Preregistration

Detecting Idiographic Personality Change

Emorie D Beck[1], Joshua J Jackson[1]

[1] Washington University in St. Louis

09. January 2019

Study Information

Title	Detecting Idiographic Personality Change

Research questions

1. Can we detect individual-level changes in multivariate time series?
2. What proportion of individuals show idiographic change over the course of 1 year?
3. Does change correspond to self-reported life events?

Hypotheses A small proportion of individuals will show change points over the course of the year. Among those that do, an above chance proportion will have change points that correspond to a life event.

Sampling Plan

Data in this study come from Fraley NSF 0443783 longitudinal study and were collected in 2004-2005. The study used online web assessments of undergraduates in romantic relationships to assess attachment, personality, and relationship variables over different time periods. In the proposed study below, we will use the sample who responded to one survey per week for one year.

Existing data	**Registration prior to analysis of the data**. The second author helped to collect the data, and the first author accessed the data two years ago for an unrelated project. No analyses were performed.
Explanation of existing data	Because the project involves coding of a permutation-based approach for which no R packages or functions currently exist, we had to create these. All code and scripts were written using simulated data to keep the authors blind.
Data collection procedures	Participants were 400 undergraduates involved in romantic relationships and collected at the Univeristiy of Illinois, Urbana-Champagne. Participants completed weekly surveys over the course of one year. Participants were offered $150 in compensation for their time.
Sample size	N = 400 participants completed weekly surveys over the course of one year. All units with at least 25 observatins will be included in this study.
Sample size rationale	Data were pre-collected, so the total sample size is constrained by the number of participants recruited in the intial study.
Stopping rule	The original planned sample was 400 participants, which also acted as the stopping rule.

Variables

Each week, data on attachment style in different relationship contexts, the Big 5 (the TIPI), relationship satisfaction, and relationship / life events were collected. In addition, each wave included one of the five following surveys at random: depressive symptoms (CESD), ECR-R attachment, Big Five personality traits (BFI), attachment features and funtions (WHOTO), physical health symptoms (PILL).

Manipulated variables	There are no manipulated variables.
Measured variables	The Ten Item Personality Scale: 1. Extraverted, enthusiastic. (E) 2. Critical, quarrelsome. (A) (-) 3. Dependable, self-disciplined. (C) 4. Anxious, easily upset. (N) 5. Open to new experiences, complex. (O) 6. Reserved, quiet. (E) (-) 7. Sympathetic, warm. (A) 8. Disorganized, careless. (C) (-) 9. Calm, emotionally stable. (N) (-) 10. Conventional, uncreative. (O) (-) Life / Relationship Events: • relevents03 = ’someone in my family passed away’; • relevents04 = ’my partner and I were separated this week due to travel’; • relevents06 = ’we found out we were pregnant’; • relevents08 = ’my partner and I got engaged’; • relevents09 = ’my partner and I moved in together’; • relevents10 = ’my partner and I got married’; • relevents11 = ’my partner and I broke up’;
Indices	We will use Big 5 composite traits. Items will be composited at each measurement point for each person for each trait. Items noted above with “(-)” will be reverse coded. We will create a variable that indexes whether participants experienced any of the events above and which week they occurred. If sample size allows, we will test both this broader composite and specific life events.

Design Plan

Study type	**Observational Study**. Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.
Blinding	No blinding is involved in this study.
Study design	This is a longitudinal study in which participants were asked to complete surveys once a week for one year.
Randomization	Participants received randomy received one of five surveys each week, in addition to the normal questions

Analysis Plan

Analyses will be based a permutation-based kernal change point analysis for detecting correlational changes in multivariate time series (Cabrieto, Tuerlinckx, Kuppens, Hunyadi, & Ceulemans, 2018).

The kernal change point project involves using moving-window correlations across the raw time series as well as shuffled permutations of the time series. We will use a window size of 10 and run 1000 permutations. Because the algorithm searches for the optimal number of change points, we will not specify this advance but will constrain the maximum number of change points to three, as this seems to the authors to be the maximum number of plausible changes over one year of weekly assessments.

