capture the processes that exist in their respective groups.

When followed up with the information theory clustering, 2 clusters were identified (Figure 12). These clusters were relatively similar in proportion to the younger and older adult numbers: 40 in cluster 1 and 22 in cluster 2. The clustering approach correctly identified 28 out of 34 (82.4%) younger adults and 16 out of 28 (57.1%) older adults for a total accuracy of 71.0%. This accuracy was significantly ($p = 0.007$) greater than the no information rate (54.8%) and had moderate agreement based on kappa (0.403).[2] Figure 13 shows the population structure with three individuals that represent the lowest JSD (closest to the population), median JSD, and highest JSD (furthest from the population).

Discussion

For well over a century the study of variation has been the backbone of psychologists' efforts to understand behavior and behavior change. Whether created via experimental manipulation or measured as it exists in nature (Cronbach, 1957; e.g., Cronbach, 1975), variation is the ore that has been dug up, assayed, weighed, and otherwise analyzed by behavioral prospectors hoping to strike it rich. Between-person or interindividual variation has been and is strongly favored in psychological research. Often, researchers infer within-person or idiographic processes based on between-person variation including many prominent theories. Indeed, some approaches to within-person measurement on some amount

[2] Values were computed using the {caret} package (version 6.0.94; Kuhn, 2008) in R.

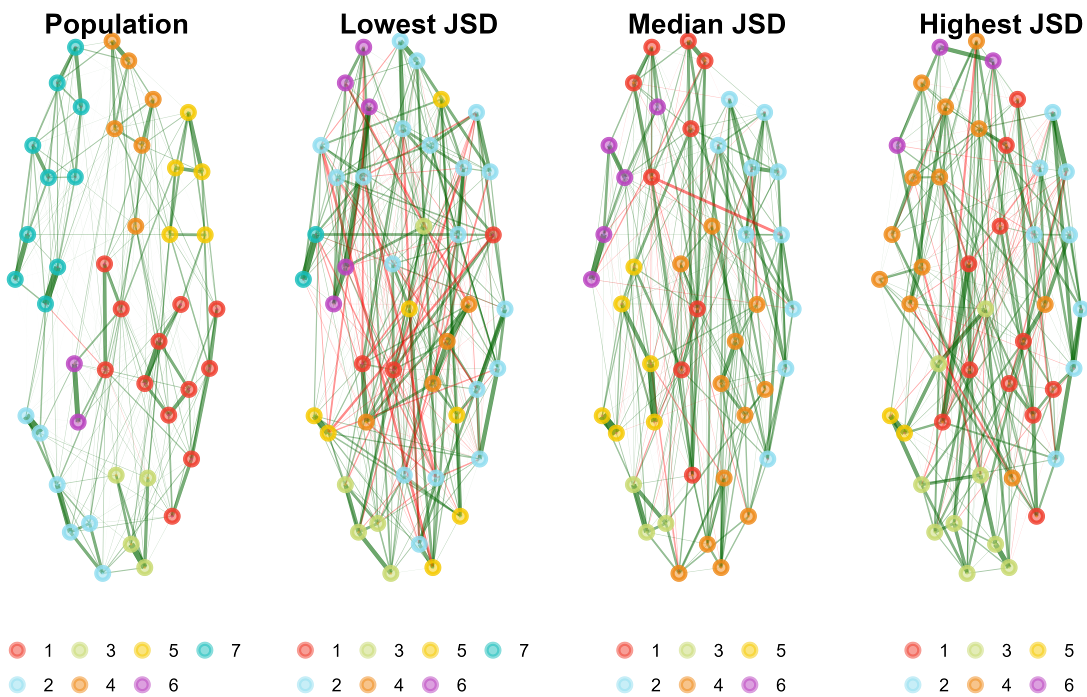


Figure 13. Examples of the individual brain network structures relative to the population structure. The population network is on left followed by the most similar, median similarity, and least similar individual networks.

of group-level structures (Gates & Molenaar, 2012). The reliance on interindividual variation has led to a field that appears to implicitly assume ergodicity is a fundamental property of psychological processes. Despite this ubiquitous assumption, many researchers have chosen to focus on how the behavior of an individual could vary from one measurement occasion to the next (Beck & Jackson, 2022; Fisher et al., 2018).

