Within-person replicability of idiographic networks

Ricarda K. K. Proppert[1]

[1] Clinical Psychology, Leiden University, The Netherlands

Author Note

Correspondence concerning this article should be addressed to Ricarda K. K. Proppert, . E-mail: ricarda.proppert@gmail.com

Abstract

Evidence-based mental health programs have long conceptualized mental disorders in terms of interactions between thoughts, feelings, behaviours and external factors.

Idiographic network models are a relatively novel way of modeling such intra-individual psychological processes. However, these methods are not without limitations, and concerns have been raised about the stability and accuracy of estimated networks.

While methods to assess network parameter accuracy have been developed for cross-sectional data, no such method exists for single-subject data. The extend to which idiographic networks are stable, or vary over time, is unknown.

In the current work, we reanalyse daily symptom records of people with personality disorders to explore the stability of idiographic networks over time, as well as the degree to which network stability varies across individuals. We further explore antecents that may relate to inter-individual variation in network stability using predictive LASSO regression.

Keywords:

Word count:

Within-person replicability of idiographic networks

Introduction

Idiographic psychological networks

Idiographic network models are of growing interest to clinical psychology because they may address two recently voiced calls in clinical psychology: Firstly, there seems to be a need for psychological research to re-orient towards idiographic methods that study intra-individual processes as opposed to group-level differences (Molenaar, 2004). Secondly, scholars are proposing a paradigm shift away from reductionism towards studying the complexity of psychological phenomena. The Network theory of mental disorders (Borsboom & Cramer, 2013; Cramer, Waldorp, Maas, & Borsboom, 2010) attempts to integrate psychology with insights and methods from complexity science, proposing a novel but well-received theoretical framework to study and understand the underpinnings of psychopathology. The network theory of mental disorders conceptualizes psychopathology as an emergent state of dynamically interacting psychological symptoms and states, as well as factors external to this system. This account seems closely aligned with established clinical practices where informal case conceptualizations in form of path diagrams are used to describe the proposed mechanisms of a given disorder (Burger et al. (2020), Scholten, Lischetzke, and Glombiewski (2020)).

With the ongoing development of statistical methods to visualize and quantify such system's structures and behaviors, called psychological networks (Epskamp, Borsboom, & Fried, 2018), an increasing body of research is now applying this framework to study symptom networks in people with mental disorders. Psychological networks consist of elements (nodes) and their pairwise interactions (edges), together forming a complex system. Nodes represent psychological or other variables, such as psychological or somatic symptoms, stressors, or behaviors. Edges represent pairwise relationships between these variables. These relationships may be directed or undirected, positive or negative, and can differ in strength. These associations and are thought to result from of a multitude of mechanisms which may not necessarily be known or defined.

Temporal stability of networks

Most of the early psychological network research focused on group-level analyses and cross-sectional comparisons, mainly for practical reasons such as availability of existing data-sets. Few studies have investigated longitudinal network stability at a group level (e.g., in veterans pre- to post-combat, Segal et al. (2020); military veterans, Stockert, Fried, Armour, and Pietrzak (2018); among adolescent earthquake survivors, Ge, Yuan, Li, Zhang, and Zhang (2019); and in patients with Anxiety and Depression, Curtiss, Ito, Takebayashi, and Hofmann (n.d.)).

Besides group-level analyses, there is growing interest in idiographic network models of intra-individual symptom dynamics. Researchers in clinical psychology in particular hope that such models could resolve what is known as the Therapist's dilemma (eg., Frumkin, Piccirillo, Beck, Grossman, and Rodebaugh (2020), Howe, Bosley, and Fisher (2020), Caviglia and Coleman (2016), Hoffart and Johnson (2020)). In clinical psychology research, not only momentary network structures are of interest, but especially *changes* in network structure over time. For example, changes in network structure may indicate therapeutic progress (Thonon, Van Aubel, Lafit, Della Libera, & Larøi, 2020) or relapse (Wichers, Groot, & Psychosystems, 2016). Even subtle changes in network structure are of interest, as they are thought to act as potential early warning signals which may predict future major change, i.e. a system's phase transition from a healthy attractor state to a disordered one. For example, signs of critical slowing down, showing as increased auto-correlations and variances of items, have been demonstrated to signal a patient's relapse into depression upon stopping antidepressant treatment (Wichers, Groot, and Psychosystems (2016); for similar work on resilience, see Kuranova et al. (2020)).

