Preregistration

Migration Experience Trajectories: A Three Mode Principle Component Analysis

Jannis Kreienkamp[1], Kai Epstude[1], Rei Tendeiro-Monden[2], Maximilian Agostini[1], Peter de Jonge[1], Laura F. Bringmann[1]

[1] University of Groningen, Department of Psychology

[2] Hiroshima University, Graduate School of Advanced Science and Engineering

14. May 2022

Study Information

Title	Migration Experience Trajectories: A Three Mode Principle Component Analysis
Description	Two important developments, within the psychological sciences in general and in migration research in particular, have been (1) a shift towards more comprehensive assessment of human experiences and (2) a growing focus on longitudinal real-life data. Within the migration research field, recent reviews have been pushing for models that take into account affects, behaviors, cognitions, and desires when it comes to migrant adaptations (Kreienkamp et al., 2022b; Ward and Szabo, 2019). At the same time prominent figures in the field have called for studies that collect longitudinal (Pettigrew, 1998, 2008a,b; Pettigrew and Tropp, 2011) and real-life experience-sampling data outside the lab (MacInnis and Page-Gould, 2015; McKeown and Dixon, 2017; Dixon et al., 2005).

In the past, such data collections were often unfeasible because they were either physically impractical or too expensive. However, recent technological developments now allow us to easily collect large amounts of experience sampling data on mobile devices (e.g., Keil et al., 2020) or using web-based applications (Arslan et al., 2020). And while generally speaking analytical methods for such, more complex, data have become more readily available (e.g., O'Donnell et al., 2021), it remains unclear how we should identify key developmental patterns — especially across multiple variables at the same time (i.e., multiple aspects). In essence, the new extensive longitudinal data comes with new forms of heterogeneity when jointly considering differences between people, over time, and across variables. Yet, past analytical advances have almost exclusively pushed for stationary lagged regression models[1] or basic trajectory models[2]. And while these two approaches are important for inferential model testing, both approaches ignore the more fundamental task of describing and unraveling the developmental data patterns.
And while these two approaches are important for inferential model testing, we still miss methods for the more fundamental task of describing and unraveling the developmental data patterns.

Such descriptive methods are not merely important for hypothesis generation, but also for identifying and understanding adaptive and maladaptive patterns. We we need to identify which variables are most important in the adaptation of migrants, we need methods to identify clusters of similar individuals and developments, and we need to assess whether such clusters differ in key adaptation markers (including, well-being, intergroup anxiety, outgroup trust, or societal participation).

We recently conducted three experience sampling studies, following the migration experiences of recent migrants to the Netherlands. In this manuscript, we aim to assess the utility of an analytical methods that has recently received increasing attention for simultaneously decomposing/considering person-, variable-, and time heterogeneity. The proposed three mode principle component analysis is of particular interest because its approach mirrors the famous data box of participants, variables, and measurement occasions (Catell's data box) for understanding

[1] meaning models that assume stable means and variances over time; e.g., vector autoregression models, dynamic structural equation models, autoregressive integrated moving average models, and cross-lagged panel analyses

[2] e.g., mixed effects models, spline regression models, and latent growth curve modeling.

psychological data.

Problem: HOWEVER, ...

- unclear how do deal with this more complex data (many variables, persons, and time points).
  - many different variables could be important.
  - how to consider them jointly (e.g., VAR, DSEM, ARIMA, cross-lagged panel analysis for testing model predictions, limited to specified lag and mean stationarity; Trajectory focused: mixed effects model, spline regressions, latent growth curve modeling, limited: often univariate outcome). At the same time, analytical methods for such more complex data have become more readily available, making the analyses more approachable (O'Donnell et al., 2021).
- heterogeneity between people, over time, and across variables.
- **unclear how to identify core/important developments.** Especially across multiple variables at he same time (i.e., multiple aspects).
- unclear which variables, time scales, and methods are useful in practice.