Intraindividual variability is no longer studied by a curious few. It has grown into a prominent psychological method over the last few decades to the point that it warrants the undivided attention of both methodologists and substantive theorists. This article takes steps toward more rigorous testing about whether researchers can safely ignore the different patterns of intraindividual variability in order to construct more general interindividual representations. We leveraged tools from dynamical systems, network science, and information theory to develop tests to the determine cost of aggregation. DynEGA paired with GLLA allows researchers to study how variables change together over time and the general structure of their relations (Golino et al., 2022). The EII provides researchers with a relative metric and bootstrap test to determine how much information is lost when aggregating individuals into a single population structure and whether that amount is significant. If significant information is lost, then the information clustering method can be applied.

The EII index is based on the concept of Kolmogorov (or algorithm) complexity that "measures" the shortest description length of an object. When describing a picture, for example, one could describe every pixel value individually which would take a lot of information. Or it could be described more succinctly by compressing redundant patterns. The compressed description length gets at the inherent complexity of the object being described. For networks, rather than describing every single edge weight individually, we can describe them in a simpler, more parsimonious way by identifying patterns and structure. The Kolmogorov complexity of a network is based on the length of the simplest and most compact description of the network—that is, by compressing the (weighted or unweighted) edge list of the network. Greater compressibility means lower Kolmogorov complexity. The EII leverages this principle. By comparing the Kolmogorov complexity of the individual networks (encoded into a single network using prime-weight encoding) with the Kolmogorov complexity of the population network, it quantifies the amount of information lost in the process of representing the former with the latter. Another way to interpret EII is as a measure of information for the individual networks with respect to the population network. For example, EII shows that more information is lost by representing the eight measures in Figure 6 as a single structure for individuals three and four (bottom of Figure 6, population 2) than for individuals one and two (top of Figure 6, population 1). Because individuals one and two on Figure 6 are similar, a single population structure representing both does not lead to a loss of information about any of them. On the other side, individuals three and four (bottom of Figure 6, population 2) are quite different, and representing them with a single population network leads to loss of information. A higher EII means the individual networks require much more information relative to what the population network can provide. Thus, information is lost aggregating to a population description, or representing all individuals with the same population network.

We propose that EII quantifies the extent to which *super-weak ergodicity* holds in the system. Super-weak ergodicity suggests that the individual structures of a system should reflect, within reasonable error, the aggregate structure of the system. This level of ergodicity is a minimum requirement of a system to be represented as an aggregate. In other words, EII captures the *topological ergodicity* or *structural ergodicity*. Systems that do not possess this property should not be aggregated because significant information is lost—such that a mere fraction of the system can be expected to reflect the aggregate system. In its present state, psychological processes are unlikely to be ergodic (Molenaar, 2004). Therefore, the measurement of ergodicity must be pursued in perpetuity. Our position, as well as the position of many others (e.g., Borsboom, Kievit, Cervone, & Hood, 2009; Fisher et al., 2018; Molenaar, 2004), is that psychological processes should be considered nonergodic until they are repeatedly demonstrated otherwise.

The results of our simulation study are very clear in terms of conditions the ergodicity information index works the best, and conditions in which it fails. In the current paper we aimed to shed light on the accuracy and performance of three EII methods: *edge list*, *unweighted*, and *weighted*. The assessment of accuracy, also known as the "hit rate," demonstrated some interesting trends. The unweighted EII method outperformed the other methods in terms of accuracy, although the differences were marginal. While the unweighted method achieved the highest accuracy at 82.98%, the weighted and edge list methods closely followed with accuracy rates of 82.97% and 82.75%, respectively.

Notably, the impact of error levels on EII accuracy is a crucial consideration. In scenarios with low error (0.125), all EII methods performed exceptionally well, maintaining accuracy at around 99.98%. However, as error increased to 0.25 and then substantially to

0.50, we observed a significant decline in accuracy for all methods, emphasizing the sensitivity of EII to measurement quality. Researchers should be cautious about employing EII in settings where data or measurement quality is compromised.

The influence of factor loadings on EII accuracy was another key aspect of our study. It became evident that higher loadings were associated with better EII performance, as accuracy improved from approximately 69% for loadings of 0.4 to around 91% for loadings of 0.8. This finding underscores the importance of high quality items/variables in intensive longitudinal research.