Empirical work is published in which differences in network structures are interpreted at face value. For example, Thonon, Van Aubel, Lafit, Della Libera, and Larøi (2020) followed three psychiatric patients over the course of treatment and interpreted changes in idiographic network structures over time as additional evidence for and a description of the observed therapeutic change. Such a substantial interpretation of network instability hinges on the assumption that differences in estimated network parameters accurately reflect change in the data-generating mechanism. This assumption rests on two conditions: First, parameter estimates need to be an accurate reflection of the true underlying relationship. Second, the underlying mechanism would have remained stable if no intervention took place (related to the assumption of stationarity made during model estimation). Neither of these conditions can currently be evaluated in rigorous ways for idiographic network models. Methods to assess the accuracy of parameter estimates have been developed for group-level estimation procedures (Epskamp, Borsboom, & Fried, 2018), but are not yet available for idiographic estimation methods. The degree to which underlying dynamic psychological processes are stable within individuals over time is unclear. Current literature investigating the temporal stability of idiographic networks mostly focuses on settings where change is expected to occur, e.g., in response to psychological treatment (Thonon, Van Aubel, Lafit, Della Libera, & Larøi, 2020), discontinuation of antidepressant medication (Wichers, Groot, & Psychosystems, 2016), or the COVID-19 pandemic (Emorie D. Beck & Jackson, 2021). To our knowledge, only one study has investigated the temporal stability of psychological networks in situations where no change would be expected. Emorie D. Beck and Jackson (2020) investigated the consistency of idiographic personality over the course of two years. They reported high consistency among contemporaneous associations, low consistency of temporal associations, and considerable interpersonal variability in the stability of networks.

Aim of this study

The present study aims to assess the stability of idiographic networks of psychopathology, as well as exploring potential person-specific factors and statistical factors predicting interpersonal variation in temporal network stability.

We will address the following research questions:

RQ1: How stable are estimated idiographic network structures over time?

RQ2: What factors, person-specific or of statistical nature, predict inter-individual variation in intra-individual replicability?

To this end, we re-analyze daily diary data collected in a clinical sample of people diagnosed with a personality disorder (Aidan G. C. Wright, Beltz, Gates, Molenaar, & Simms, 2015; Aidan G. C. Wright, Hopwood, & Simms, 2015; Aidan G. C. Wright & Simms, n.d.). Participants (N=116) provided once-daily ratings of their mood, behaviors, and daily stressors over the course of 100 consecutive days. To assess intra-individual stability of idiographic networks, we will fit separate idiographic networks on the first and last 50 days of measurement. Network structures of each individual's first vs. last 50 days of measurements will be compared by correlating estimated path estimates, with high correlation indicating high temporal stability. Finally, person-specific antecedents and statistical factors will be tested for their ability to predict interindividual variation in network stability using predictive LASSO regression.

Data Set

A summarized version of the data are published HERE, an overview of all available variables can be found in Appendix A.

The original study design by which the data were collected are described in detail in previous publications (Aidan G. C. Wright, Beltz, Gates, Molenaar, & Simms, 2015; Aidan

G. C. Wright, Hopwood, & Simms, 2015; Aidan G. C. Wright & Simms, n.d.). The study investigated daily dynamics in affect, stress, andexpressions of personality disorder, along with some lifestyle variables such as self-reported sleep, drug and alcohol use, and overall functioning. Participants were recruited from an ongoing clinical study (N=628, Simms et al. (2011)) that targeted individuals who had received psychiatric treatment within the previous two years, recruited via flyers distributed at mental health clinics across Western New York, USA. They were invited for study participation if they met the diagnostic requirements of any personality disorder during the initial structured clinical interview conducted for the parent study (SCID-II, TODO REF), and if they had daily internet access via a computer or mobile device. 116 participants were initially enrolled, of which 101 participants completed at least 30 of the desired 100 daily measurements.

Participants completed daily measurements over the course of 100 consecutive days. Participants completed surveys at roughly the same time each night, depending on their individual schedule. Therefore, therefore, this data set is expected to meet the assumption of roughly equal time intervals in between measurements (Epskamp, 2020). Furthermore, the original authors reported stability of means and variances of expressions of daily psychopathology in Aidan G. C. Wright and Simms (n.d.), indicating that the assumption of stationarity may be realistic (Epskamp, 2020). Participants started measurements asynchronically, so that no sample-wide history effects should affect our results on longitudinal within-person stability of estimated networks.