Variance Test: tests whether the variance in moving window correlation similarity across observation within a window exceeds chance variability in permutations (shuffles) of the data set. This indicates whether there are any change points in the data. If there are, the variance in moving window correlations should be greater in the raw data than in the permuted data sets.

Variance Drop Test: Because the KCP solution by construction improves with more change points (because variance will be decreased when window sizes are smaller and more homogeneous), the variance drop test tests whether the drop in variance by adding an additional change point in the raw data exceeds the drop in variance in the permuted data sets in at least 97.5% of cases.

Steps: 1. Calculate moving window correlations.

2. Calculate Gaussian similarity of correlations within all possible combinations of windows.

3. Repeat steps 1 and 2 for 1000 permutations of the raw data.

4. Perform the Variance test using the Gaussian similarity when K = 0.

5. Perform the Variance drop test using the max drop in average Variance of Gaussian similarity ($\hat{R}$).

6. Make a conclusion on whether there is a change point using a combination of 4 and 5.

Statistical models

Gaussian Similarity Measure:

$$Gk(\mathbf{R}_i, \mathbf{R}_j) = exp\Big(\frac{-||\mathbf{R}_i - \mathbf{R}_j||^2}{2h_R^2}\Big)$$

where, h_R is obtained by computing the median Euclidean distances between all R_i's.

Calculating variance within a window:

$$\hat{V}_{p,\tau_1,\tau_2,\ldots,\tau_K} = (\tau_p - \tau_{p-1}) - \frac{1}{\tau_p - \tau_{p-1}} \sum_{i=\tau_{p-1}+1}^{\tau_p} \sum_{j=\tau_{p-1}+1}^{\tau_p} Gk(\mathbf{R}_i, \mathbf{R}_j)$$

Calculating average variance within a window:

$$\hat{R}(\tau_1, \tau_2, \ldots, \tau_K) = \frac{1}{n} \sum_{p=1}^{K+1} \hat{V}_p, \tau_1, \tau_2, \ldots, \tau_K$$

Choosing the Change Points

$$\tau_1, \tau_2, \ldots, \tau_K = arg\ min\ \hat{R}(\tau_1, \tau_2, \ldots, \tau_K) = arg\ min \frac{1}{n} \sum_{p=1}^{K+1} \hat{V}_p, \tau_1, \tau_2, \ldots, \tau_K$$

Calculating the Variance Test:

$$P_{variancetest} = \frac{\#(\hat{R}_{min,K=0} > \hat{R}_{min,K=0,perm})}{B}$$

Calculating the Variance Drop Test:

$$P_{variancedroptest} = \frac{\#(max\ variance\ drop_{perm} > max\ variance\ drop)}{B}$$

where $max\ variance\ drop = \hat{R}_{min,K} - \hat{R}_{min,K-1}$

Transformations Items noted above with a "(-)" will be reverse coded.

Follow-up analyses

Inference criteria We are using a permutation based approach that involves two tests (the variance drop test and the variance test). As such, we will use a Bonferroni correction ($\frac{\alpha}{2}$), such that the raw observed value should be greater than 97.5% of the permuted values (variance and variance drop).

Data exclusion Participants with fewer than 25 responses will be excluded because of inadequate power to detect change.

Missing data Missing data will left as missing to prevent correlating across uneven time windows.

Exploratory analyses (optional) As a follow up, we will test whether identified change points coincide with life event experiences.

It is possible that only a small subset of the population will show reliable change points. Of those that do, it is likely that only a subset will report one of the 7 life events we collected. Of those that do, we will report the proportion of individuals whose life event experiences fall +/- 2 weeks from the empirically derived change points.

Analysis scripts (optional) Attached with this preregistration is an R script with a series of functions written to succinctly run the KCP permutation procedure given a data frame of multivariate time series data.

Other

Other (Optional)

References

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., & Ceulemans, E. (2018). Testing for the presence of correlation changes in a multivariate time series: A permutation based approach. **Scientific reports, 8**(1), 769.