Similarly, the number of variables in the reference group was found to have a moderate impact on accuracy. In cases with four variables per factor, accuracy was approximately 75% for all methods, while increasing the number to six variables per factor improved accuracy to around 90%. This suggests that researchers should consider the number of variables per factor when designing their data collection strategies, since it impacts the accuracy of EII. Finally, the simulation results also shows that conditions with more factors than items per factor (e.g., four factors with two items each) can compromise the accuracy of EII, and should be avoided in empirical research.

Overall, our study provides valuable insights into the performance of EII methods across diverse conditions. It's clear that the unweighted EII is a robust choice, demonstrating high accuracy and offering better differentiation between conditions. However, it's important to remain mindful of the potential limitations, particularly the sensitivity to data and measurement quality and the dependence on moderate or high factor loadings. By adhering to these considerations, researchers can make the most of the Ergodicity Information Index in their investigations. Another limitation of the simulation is that we did not implement cross-regressive parameters. The inclusion of error can add unsystematic cross-loadings to the structure; however, these patterns were not systematically manipulated. Future studies should evaluate the effect of the cross-regressive parameters on the performance of EII.

Our empirical examples take two commonly aggregated psychological phenomena, personality and (resting state) brain activity, and examined the extent to which they lose information when aggregated into a single population network. We show that the personality and brain networks were nonergodic. For personality, these results question the extent to which the Big Five generalize to individuals (Borkenau & Ostendorf, 1998). The information theory clustering revealed that there was a large portion of the sample ($N = 64$) that were similar enough to form a considerable cluster. There were, however, many other smaller and somewhat fragmented clusters.

For the brain networks, there were two age groups that allowed us to leverage both determining whether younger and older adult brain networks could be collapsed to form a single aggregate network but also whether these groups could be aggregated separately. Our results found evidence for the latter but not the former—in line with expectations as well as previous results using the sample (Wahlheim et al., 2022). The information clustering approach performed moderately well to disentangle the age groups in the sample demonstrating the approach can identify group-level differences when they exist. Taken together, our results suggest that distinct processes can be lost if modeled as a single aggregate structure.

What are the consequences of our findings? For personality, the best case seems to be that measurement error and sampling variability could affect the extent to which our approach can determine ergodicity. There are many underlying factors that make this interpretation plausible such as the specific personality scale used (BFI-2). Our findings, however, are neither the first to suggest that the Big Five does not hold for individual people (e.g., Borkenau & Ostendorf, 1998). We believe that our results add to the growing evidence that aggregate structures in personality are not isomorphic to individual structures (for a thorough review, see Richters, 2021).

For brain networks, we demonstrate the importance of identifying groups in the data.

Our results focused on resting-state functional connectivity where participants are not performing a specific task, allowing their thoughts to flow freely. Younger and older adults appear to have different underlying processes that when separated are ergodic. Because of the small sample size, we are cautious to suggest that younger and older adult hippocampal and default mode network connectivity are ergodic across all individuals. Given the precedent in the neuroimaging literature, more work is necessary to determine how common finding ergodicity is in these samples (Lurie et al., 2020; Medaglia et al., 2011).

Psychology aims to understand the thoughts, emotions, and behavior of the person and people. Despite repeated calls and manifestos, the study of people continues to dominate psychology. With modern technology, everyday thoughts, emotions, and behaviors of the person have never been more accessible. As psychologists begin to emphasize intraindividual processes over interindividual predictions, statistical models that answer questions about the dynamics of systems are needed (Epskamp et al., 2018b; e.g., Gates & Molenaar, 2012; Moulder, Martynova, & Boker, 2023; Sterba & Bauer, 2010).

To the best of our knowledge nowhere is it written that to build a science of behavior the aggregation of information between individuals should take precedence over the aggregation of information within individuals. Priorities can and should be data-driven. The present work provides one statistical tool to understand whether intraindividual structures can be reasonably aggregated into a single interindividual structure and another tool to determine possible sub-aggregations when they cannot. Together, these tools allow researchers to establish generalizability starting with the person rather than searching for it across people (Nesselroade & Molenaar, 2016).