Analysis plan

Data pre-processing:

1) Missing data:

Participants with more than 30 days of non-response, or more than 15 in either the first or last 50 days of measurement, will be excluded from the analyses. Remainign missing data points of daily diary variables will be imputed using the Kalman Filter (Harvey, 1989) , which has been shown to perform well in simulation studies (Mansueto, Wiers, van Weert, Schouten, & Epskamp, 2020).

2) Detrending linear effects

Variables used for network estimation will be linearly detrended across the 100 day time course, at the level of the individual.

RQ1: Temporal stability estimation.

Variable selection. Because we expect high heterogeneity in sample in terms of which items apply to an individual's experience and variables with high negative skew (e.g. most responses being zero) are problematic for model estimation (violates model assumptions and can lead to non-convergence). We further want to make the variable selection process reproducible by basing the decision process on statistical criteria. Selecting variables purely based on statistical properties may result in networks that include only highly similar, closely related variables. Also, as resulting estimates of network stability depend on which variables are included in the network, having vastly different idiographic networks may confound these estimates. Therefore, want to optimize variable selection in a way that makes idiographic networks somewhat comparable across individuals, while capturing the unique behaviors that are related to individual psychopathology. We thus took a hybrid approach of variable selection based ontheoretical as well as statistical grounds:

Each idiographic network will include 3 composite variables (sum scores) which are common across participants: Positive Affect, Negative Affect, and Stress Severity. Further, we will include single-item variable capturing daily functioning. Per subject, two additional variables will be selected according to their rank on the following scoring algorithm:

Ranking metric = (1 - Shapiro-Wilk test statistic) * prop completed assessments T1 * proportion completed assessments T2

The Shapiro-Wilk test statistic tests the null hypothesis that a variable is sampled from a normal distribution, ranging from 0 to 1 Shapiro & Wilk (1965).

Network estimation. Idiographic network models will be estimated on an individual basis. Contemporaneous and temporal (Lag-1) associations will be estimated in form of a Gaussian Graphical Model using the open-source R package *graphicalVAR* (Epskamp, Waldorp, Mõttus, & Borsboom, 2018).

Lag-1 VAR models estimate temporal associations by predicting each variable by all variables in the network at the previous time point, including itself, using multivariate linear regression. The tuning parameter */gamma*, controlling the degree of regularization, will be set to 0.5 (default in R). Should this result in predominantly empty networks, */gamma* will be reduced to 0.25 or 0. All applied variations to */gamma* and their effects on the results will be reported.

Network comparisons. As an index of temporal stability, we compare idiographic networks estimated for the first and last 50 days of measurement by correlating estimated network edge weights. Other conceptualizations and tests of network similarity have been proposed, the most popular one being the Network Comparison Test (NCT, Borkulo et al. (2017)). However, performance of the NCT in sample sizes as small as 50 is unknown, as it has been tested and applied mostly for large-scale cross-sectional studies.

Exploratory analysis of contributing factors: The following baseline variables will be included tested as predictors:

- Sex

- Age

- Income

- past six months: Happy

- past six months: Mobility

- past six months: Impulse

- past six months: Relationships

- past six months: Work

- past week: Suicidality

- past year: Operation

- Handicap

- Cigarette

- Alcohol

- Substance

- Time since last psychological treatment (including “never”)

- Treatment provider (as ordinal scale)

- Comorbid / previously diagnosed depression

- Comorbid / previously diagnosed Anxiety

- Comorbid / previously diagnosed Substance abuse or other addiction (merge level 2 and 3, see codebook)

- Comorbid / previously diagnosed Schizophrenia

- Comorbid / previously diagnosed Eating Disorder

- Relationship / Family problems

- Life Satisfaction (Mean, Satisfaction with Life Scale)

- Neuroticism (NEO-FFI)

- Extraversion (NEO-FFI)

- Openness (NEO-FFI)

- Agreeableness (NEO-FFI)

- Conscientiousness (NEO-FFI)

The following statistical aspects will be included as predictors:

- Total number of imputed data points per individual: Due to the imputation process, higher proportions of missing data may inflate estimates of temporal stability