Solution:

- 3MPCA uses the data cube

little data has thus far investigated the development of migration experiences and no research has assessed the co-development of multiple experience aspects. Yet, understanding how people differ in their migration trajectories, can be crucial in understanding adaptive and maladaptive patterns. Simultaneously clustering the person-, experience aspect-, and time level of migration experiences, using a three mode PCA, may allow to identify clusters of similar developments and whether these clusters differ in key adaptation markers (including, well-being, anxiety, trust, or societal inclusion).

Hypotheses We do not have hypotheses in the traditional sense. Our analysis plan is based on the aim of describing and understanding a complex data set of extensive longitudinal data. We propose that the dimension reduction procedure we employ will identify meaningful patterns and developments within the data, which are useful to migration researchers and practitioners.

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 participant blinding is involved in this study.
Study design	All three studies used an extensive longitudinal design. Using a daily diary format, for at least 30 days participants received a short survey twice per day (at around 12pm and 7pm). We additionally included a longer pre- and post measurement survey the days before and after the extensive longitudinal data collection.
Randomization	No randomization is involved in this study.

Sampling Plan

Existing data	**Registration following analysis of the data**: As of the date of submission, you have accessed and analyzed some of the data relevant to the research plan. This includes preliminary analysis of variables, calculation of descriptive statistics, and observation of data distributions. Please see cos.io/prereg for more information.
Explanation of existing data	The data was collected as part of a larger collaboration on daily intergroup relations. A sub-sample of variables has recently been accessed by the research team for an unrelated analysis (Kreienkamp et al., 2022a). Additionally, several of the variables were prepared for a graphical presentation of the dataset. Thus far, none of the proposed analyses have been conducted and none of the previous analyses have been related to dimension reductions or longitudinal trajectories.

Data collection procedures

For all three studies, the data was collected in a three-step procedure:

1. Entry Survey: A pre-measurement questionnaire (appr. 25 minutes) including demographic information, and relations to the Dutch majority (payment: 2 Euros).
2. Experience Recaps: At least 30 days of short reflection surveys (appr. 3—5 minutes) on intergroup interactions twice a day (payment: 1 Euro per Recap; up to 2 Euros per day).
3. Conclusion Survey: On the last day, we conclude with a post-measurement questionnaire (appr. 25 minutes) with some questions on habits and reflections on the study (payment: 2 Euros).

For the third study, participants had the option of continuing the study for an unspecified amount of time. After the initial 30 day duration, participants were offered the possibility to continue participating in the study either with payment if daily diary measures were missed during the initial study phase or without payment after a total of 60 daily diary measurements were completed. After the initial 30-day period, participants receive automated feedback visualizing the development of their own well-being, attitudes, and motive responses as an additional initiative and to give participants access to their own data and to compensate study participation.

Sample size

For Study 1 (initial preliminary study) our target sample size is a sum of 1,000 daily diary measurements. As we expected a completion rate of around 80% we aimed to recruite 20 participants (20 participants X 50 measurements). We ultimately over-sampled slightly to compensate for potential drop outs given the length and intensity of the study. We then used several key relationships (relevant to the broader research team) to estimate sufficient power for a range of different analyses (i.e., using power simulations).

For Studies 2 and 3, our target sample size is a sum of 4,000 daily diary measurements. With 100% completion rate that would be archived with 67 participants (60 daily diary responses each). Given that we expect some incomplete daily diary measurements, we aimed to recruit 80 participants. For both studies we again over-sampled slightly to account for expected drop outs within each target population.

Sample size rationale

The targeted sample size depended on a combination of different factors. Different analyses were planned as part of the collaboration and budgeting was a practical constraint. Some analyses were planned based on (1) the pre- to post measurements, (2) the dynamic developments over the daily diary measurements, or (3) the contemporary effects within the daily diary measurements.

Power considerations of mixed effects model such as with extensive longitudinal data are difficult to estimate because of the complex covariance structures. Simulation studies based on the first sample within this project indicated that with well-distributed scales, and small to medium effect sizes, 70-80 participants with at least seven daily diary measurements and a simple pre–post survey were sufficiently powerful (power = .8, alpha = .05) to answer most of the key research questions of the collaboration.