Despite the significant contributions our new methods have made to the field of ergodicity analysis, it is crucial to approach the interpretation of the ergodicity information index (EII) and its test with caution. In the context of psychological measures, particularly when dealing with what we refer to as “super-weak ergodicity” (or “topological/structural ergodicity"), as indicated by the EII and its test, it becomes possible to compute scores that capture the underlying latent dynamic processes. Consider emotions, for instance. Imagine a research scenario examining positive and negative emotions through a set of items for each of these factors in an intensive longitudinal study.

If the EII and its test suggest that the structure, as estimated through dynamic exploratory graph analysis, allows for representing all individuals with the same population structure (such as a network with two communities or factors) without significant loss of information, then it becomes feasible to compute network scores for positive and negative emotions for all individuals. However, it's essential to note that super-weak ergodicity (or topological/structural ergodicity) does not necessarily imply the ergodicity of the latent dynamic process. Some individuals might exhibit more variability in positive emotions compared to negative emotions, while others may not display liability in either positive or negative emotions. Consequently, representing these diverse emotion dynamics with a single time series of emotion scores per factor may fail to capture the distinctions within these two groups of individuals at the latent dynamics level.

In such cases, the techniques proposed by Janczura and Weron (2015), Domowitz and El-Gamal (2001), and Loch et al. (2016) can be employed on the time series of emotions (both positive and negative) for each individual to investigate a second tier of ergodicity, which we can term *dynamic ergodicity*. Therefore, the interpretation of "super-weak ergodicity" or "topological/structural" ergodicity and "dynamic ergodicity" parallels the interpretation of "structural invariance" and "metric invariance" in the realm of psychometrics. Structural invariance is a necessary but not sufficient condition for metric invariance, just as super-weak or structural ergodicity is necessary but not sufficient for dynamic ergodicity. Super-weak and dynamic ergodicity thus represent two distinct levels of analysis, one related to the structural aspects and the other focused on the dynamic characteristics of latent trends.

Limitations and Future Directions

The exploration of the ergodic nature of intensive longitudinal data in our research offers valuable insights for the field of psychology and other areas in which ergodicity is an important issue. Yet it's essential to acknowledge the broader landscape of time series analysis. While our simulation have addressed critical conditions to investigate the validity of the ergodicity information index, future research should consider expanding the work implemented in the current paper. For example, we don't know the effect of the adding more than two groups (for the "NotEq" condition) in the accuracy of EII or if using a different data generation mechanism (e.g., Dumped Linear Oscillator model) will affect our new index. Further, conditions likely to be key in the performance of the EII will need to be thoroughly tested in future simulation studies such as fewer individuals, shorter time series, different autoregressive coefficients, and other network estimation methods (e.g., Williams & Rast, 2019; Williams et al., 2020).

Real-world data can be influenced by various factors not considered in our simulation. Issues such as the presence of influential observations, ceiling and floor effects, or localized periods of ergodic behavior interspersed with instability offer a promising avenue for future research. Additionally, individual-level seasonality effects in behavioral data, such as mood changes tied to weekdays versus weekends or regular social interactions, may influence ergodicity. Incorporating these factors into future research can offer a more comprehensive understanding of ergodicity in various contexts.

In addition, we introduce a bottom-up clustering approach based on information theory to identify the number of possible subgroups in the data. Our introduction and demonstration to this approach involves one simulated and two empirical examples. More extensive testing is needed to validate this approach such as comparing it to contemporary methods such as Group Iterative Multiple Model Estimation (GIMME; Gates & Molenaar, 2012; Lane, Gates, Pike, Beltz, & Wright, 2019) and vector autoregression methods (Park, Fisher, Chow, & Molenaar, 2023).

In summary, our research provides valuable insights into the ergodic nature of intensive longitudinal data. However, the field of time series analysis is multifaceted, with various methodologies and perspectives. We acknowledge the need for further exploration, considering alternative explanations and factors that may influence ergodicity in real-world data. This discussion aims to pave the way for continued research in this area, ultimately advancing our understanding of time series dynamics and ergodic and nonergodic processes.

References

Adolf, J., Schuurman, N. K., Borkenau, P., Borsboom, D., & Dolan, C. V. (2014). Measurement invariance within and between individuals: A distinct problem in testing the equivalence of intra-and inter-individual model structures. *Frontiers in Psychology*, *5*, 883.