RQ2: LASSO predictive regression

To test which person-specific attributes or statistical factors predict temporal network stability, we use predictive regression with least absolute shrinkage and selection operator (LASSO, Tibshirani (1996), McNeish (2015)) regularization. In predictive regression, a model is optimized for it's ability to predict the outcome variable in a novel sample (Westfall & Yarkoni, 2016). This is in contrast to the more commonly used explanatory regression, in which model parameters are optimized to explain maximum variance in the observed sample, which can lead to overfitting and poor out-of sample utility. LASSO regularization shrinks small beta coefficients to zero, effectively excluding less relevant predictors from the model. The amount of shrinkage applied, determined by the tuning parameter lambda, is optimized for predictive accuracy using K-fold cross-validation. Using LASSO regularization, we can explore a wide range of potential predictors while selecting a parsimonious model without relying om arbitrary significance cut-offs or prior theory for variable selection. Data will be spit at random into a training set (80% of cases) and test set (20% of cases), fixing the random number generator at 1821 for reproducibility of the analysis. The training set will be used for model estimation using K-fold cross validation with 5 folds. Missing data (e.g. due to network model non-convergence or missing baseline variables) will be handled using listwise deletion. The model with least prediction error during cross validation will be tested for out-of sample predictive accuracy in the test set. If predictive accuracy is satisfactory (root mean squared error <.20, meaning predicted correlation coefficients are within +/- .10 of original correlation coefficient estimates , the resulting model will be fit on the complete data set to extract unbiased parameter estimates and estimated variance explained.

Results

Discussion

Implications.

- relate to stationarity assumption
- relate to idea of monitoring change, critical slowing down, phase transitions, ROM
- implications for feasibility of this approach for research and applications, ie power, study designs
- propose ideas for further research on this

Implications of high stability.

- assumption of stationarity realistic?

Implications of low stability.

- underpowered? (in)stability of estimates?
- warrants caution regarding the inferences we draw (momentary impression of item correlations in certain period of time vs. stable process which extends beyond this period and could inform interventions (TODO: related to predictive value of networks?)
- measurement error assumption: are we just capturing noise? (unlikely but still an issue to think about)
- is extending measurement period instead of increasing frequency a good solution, eg. with planned missingness, or is non-stationarity and low replicability too much of concern?

Limitations.

- ACCURACY OF ESTIMATES UNKNOWN: Uncertainty around parameter estimates is unknown. We cannot know whether dissimilarity of networks is due to unreliable estimates (too low power, measurement error) or because underlying data-generating process changes (non-stationarity, important variables missing from network)
- limitations of this data set for this research q (e.g. sample size / power)
- constraints to generalizability (population, designs, time frame, estimation method)
- We focus on similarity of global network structure, but there are many more ways to descibe and interpret networks which may or may not be relevant for replicability: network comparison test, predictive networks models, sensitivity and specificity of recovered edges if true network (assumed to be) known
- Kalman imputation assumes MCAR, but likely there is some bias.
- Important assumption made by (idiographic) network models is that constructs were measured without error. Little published research on this, but eg. SchreuderEtAl2020 assessed participants interpretation of EMA items over the course of 6 months and concluded interpretation was consistent. Changes in item interpretation also known as measurement invariance or response shift bias.

Conclusion.

References

Beck, Emorie D., & Jackson, J. J. (2020). Consistency and change in idiographic personality: A longitudinal ESM network study. *Journal of Personality and Social Psychology*, *118*(5), 1080–1100. https://doi.org/10.1037/pspp0000249

Beck, Emorie D., & Jackson, J. J. (2021). Idiographic personality coherence: A quasi experimental longitudinal ESM study. *European Journal of Personality*, 08902070211017746. https://doi.org/10.1177/08902070211017746

Borkulo, C. van, Bork, R. van, Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R., … Waldorp, L. (2017). *Comparing network structures on three aspects: A permutation test.*

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Burger, J., Veen, D. C. van der, Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine*, *18*(1), 99. https://doi.org/10.1186/s12916-020-01558-1

Caviglia, G., & Coleman, N. (2016). Idiographic network visualizations. *Leonardo*, *49*(5), 447–447. https://doi.org/10.1162/LEON_a_01267

Cramer, A. O. J., Waldorp, L. J., Maas, H. L. J. van der, & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–150. https://doi.org/10.1017/S0140525X09991567

Curtiss, J., Ito, M., Takebayashi, Y., & Hofmann, S. G. (n.d.). *Longitudinal network stability of the functional impairment of anxiety and depression*. 10.