The ultimate sampling procedure decision was made as a practical balancing of the number of participants and the number of measurements provided by each participants.

Stopping rule

Participants were recruited until the targeted participant number finish the pre-measurement. Invitations to complete additional daily diary measurements (in Study 3) were extended until participants chose to leave the study or at the most until two months (i.e., 64 days) after the initial entry survey (i.e., from the pre measurement survey).

Variables

Manipulated variables

Not applicable given that the study design is observational.

Measured variables

Given the circumstance that we utilize three independent experience sampling studies for our analyses, the variables are at times slightly different between studies. Additionally, given that the analysis we aim to undertake (i.e., 3MPCA) uses a large number of variables for the multiple imputations, dimension reduction, and correlational analyses, we will not list all items individually. Instead, we provide full variable information is available in the following data sheets: 'VariableSelectionDaily.xlsx' and 'VariableSelectionPrePost.xlsx'.

Key variables

Most of the variables are identical across all studies, however some differ slightly in content or wording. However, for each study we aimed to collect the following types of variables. We collected substantially more motivational and affecive variables, as they have been understudied in the past.

1. Motivational variables
    a. core motive fulfillment (i.e., "*During your interaction with -X- (/this morning) your goal (-TEXT-) was fulfilled.*")
    b. goal importance ratings of 10 individual goals (e.g., "*career goals*", "*health / fitness goals*")
    c. self determination theory needs (i.e., autonomy, relatedness, competence)
2. Affective states
    a. experienced well-being (happiness, energy)
    b. general emotional state scale (e.g., "*How do you feel right now? not angry at all to very angry*")
3. Behavioral self-reports (Studies 2 and 3)
    a. pro-social behavior (e.g., "*Made demeaning, rude or derogatory remarks about someone.*")
    b. anti-social behavior (e.g., "*I was there to listen to someone's problems.*")
4. Cognitive measures
    a. outgroup attitude ("*After the interaction, how favorably do you feel towards ... / At the moment, how favorably do you feel towards ... the Dutch.*")
    b. Allport's contact condition perceptions (i.e., equal status, shared goal, cooperative, voluntary)
    c. interaction quality ratings (e.g., "*Overall, the interaction was... Unpleasant to Pleasant*")

Interpretation correlation variables

To interpret the components identified by the 3MPCA, we would like to correlate the component scores with a range of person- or time point specific variables. While the full list of variables is available in the accompanying data sheets (see 'VariableSelectionDaily.xlsx' and 'VariableSelectionPrePost.xlsx'), we provide a few examples of the targeted categories below.

1. demographic variables
   a. age
   b. gender
   c. ...
2. societal participation
   a. language comprehension and -use
   b. work inclusion
   c. ...
3. adaptation indicators
   a. social identification
   b. intergroup anxiety
   c. ...
4. everyday life obstacles
   a. discrimination experiences
   b. negative life events
   c. ...
5. individual differences
   a. big five personality scores
   b. attachment style
   c. ...

Auxiliary variables for multiple imputation

For the missing data imputation we follow the procedure reported by Monden et al. (2015), and include any pre-, post-, or daily dairy variables that are significantly ($p < .01$) and meaningfully ($r > .3$) correlated with (1) the key variables or (2) missingness on the key variables. For an overview of the variables that qualify as auxiliary variables see the enclosed data sheets (see 'VariableSelectionDaily.xlsx' and 'VariableSelectionPrePost.xlsx'). Additional information about the multiple imputation procedure is also available at Statistical models 2.

We also provide an exemplary codebook (using Study 3 as an example; see 'Codebook_AOT-M_ItemsPerSection.xlsx'). The full survey files will be available as part of the main OSF repository connected to this preregistration.