Almendral, J. A., & Dıaz-Guilera, A. (2007). Dynamical and spectral properties of complex networks. *New Journal of Physics*, *9*(6), 187. https://doi.org/10.1088/1367-2630/9/6/187

Anderson, H., T. W. & Rubin. (1958). Statistical inference in factor analysis. *Proceedings of the 3rd Berkeley Symposium on Mathematics, Statistics, and Probability*, *5*, 111–150.

Beck, E. D., & Jackson, J. J. (2022). Idiographic personality coherence: A quasi experimental longitudinal ESM study. *European Journal of Personality*, *36*(3), 391–412. https://doi.org/10.1177/08902070211017746

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison, WI: University of Wisconsin Press.

Boker, S. M. (2018). *Longitudinal multivariate psychology* (E. Ferrer, S. M. Boker, & K. J. Grimm, Eds.). Routledge.

Boker, S. M., Deboek, 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.), *The notre dame series on quantitative methodology. Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 161–178). Routledge/Taylor & Francis Group.

Borkenau, P., & Ostendorf, F. (1998). The Big Five as states: How useful is the five-factor model to describe intraindividual variations over time? *Journal of Research in Personality*, *32*(2), 202–221. https://doi.org/10.1006/jrpe.1997.2206

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 67–97). New York, NY: Springer. https://doi.org/10.1007/978-0-387-95922-1_4

Cattell, R. B. (1965). *Studies in psychology* (C. Banks & P. L. Broadhurst, Eds.). University of London Press London.

Cattell, R., Cattell, A., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysiological source traits in a normal individual. *Psychometrika*, *12*(4), 267–288.

Chauhan, S., Girvan, M., & Ott, E. (2009). Spectral properties of networks with community structure. *Physical Review E*, *80*(5), 056114. https://doi.org/10.1103/PhysRevE.80.056114

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771. Retrieved from https://www.jstor.org/stable/20441500

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*(2), 116.

De Domenico, M., Nicosia, V., Arenas, A., & Latora, V. (2015). Structural reducibility of multilayer networks. *Nature Communications*, *6*(1), 1–9. https://doi.org/10.1038/ncomms7864

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

Diehl, M., Hooker, K., & Sliwinski, M. J. (2014). *Handbook of intraindividual variability across the life span*. East Sussex, BN: Routledge.

Domowitz, I., & El-Gamal, M. A. (2001). A consistent nonparametric test of ergodicity for time series with applications. *Journal of Econometrics*, *102*(2), 365–398.

Downey, R. G., & Hirschfeldt, D. R. (2010). *Algorithmic randomness and complexity*. Springer Science & Business Media.

Epskamp, M., S. (2018). Network psychometrics. In B. Irwing Paul (Ed.), *The wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 953–986). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118489772.ch30

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. https://doi.org/10.18637/jss.v048.i04

Epskamp, S., & Fried, E. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network pschometrics: Combining network and latent variable models. *Psychometrika*, *82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018b). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018a). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106–E6115.

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology*, *126*(8), 1044.

Foygel, R., & Drton, M. (2010). Extended bayesian information criteria for gaussian graphical models. *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1*, *1*, 604–612. Vancouver, Canada.

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

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026

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., & Christensen, A. P. (2019). *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., Christensen, A. P., Moulder, R., Kim, S., & Boker, S. M. (2022). Modeling latent topics in social media using dynamic exploratory graph analysis: The case of the right-wing and left-wing trolls in the 2016 US elections. *Psychometrika*, *87*(1), 156–187.

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

Gomes, C. M. A., & Golino, H. F. (2015). Factor retention in the intra-individual approach: Proposition of a triangulation strategy. *Avaliaçao Psicologica: Interamerican Journal of Psychological Assessment*, *14*(2), 2.

Gonzales, J. E., & Ferrer, E. (2014). Individual pooling for group-based modeling under the assumption of ergodicity. *Multivariate Behavioral Research*, *49*(3), 245–260.

Guttman, L. (1953). Image theory for the structure of quantitative variates. *Psychometrika*, *18*(4), 277–296.

Hamaker, E. (2022). The curious case of the cross-sectional correlation. *Multivariate Behavioral Research*, 1–12.

Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S. W. S. (2008). Intraindividual variability, cognition, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 491–556). Psychology Press.