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3

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

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). 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

Frumkin, M. R., Piccirillo, M. L., Beck, E. D., Grossman, J. T., & Rodebaugh, T. L. (2020). Feasibility and utility of idiographic models in the clinic: A pilot study. *Psychotherapy Research*, *0*(0), 1–15. https://doi.org/10.1080/10503307.2020.1805133

Ge, F., Yuan, M., Li, Y., Zhang, J., & Zhang, W. (2019). Changes in the network structure of posttraumatic stress disorder symptoms at different time points among youth survivors: A network analysis. *Journal of Affective Disorders*, *259*, 288–295. https://doi.org/10.1016/j.jad.2019.08.065

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hoffart, A., & Johnson, S. U. (2020). Within-person networks of clinical features of social anxiety disorder during cognitive and interpersonal therapy. *Journal of Anxiety Disorders*, *76*, 102312. https://doi.org/10.1016/j.janxdis.2020.102312

Howe, E., Bosley, H. G., & Fisher, A. J. (2020). Idiographic network analysis of discrete mood states prior to treatment. *Counselling and Psychotherapy*

Research, *20*(3), 470–478. https://doi.org/https://doi.org/10.1002/capr.12295

Kuranova, A., Booij, S. H., Menne-Lothmann, C., Decoster, J., Winkel, R. van, Delespaul, P., … al., et. (2020). Measuring resilience prospectively as the speed of affect recovery in daily life: A complex systems perspective on mental health. *BMC Medicine*, *18*(1), 36. https://doi.org/10.1186/s12916-020-1500-9

Mansueto, A. C., Wiers, R., van Weert, J., Schouten, B. C., & Epskamp, S. (2020). *Investigating the Feasibility of Idiographic Network Models*. Retrieved from https://osf.io/hgcz6

McNeish, D. M. (2015). Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965

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

Scholten, S., Lischetzke, T., & Glombiewski, J. (2020). *Toward data-based case conceptualization: A functional analysis approach with ecological momentary assessment*. PsyArXiv. https://doi.org/10.31234/osf.io/prg7n

Segal, A., Wald, I., Lubin, G., Fruchter, E., Ginat, K., Ben Yehuda, A., … Bar-Haim, Y. (2020). Changes in the dynamic network structure of PTSD symptoms pre-to-post combat. *Psychological Medicine*, *50*(5), 746–753. https://doi.org/10.1017/S0033291719000539

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611. https://doi.org/10.2307/2333709

Simms, L. J., Goldberg, L. R., Roberts, J. E., Watson, D., Welte, J., & Rotterman, J. H. (2011). Computerized adaptive assessment of personality disorder: Introducing the CAT–PD project. *Journal of Personality Assessment*, *93*(4), 380–389. https://doi.org/10.1080/00223891.2011.577475

Stockert, S. H. H. von, Fried, E. I., Armour, C., & Pietrzak, R. H. (2018). Evaluating the stability of DSM-5 PTSD symptom network structure in a national sample of u.s. Military veterans. *Journal of Affective Disorders*, *229*, 63–68. https://doi.org/10.1016/j.jad.2017.12.043

Thonon, B., Van Aubel, E., Lafit, G., Della Libera, C., & Larøi, F. (2020). Idiographic analyses of motivation and related processes in participants with schizophrenia following a therapeutic intervention for negative symptoms. *BMC Psychiatry*, *20*(1), 464. https://doi.org/10.1186/s12888-020-02824-5

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288. Retrieved from https://www.jstor.org/stable/2346178

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLOS ONE*, *11*(3), e0152719. https://doi.org/10.1371/journal.pone.0152719

Wichers, M., Groot, P. C., & Psychosystems, E. G., ESM Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*(2), 114–116. https://doi.org/10.1159/000441458

Wright, Aidan G. C., Beltz, A. M., Gates, K. M., Molenaar, P. C. M., & Simms, L. J. (2015). Examining the dynamic structure of daily internalizing and externalizing behavior at multiple levels of analysis. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01914

Wright, Aidan G. C., Hopwood, C. J., & Simms, L. J. (2015). Daily interpersonal and affective dynamics in personality disorder. *Journal of Personality Disorders*, *29*(4), 503–525. https://doi.org/10.1521/pedi.2015.29.4.503

Wright, Aidan G. C., & Simms, L. J. (n.d.). *Stability and fluctuation of personality disorder features in daily life*. 16.

Yazici, B., & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, *77*(2), 175–183. https://doi.org/10.1080/10629360600678310