Indices

1. Mean Allport's conditions. We create a mean-averaged index of Allport's conditions in response to past findings indicating that the conditions are best conceptualized jointly and as functioning together rather than as fully independent factors (Pettigrew and Tropp, 2006, p. 766). Similar to past studies we thus hope to build a global indicator (e.g., see Pettigrew and Tropp, 2006). As with other indices we will ensure that the individual items indeed relate to a common latent construct and are meaningfully combine in an index. If this is not possible we will create sub-indices and/or assess the impact of the conditions separately.
   a. The interaction with [name interaction partner] was on equal footing (same status)
   b. [name interaction partner] shared your goal ([free-text entry interaction key need])
   c. The interaction with [name interaction partner] was cooperative
   d. The interaction with [name interaction partner] was voluntary
2. Mean belongingness during intergroup contact
   a. I shared information about myself.
   b. [name interaction partner] shared information about themselves.
3. Mean alternative interaction quality definition ("Overall, the interaction with [name interaction partner] was . . . ")
   a. Unpleasant to Pleasant
   b. Superficial to Meaningful
   c. Ineffective to Effective
   d. Unimportant to Important

4. Mood Subscales (MDMQ)
   a. alertness
   b. calmness
   c. valence
5. Mean anti-social behaviors
   a. Put someone down
   b. Show little attention in someones opinion
   c. Demeaning remarks
   d. Inappropriately addressing someone
   e. Ignored or excluded someone
   f. Doubt someones judgement
   g. Unwanted attempts of personal matters
6. Mean pro-social behaviors
   a. Listen to someones problems
   b. Cheer someone up
   c. Help someone get things done
   d. Help someone with responsibilities

Analysis Plan

Statistical models In our analysis, we follow the procedure outlined by Monden et al. (2015)

1. **Sample Selection.** Our sample selection consists of three main steps. We (1) aggregate the key variables over time to archive a reasonably interpretable number of time points and the remove a first proportion of missing data. We then address the selection of time point and participants. Given that multiple imputation procedures even work well with large proportions of missing data [assuming missingness at random; Madley-Dowd et al. (2019)], we decided on a general criterion of less than 45% missingness to balance sample size retention and bias in the multiple imputation model. Thus, we then (2) select the time points for which we have responses from 55% of the participants and (3) select participants who have responded to at least 55% of the surveys across the selected time range.

1. Time point aggregation: Given that little data is available on the meaningful time scales of the selected psychological variables, we chose to choose the appropriate time scales using variance decomposition (e.g., see Ram et al., 2014). This is to say that we create multi-level unconditional means models (without predictors) that include possible nested time scales as levels. For example, assess the explained variances for measurements nested in days, nested in weeks, nested in months.

   Potentially nested measurements:
   1 - 2 - 4 - 8 - 16 - 32 - 64
   1 - 2 - 6 - 12 - 24 - 48
   1 - 2 - 10 - 30 - 60
   1 - 2 - 14 - 28 - 56
   Weird but possible (uneven):
   1 - 3 - 6 - 12 - 24 - 48
   1 - 5 - 10 - 30 - 60
   1 - 7 - 14 - 28 - 56

   We then chose the ideal time scale for each of the key variables.

   Question: Should the aggregation happen after the multiple imputation? The current procedure would reduce missing data but the aggregated values would be based on differing amounts of data (depending on the amount of missingness). Yet again, a daily or weekly summary this is probably still more accurate then a single later recall response.

2. **Multiple Imputation.** We then create 20 imputed datasets to perform the analyses on. As outlined by Monden et al. (2015), we will use the key variables themself as well as auxiliary variables to impute missing values. Auxiliary variables are any pre-, post-, or daily dairy variables that are significantly ($p < .01$) and meaningfully ($r > .3$) correlated with (1) the key variables or (2) missingness on the key variables.

3. **Three-way ANOVA.** We assess the percentages of explained variance for the person-, variable-, and time aspects — which offers an indication of whether

a 3MPCA is useful for the dataset. Based on the grand mean centered variables we conduct a fixed-effects three-way ANOVA of the person-, variable-, and time modes as well as their various interactions. We are particularly interested in the highest order interaction term Person * Variable * Time + error, as a large amount of variance in this effect would speak towards the interdependence of the three modes.