Hunter, M. D., Fisher, Z., & Geier, C. F. (2023). What ergodicity means for you. *PsyArxiv*. https://doi.org/https://doi.org/10.31234/osf.io/25qsu

Janczura, J., & Weron, A. (2015). Ergodicity testing for anomalous diffusion: Small sample statistics. *The Journal of Chemical Physics*, *142*(14).

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, *106*(4), 620.

Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics*, *2*(1-4), 157–168.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*, 1–26. https://doi.org/10.18637/jss.v028.i05

Lamiell, J. T. (1998). Nomothetic and idiographic: Contrasting windelband's understanding with contemporary usage. *Theory & Psychology*, *8*(1), 23–38. https://doi.org/10.1177/0959354398081002

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*, *24*(1), 54. https://doi.org/10.1037/met0000192

Lauritzen, S. L. (1996). *Graphical models* (Vol. 17). Oxford: Clarendon Press.

Liu, S., Zhou, Y., Palumbo, R., & Wang, J.-L. (2016). Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data. *Psychological Methods*, *21*(3), 291.

Loch, H., Janczura, J., & Weron, A. (2016). Ergodicity testing using an analytical formula for a dynamical functional of alpha-stable autoregressive fractionally integrated moving average processes. *Physical Review E*, *93*(4), 043317.

Lurie, D. J., Kessler, D., Bassett, D. S., Betzel, R. F., Breakspear, M., Kheilholz, S., et al.others. (2020). Questions and controversies in the study of time-varying functional connectivity in resting fMRI. *Network Neuroscience*, *4*(1), 30–69. https://doi.org/10.1162/netn_a_00116

Martin-Löf, P. (1966). The definition of random sequences. *Information and Control*, *9*(6), 602–619.

Medaglia, J. D., Ramanathan, D. M., Venkatesan, U. M., & Hillary, F. G. (2011). The challenge of non-ergodicity in network neuroscience. *Network: Computation in Neural Systems*, *22*(1-4), 148–153.

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Molenaar, P. C., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). Springer.

Molenaar, P. C., & Nesselroade, J. R. (2012). Merging the idiographic filter with dynamic factor analysis to model process. *Applied Developmental Science*, *16*(4), 210–219. https://doi.org/10.1080/10888691.2012.722884

Morzy, M., Kajdanowicz, T., & Kazienko, P. (2017). On measuring the complexity of networks: Kolmogorov complexity versus entropy. *Complexity*, *2017*.

Moulder, R. G., Martynova, E., & Boker, S. M. (2023). Extracting nonlinear dynamics from psychological and behavioral time series through HAVOK analysis. *Multivariate Behavioral Research*, *58*(2), 441–465. https://doi.org/10.1080/00273171.2021.1994848

Nesselroade, J. R., & Ford, D. H. (1985). P-technique comes of age: Multivariate, replicated, single-subject designs for research on older adults. *Research on Aging*, *7*(1), 46–80. https://doi.org/10.1177/0164027585007001003

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. M. (2002). Dynamic factor analysis models for representing process in multivariate time-series. In D. S. Moskowitz & S. L. Hershberger (Eds.), *Multivariate applications book series. Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 235--265). Lawrence Erlbaum Associates Publishers.

Nesselroade, J. R., & Molenaar, P. C. (2016). Some behaviorial science measurement concerns and proposals. *Multivariate Behavioral Research*, *51*(2-3), 396–412.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). *Statistical strategies for small sample research* (R. H. Hoyle, Ed.). London: Sage.

Nesselroade, J. R., & Molenaar, P. C. M. (2010). *The handbook of life-span development, vol. 1. Cognition, biology, and methods* (W. F. O. & R. M. Lerner, Ed.). John Wiley & Sons, Inc. https://doi.org/https://doi.org/10.1002/9780470880166.hlsd001002

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

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

Oertzen, T. von, Schmiedek, F., & Voelkle, M. C. (2020). Ergodic subspace analysis. *Journal of Intelligence*, *8*(1), 3.

Park, J. J., Fisher, Z. F., Chow, S.-M., & Molenaar, P. C. (2023). Evaluating discrete time methods for subgrouping continuous processes. *Multivariate Behavioral Research*, 1–13. https://doi.org/10.1080/00273171.2023.2235685

Peters, O. (2019). The ergodicity problem in economics. *Nature Physics*, *15*(12), 1216–1221.

