

Preregistration

Migration Experience Trajectories: A Three Mode Principle Component Analysis

Jannis Kreienkamp¹, Kai Epstude¹, Rei Tendeiro-Monden², Maximilian Agostini¹, Peter de Jonge¹, Laura F. Bringmann¹

¹ University of Groningen, Department of Psychology

² Hiroshima University, Graduate School of Advanced Science and Engineering

13. 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., 2022 ; 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 models¹ or basic trajectory models². 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 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

¹meaning 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

²e.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 the 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 any hypotheses in the traditional sense. Our analysis plan is based on
-------------------	---

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. Not relevant, no insight for the current analysis.
Data collection procedures	For all three studies, the data was collected in a three-step procedure: <ol style="list-style-type: none"> 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	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.
--------------------	---

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 two samples 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. Given that this study primarily focuses on the contemporaneous effects within the daily diary measurements, power should be high enough.

Stopping rule	Participants will be recruited until 80 participants finish the pre-measurement. Invitations to complete additional daily diary measurements will be 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	Key variables: <ol style="list-style-type: none"> 1. Whether an interaction happened during the past daytime period <ol style="list-style-type: none"> a. Did you meet a Dutch person this [morning/afternoon]? b. Did you meet a Non-Dutch person [morning/afternoon]? 2. Number of interactions during the past daytime period <ol style="list-style-type: none"> a. How many interactions with Dutch people did you have this [morning/afternoon]? b. How many interactions with Non-Dutch people did you have this [morning/afternoon]? 3. Key need assessment <ol style="list-style-type: none"> a. What was your main goal during the interaction with [name interaction partner]? b. During the interaction with [name interaction partner] your goal [free-text entry interaction key need] was fulfilled. 4. Allport's Conditions (equal status, common goal, collaboration, and structural support)
---------------------------	---

PREREGISTRATION: MIGRATION EXPERIENCE TRAJECTORIES

- 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

5. Interaction Quality

- a. Overall, the interaction with [name interaction partner] was: Unpleasant to Pleasant

6. Outgroup Attitudes

- a. After the interaction, how favorably do you feel towards the Dutch.

Control variables:

1. Self Determination Theory needs (“During the interaction with [name interaction partner]”)
 - a. I was myself.
 - b. I felt competent.
 - c. I shared information about myself.
 - d. [name interaction partner] shared information about themselves.
2. General key need during daytime
 - a. What was your main goal during this [morning/afternoon]?
 - b. During this [morning/afternoon] your goal ([free-text entry daytime key need]) was fulfilled
 - c. How important was your goal ([free-text entry daytime key need]) to you:
3. Student Goals (“Please rate the importance of each goal during the interaction with [name interaction partner]”):
 - a. Social support / social connection goals
 - b. Romantic or sexual relationship goals
 - c. Academic goals
 - d. Career goals
 - e. Financial goals

- f. Health / fitness goals
- g. Leisure / fun goals
- h. Personal improvement / growth goals
- i. Service / helping goals
- j. Spiritual / religious goals

Additional Analyses:

1. Generalization from interaction partner to outgroup
 - a. After the interaction, how favorably do you feel towards [name interaction partner].
 - b. The interaction with [name interaction partner] was representative of the Dutch.
2. 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
3. Potential moderation by the relationship with interaction partner
 - a. How close do you feel to [name interaction partner]?
 - b. How much do you trust [name interaction partner]?
 - c. What is your relationship with [name interaction partner]?
4. Potential moderation by key need characteristics
 - a. [name interaction partner] helped fulfill your goal [free-text entry interaction key need].
 - b. How important was your goal ([free-text interaction key need]) to you?
5. Potential moderation by the interaction characteristics
 - a. How long was the interaction? (in minutes)
 - b. What was the main language during the interaction?
 - c. The interaction was: With a single other / In a group

Full variable information is available in the codebook ('Codebook_AOT-M_ItemsPerSection.xlsx')

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 (?, p. 766). Similar to past studies we thus hope to build a global indicator (e.g., see ?). 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

Analysis Plan

Statistical models

We will use a sequential analysis strategy in line with our proposed hypotheses. Given the nested structure of much of our data we test many of our hypotheses using a multilevel approach, where y_{ti} denotes the response at measurement occasion t ($t = 1, \dots, T_i$; level 1) for individual i ($i = 1, \dots, n$; level 2). It should be noted that

we will follow a hierarchical modeling approach. We follow the common four-step procedure (e.g., ?): (1) Test whether enough variation exists within and between participants to justify a multilevel structure, (2) (sequentially) add key predictors, (3) check whether a random slope explains an adequate amount of variance, and (4) check for autocorrelations and heteroscedasticity. For brevity we will only present the full multilevel regression formulas below.