Question: What is the DV of the three-way ANOVA?

4. **Data preprocessing.** For the three mode PCA, we center (across participants but within time point) and to normalize (within variable but across time points) all key variables. The between-person centering, ensures that all variations are around the mean trend (which is removed by the centering). The normalization within variable are important for ensuring equal variances across variables, which ensures equal weighting of the variables in the 3MPCA. For methodological detail see Transformations.
5. **Selection of 3MPCA model complexity** We then run the 3MPCA. We use a generalized scree plot to select the appropriate number of components for each of the three modes (i.e., person, variable, time). To test the stability of the solution results are compared across the 20 imputed datasets.

Question: Additionally split-half procedures? How does this work is this the same as split-half reliability?

6. **3MPCA model fitting.** For the selected complexities we then extract simple component structures (i.e., for each mode individually) as well as the array core (the full 3MPCA array, which holds the weights of all combinations of the three modes). A Joint Orthomax orthogonal rotation and standard weights (when necessary) are used to extract human interpretable component scores.

This procedure is done on all 20 imputed data sets.

> **Question:** What exactly does the core array entail? In the @Monden2015 article it says: "the core-array, which describes the patterns of interaction between the symptom- and time-components in each person-component.". Additionally, in a more mathematical piece by Jorge it reads: "a three-way array called the core array, which holds the weights for the joint impact of any triple of components (one from each mode)." Does this mean that this is basically an array of loading scores in the lower dimensional space (similar to fitted values in regressions)? Or does it store the weights (i.e., interactions scores) that combines the loading scores of all three dimension reductions (not sure I express this right :-D)? I think this comes down to: What is the interpretation of a core element?

7. **Generalized Procrustes rotation.** In order to combine the three individual component structure from the 20 data sets, we use a generalized Procrustes rotation — calculating the average of each compontent and core array.
8. **Explained variance.** As fit indices we then calculate (1) fit percentage of the estimated array for each imputed data set (i.e., explained variance around the removed general trend) as well as (2) an *overall* fit percentage of the estimated array (i.e., explained variance including the general trend). Both these fit metrices should give an indication of how much of the original variance the reduced component structure was able to capture.
9. **Component interpretation** To embed and interpret the resulting components (especially the person components), we calculate correlations of the component scores with key adaptation markers (including, well-being, anxiety, trust, or societal inclusion). If components remain unclear still we can additionally perform k-means clustering on the lower dimensional space to identify clearer groups which can be compared on the adaptation markers if necessary.

Transformations

For three mode PCA it is custom to center (across participants but within time point) and to normalize (within variable but across time points). This is done to create a meaningful zero value for each time point (to compare participants) and a variance of 1 across all time points of a variable (removing any overall trends).

We provide the main equations of the procedure below, where individual i ($i =$

1, ..., I) reported variable k ($k = 1,\ ...,\ K$) at time point t ($t = 1,\ ...,\ T$), so that I=number of persons, K=number of items, T=number of time points are within the dataset. For further information as well as an example illustration see Supplemental Material 3 in Monden et al. (2015).

$$\overline{x}_{kt} = \sum_{i=1}^{I} x_{ikt} \tag{1}$$

$$\sigma_k = \sqrt{\sum_{i=1}^{I} \sum_{k=1}^{K} \frac{(x_{ikt} - \overline{x}_{kt})^2}{I * T}} \tag{2}$$

$$z_{ikt} = \frac{x_{ikt} - \overline{x}_{kt}}{\sigma_k} \tag{3}$$

Inference criteria

Many of the proposed procedures are more descriptive rather then inferential. Importantly, we will use amounts of variances explained as our main fit indices. If we perform inferential analyses, we will use the standard p<.05 criteria for determining whether the correlation and regression coefficients are statistically significant. We will place particular emphasis on effect sizes in our interpretations of the results.