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. In Pi. Yolum, T. Güngör, F. Gürgen, & C. Özturan (Eds.), *Computer and information sciences - ISCIS 2005* (pp. 284–293). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11569596_31

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

Ram, N., Gerstorf, D., Lindenberger, U., & Smith, J. (2011). Developmental change and intraindividual variability: Relating cognitive aging to cognitive plasticity, cardiovascular lability, and emotional diversity. *Psychology and Aging*, *26*(2), 363.

Richters, J. E. (2021). Incredible utility: The lost causes and causal debris of psychological science. *Basic and Applied Social Psychology*, *43*(6), 366–405. https://doi.org/10.1080/01973533.2021.1979003

Rosenstein, M. T., Collins, J. J., & De Luca, C. J. (1993). A practical method for calculating largest lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, *65*(1-2), 117–134. https://doi.org/10.1016/0167-2789(93)90009-P

Santoro, A., & Nicosia, V. (2020). Algorithmic complexity of multiplex networks. *Physical Review X*, *10*(2), 021069.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2020). Training working memory for 100 days: The COGITO study. In *Cognitive and working memory training: Perspectives from psychology, neuroscience, and human development* (pp. 40–57). Oxford University Press.

Schmiedek, F., Lövdén, M., Oertzen, T. von, & Lindenberger, U. (2020). Within-person structures of daily cognitive performance differ from between-person structures of cognitive abilities. *PeerJ*, *8*, e9290.

Shen, A., Uspensky, V. A., & Vereshchagin, N. (2022). *Kolmogorov complexity and algorithmic randomness* (Vol. 220). American Mathematical Society.

Solomonoff, R. J. (1964). A formal theory of inductive inference. *Information and Control*, *7*(1), 1–22. https://doi.org/10.1016/S0019-9958(64)90223-2

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented developmental psychopathology research. *Development and Psychopathology*, *22*(2), 239–254. https://doi.org/10.1017/S0954579410000015

Takens, F. (1981). Detecting strange attractors in turbulence. In *Lecture notes in mathematics (vol. 898)* (pp. 366–381). Springer. https://doi.org/10.1007/BFb0091924

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288.

https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Velupillai, K. V. (2011). Non-linear dynamics, complexity and randomness: Algorithmic foundations. *Journal of Economic Surveys*, *25*(3), 547–568.

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213.

Wahlheim, C. N., Christensen, A. P., Reagh, Z. M., & Cassidy, B. S. (2022). Intrinsic functional connectivity in the default mode network predicts mnemonic discrimination: A connectome-based modeling approach. *Hippocampus*, *32*(1), 21–37. https://doi.org/10.1002/hipo.23393

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244. https://doi.org/10.2307/2282967

West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality*, *59*(3), 609–662.

Whitney, H. (1936). Differentiable manifolds. *The Annals of Mathematics*, *37*(3), 645–680. https://doi.org/10.2307/1968482

Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*(3), 263–311. https://doi.org/10.1207/s15327906mbr2803_1

Williams, D. R., & Rast, P. (2019). Back to the basics: Rethinking partial correlation network methodology. *British Journal of Mathematical and Statistical Psychology*, 1–25. https://doi.org/doi: 10.1111/bmsp.12173

Williams, D. R., & Rast, P. (2020). Back to the basics: Rethinking partial correlation network methodology. *British Journal of Mathematical and Statistical Psychology*, *73*(2), 187–212. https://doi.org/10.1111/bmsp.12173

Williams, D. R., Rast, P., Pericchi, L. R., & Mulder, J. (2020). Comparing gaussian graphical models with the posterior predictive distribution and bayesian model selection. *Psychological Methods*, *25*(5), 653.

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

Wright, A. G., & Zimmermann, J. (2019). Applied ambulatory assessment: Integrating idiographic and nomothetic principles of measurement. *Psychological Assessment*, *31*(12), 1467.

Zenil, H., Kiani, N. A., & Tegnér, J. (2018). A review of graph and network complexity from an algorithmic information perspective. *Entropy*, *20*(8), 551.

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 377–402. https://doi.org/doi.org/10.1080/10705510802154281