1. Contact Hypothesis (partially using between participant aggregates to meaningfully include interaction frequency)

a. Correlation:

$$r_{ContactFreq, Attitude} \neq 0 \quad (1)$$

b. Regression:

$$\begin{aligned} \text{Level 1: } Attitude_{ti} = & \beta_{0i} + \beta_{1i} OutgroupInteraction_{ti} + \\ & \beta_{2i} NonOutgroupInteraction_{ti} + e_{ti} \end{aligned}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (2)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

c. Regression:

$$Attitude = ContactFreq \times AverageQual \quad (3)$$

2. Allport's Conditions

a. Regression:

$$\text{Level 1: } Attitude_{ti} = \beta_{0i} + \beta_{1i} AllportConditions_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (4)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

b. Regression:

$$\text{Level 1: } InteractionQuality_{ti} = \beta_{0i} + \beta_{1i} AllportConditions_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (5)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

c. Regression:

$$\begin{aligned} \text{Level 1: } \textit{Attitude}_{ti} &= \beta_{0i} + \beta_{1i} \textit{AllportConditions}_{ti} + \\ &\quad \beta_{2i} \textit{InteractionQuality}_{ti} + e_{ti} \end{aligned}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (6)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

3. Key Need fulfillment

a. Regression:

$$\text{Level 1: } \textit{Attitude}_{ti} = \beta_{0i} + \beta_{1i} \textit{KeyNeedFulfill}_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (7)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

b. Regression:

$$\text{Level 1: } \textit{InteractionQuality}_{ti} = \beta_{0i} + \beta_{1i} \textit{KeyNeedFulfill}_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (8)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

c. Regression:

$$\begin{aligned} \text{Level 1: } \textit{Attitude}_{ti} &= \beta_{0i} + \beta_{1i} \textit{KeyNeedFulfill}_{ti} + \\ &\quad \beta_{2i} \textit{InteractionQuality}_{ti} + e_{ti} \end{aligned}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (9)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

d. Regression:

$$\begin{aligned} \text{Level 1: } \textit{Attitude}_{ti} &= \beta_{0i} + \beta_{1i} \textit{KeyNeedFulfill}_{ti} + \\ &\quad \beta_{2i} \textit{InteractionQuality}_{ti} + \\ &\quad \beta_{3i} \textit{KeyNeedFulfill} * \textit{InteractionQuality}_{ti} + e_{ti} \end{aligned}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

$$\beta_{3i} = \gamma_{30} + u_{3i}$$

$$(10)$$

e. Regression:

$$\text{Level 1: } Attitude_{ti} = \beta_{0i} + \beta_{1i}KeyNeedFulfill_{ti} + \beta_{2i}Autonomy_{ti} + \beta_{3i}Competence_{ti} + \beta_{4i}Relatedness_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

$$\beta_{3i} = \gamma_{30} + u_{3i}$$

$$\beta_{4i} = \gamma_{40} + u_{4i}$$

(11)

4. Comparison with Allport's Conditions

a. Model Comparison:

$$AIC_{KeyNeedModel} < AIC_{AllportModel} \quad (12)$$

b. Regression:

$$\text{Level 1: } Attitude_{ti} = \beta_{0i} + \beta_{1i}KeyNeedFulfill_{ti} + \beta_{2i}AllportConditions_{ti} + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i} \quad (13)$$

$$\beta_{1i} = \gamma_{10} + u_{1i}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}$$

All multilevel assumptions are tested as usual including (e.g., for random slopes model with j within person predictors):

$$\text{Level 1 Variance: } e_{ti} \sim \mathcal{N}(0, \sigma^2) \quad (14)$$

$$\text{Level 2 Variance: } \begin{bmatrix} u_{0i} \\ \vdots \\ u_{ji} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00}^2 & & \\ \vdots & \ddots & \\ \tau_{j0} & \dots & \tau_{jj}^2 \end{bmatrix} \right) \quad (15)$$

Transformations

We will mean-center all level-one predictors within participants (i.e., cluster mean centering). This is relevant for any predictor variable during the interaction (e.g., key need fulfillment, interaction quality, ...). We mean-center within-person predictors, to meaningfully distinguish within and between person variances and because we obtain a meaningful intercept when having a random slope model (e.g., ?).

Inference criteria	We will use the standard $p < .05$ criteria for determining whether the correlation and regression coefficients are statistically significant. However, given the large sample size we will place a 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	Given the (expectedly) large sample size of contemporaneous measurements, we will remove all instances during which a participant did not respond to any of the relevant variables. Given that we are not necessarily interested in individual trajectories in our analyses, we will use all available data (i.e., will not remove entire participants but rather individual incomplete daily diary measurements).
Exploratory analyses (optional)	<p>? proposed in their integrative theory of intergroup contact that “group salience” should moderate the positive effect of intergroup contact on outgroup attitudes. They argue that especially when the outgroup membership is salient, the positive attitudes should generalize from the outgroup individual to the general outgroup. We have data to test this hypothesis within a large intensive longitudinal dataset of daily intergroup interactions. We do this by separately assessing attitudes towards the outgroup interaction partner and the outgroup generally. And by assessing whether perceiving the interaction as representative of the outgroup moderates the positive contact effects.</p> <p>To build confidence in our conclusions about the role of contact quality we calculate an alternative measure of interaction quality that includes perceived pleasantness, meaning, effectiveness, and importance of the interaction.</p> <p>We expect that certain contextual characteristics may be related to the positive effects of high quality outgroup interactions. Therefore, we will look for the effect of key need fulfillment on outgroup attitudes while also including information about the interaction partner (perceived closeness, trust, relationship), the interaction</p>

characteristics (duration, interaction language, group vs. dyadic), as well as characteristics of the key need (importance of the need, fulfillment due to the interaction partner).

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., Bringmann, L. F., Engler, R. F., de Jonge, P., and Epstude, K. (2022). 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.
- McKeown, S. and Dixon, J. (2017). The “contact hypothesis”: Critical reflections and future directions. *Social and Personality Psychology Compass*, 11(1):1–13.

- 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. (2011). *When groups meet: The dynamics of intergroup contact*. Psychology Press.
- 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.