Data exclusion

No checks will be performed to determine eligibility for inclusion besides verification that each subject answered each of the variables of interest for a given analysis. Outliers will generally be included in analyses, however we will use sensitivity analyses to assess the robustness of the results to outliers.

Missing data

The sample selection based on proportions of missing data and the multiple imputation procedures are described in Statistical models.

Exploratory analyses (optional)

A recent alternative for trajectory dimension reduction has been the use of LSTM auto encoders. The encoder decoder classification procedure has the advantage of being extremely flexible in the use of the data set and could be interpreted in a similar way as the 3MPCA (see Statistical models 9).

Other

Other (Optional)	Not applicable.

References

Arslan, R. C., Walther, M. P., and Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, 52(1):376–387.

Dixon, J., Durrheim, K., and Tredoux, C. (2005). Beyond the optimal contact strategy: A reality check for the contact hypothesis. *American Psychologist*, 60(7):697–711.

Keil, T. F., Koschate, M., and Levine, M. (2020). Contact Logger: Measuring everyday intergroup contact experiences in near-time. *Behavior Research Methods*, 52(4):1568–1586.

Kreienkamp, J., Agostini, M., Bringmann, L. F., de Jonge, P., and Epstude, K. (2022a). Psychological Needs During Intergroup Contact: Three Experience Sampling Studies [Manuscript submitted for publication].

Kreienkamp, J., Bringmann, L. F., Engler, R. F., de Jonge, P., and Epstude, K. (2022b). The Migration Experience: A Conceptual Framework and Systematic Review of Psychological Acculturation [Manuscript submitted for publication].

MacInnis, C. C. and Page-Gould, E. (2015). How can Intergroup Interaction be bad If intergroup contact Is good? Exploring and reconciling an apparent paradox in the science of intergroup relations. *Perspectives on Psychological Science*, 10(3):307–327.

Madley-Dowd, P., Hughes, R., Tilling, K., and Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, 110:63–73.

McKeown, S. and Dixon, J. (2017). The "contact hypothesis": Critical reflections and future directions. *Social and Personality Psychology Compass*, 11(1):1–13.

Monden, R., Wardenaar, K. J., Stegeman, A., Conradi, H. J., and De Jonge, P. (2015). Simultaneous decomposition of depression heterogeneity on the person-, symptom- and time-level: The use of three-mode principal component analysis. *PLoS ONE*, 10(7):1–16.

O'Donnell, A. W., Friehs, M. T., Bracegirdle, C., Zúñiga, C., Watt, S. E., and Barlow, F. K. (2021). Technological and analytical advancements in intergroup contact research. *Journal of Social Issues*, 77(1):171–196.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49:65–85.

Pettigrew, T. F. (2008a). Future directions for intergroup contact theory and research. *International Journal of Intercultural Relations*, 32(3):187–199.

Pettigrew, T. F. (2008b). Reflections on Core Themes in Intergroup Research. In Wagner, U., Tropp, L. R., Finchilescu, G., and Tredoux, C., editors, *Improving Intergroup Relations*, pages 281–303. Blackwell Publishing Ltd., Oxford, UK.

Pettigrew, T. F. and Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5):751–783.

Pettigrew, T. F. and Tropp, L. R. (2011). *When groups meet: The dynamics of intergroup contact*. Psychology Press.

Ram, N., Conroy, D. E., Pincus, A. L., Lorek, A., Rebar, A., Roche, M. J., Coccia, M., Morack, J., Feldman, J., and Gerstorf, D. (2014). Examining the Interplay of Processes Across Multiple Time-Scales: Illustration With the Intraindividual Study of Affect, Health, and Interpersonal Behavior (iSAHIB). *Research in Human Development*, 11(2):142–160.

Ward, C. and Szabo, A. (2019). Affect, Behavior, Cognition, and Development. In Matsumoto, D. and Hwang, H. C., editors, *The Handbook of Culture and Psychology*, pages 640–692. Oxford University Press.