DRAFT: Do not distribute outside of LIFE Academy

The new person-specific paradigm in psychology

Peter C.M. Molenaar
The Pennsylvania State University
Email: pxm21@psu.edu

1. Introduction

During about a century psychological data analysis proceeds according to a standard approach that is focused on the inference about states of affairs obtaining within a population of homogeneous human subjects. The information which is employed to infer about such states of affairs is of a particular type, namely inter-individual variation (variation between subjects; individual differences). This approach underlies all standard statistical analysis techniques such as analysis of variance, regression analysis, path analysis, factor analysis, multilevel and mixed modeling techniques. Whether the data are obtained in cross-sectional or longitudinal designs, the statistical analysis always is focused on the structure of inter-individual variation at the population level. Parameters and statistics of interest are estimated by pooling across subjects, where these subjects are assumed to be homogeneous in all relevant respects. This is the hall-mark of analysis of inter-individual variation: the sums defining the estimators in statistical analysis are taken over different subjects randomly drawn from a population of presumably homogeneous subjects

It would seem that inferences about states of affairs at the population level obtained by pooling across subjects constitute general findings that apply to each individual subject in a homogeneous population. Yet such applications to individual subjects involve a shift in level, namely from the level of inter-individual variation in the population to the level of intra-individual variation characterizing the life histories of individual subjects. Is this shift between levels allowed?

It will be shown that in general the answer is negative. Knowledge about the structure of inter-individual variation in the population usually cannot be applied at the level of individual subjects making up this population. This is the direct consequence of general mathematical-statistical theorems – the so-called classical ergodic theorems. In what follows we will give a heuristic description of some of the wide-ranging consequences of these theorems for psychological methodology and statistical analysis. These consequences are so profound that the changes required in order to accommodate them will involve a Kuhnian paradigm change in psychological science.

2. Heuristic description of ergodicity

Each real human being can be conceived of as a high-dimensional integrated system whose behavior evolves as function of place and time. The system includes important functional subsystems such as the perceptual, emotional, cognitive and physiological systems, as well as their dynamic interrelationships. A subject's time-dependent behavior can be represented as a trajectory (life history) in a high-dimensional behavior space. The complete set of life histories of a population of human subjects can be represented as an ensemble of trajectories in the same behavior space.

With respect to an ensemble of trajectories in behavior space, inter-individual variation is defined as follows: (i) select a fixed subset of variables; (ii) select one or more fixed time points as measurement occasions, (iii) determine the variation of the scores on the selected variables at the selected time points by pooling across subjects. Analysis of inter-individual variation thus defined is called R-technique by Cattell (1952). In contrast, intra-individual variation is defined as follows: (i) select a fixed subset of variables; (ii) select a fixed subject; (iii) determine the variation of the scores of the single subject on the selected variables by pooling across time points. Analysis of intra-individual variation thus defined is called P-technique by Cattell (1952).

Ergodicity addresses the following fundamental question: Given the same set of selected variables, under which conditions will an analysis of inter-individual variation yield the same results as an analysis of intra-individual variation? To illustrate: under which conditions will factor analysis of inter-individual variation (R-technique) yield a solution that is equal to the one obtained in factor analysis of intra-individual variation (P-technique)?

The general answer to this question is provided by the classical ergodic theorems (cf. Molenaar, 2004; Molenaar et al., 2003): An analysis of inter-individual variation yields the same results as an analysis of intra-individual variation only if the ensemble of time-dependent trajectories in behavior space obeys two rigorous conditions. Firstly, the trajectories of all subjects in the ensemble have to obey exactly the same dynamical law (homogeneity of the ensemble). Secondly, each trajectory should have constant statistical characteristics in time (stationarity, i.e., constant mean level, variance and correlation). The two conditions for ergodicity will be further explained shortly. In case either one (or both) of these two conditions is not met, the psychological process concerned is non-ergodic, i.e., its structure of inter-individual variation will differ from its structure of intra-individual variation. For a non-ergodic process, the results obtained in standard analysis of inter-individual variation do not apply at the individual level of intra-individual variation, and vice versa.

3. Explanation and illustration of the conditions for ergodicity

The first condition for ergodicity is that the trajectory of each subject in the ensemble has to obey the same dynamical law (homogeneity of the ensemble). This implies that the statistical model describing the trajectory of each subject is invariant across subjects. For instance in the context of Cattell's Data Box, the condition implies that the trajectories of all subjects should obey the same P-technique factor model in which the number of factors and the factor loadings are invariant across subjects. If the condition of ensemble homogeneity is violated, the process concerned is non-ergodic and its structure of inter- and intra-individual variation will differ. The next section gives an empirical illustration.

3.1 Heterogeneous Big Five personality processes We consider factor analysis of multivariate time series of scores on a personality test, obtained in a replicated time series design by Borkenau[1]; the reader is referred to Borkenau & Ostendorf (1998) for a detailed description. The replicated time series design involves N=22 subjects measured at T=90 consecutive days with (parallel versions of) the same questionnaire composed of 30 items (6 items for each Big Five personality factor). Standard oblique factor analysis of a robust estimate of the (30,30)-dimensional inter-individual covariance matrix corroborates the nominal Five Factor model. That is, a confirmatory 5-factor model in which the loadings conform to the nominal Five Factor structure yields a satisfactory fit to the inter-individual covariance matrix.

However, it is found that factor analysis of the 30-variate time series of *each* individual subject yields solutions that do *not* conform to the nominal Five Factor structure. Hence we have here a clear example of a heterogeneous ensemble: the intra-individual covariance structures differ between subjects in the number of factors, their loadings and sequential dependencies. Because the measured personality process violates the condition of ensemble homogeneity, it is non-ergodic and therefore the inter-individual (Big Five) structure cannot be generalized to the level of intra-individual variation.

Recently a hybrid model has been developed with which it possible to simultaneously analyze the total variation in a replicated time series design at both the inter- and intra-individual levels (Hamaker et al., 2007). Figure 1 depicts the contribution to the total variation of item scores by the inter-individual (Five Factor) and intra-individual factors thus obtained. Although this decomposition is initially obtained for each individual subject separately, total variation in Figure 1 is defined as the average over all subjects k=1,…,22 and all times t=1,2,…,90. The results are presented according to the nominal (Big Five) 5-factor structure underlying the test. This way of presentation has been chosen because it is convenient; it is not mandatory and other ways of presentation are possible because the decomposition is obtained per item/subject.

For each of the 30 items three bars are presented. The first white bar on the left is the total observed variation, averaged across subjects and times (inserted line denotes standard deviation across subjects and times). The second light-grey bar in the middle is the amount of *intra*-individual variation explained by the subject-specific dynamic factor models, also averaged across subjects and times (again, inserted line denotes standard deviation). The third dark-grey bar on the right is the *inter*-individual variation explained by the common Big Five factor model. Clearly, on average the contribution of the intra-individual variation to the total variation of each item is consistently larger than the contribution of the inter-individual variation.

[1] The author would like to thank Peter Borkenau for kindly making available the data.

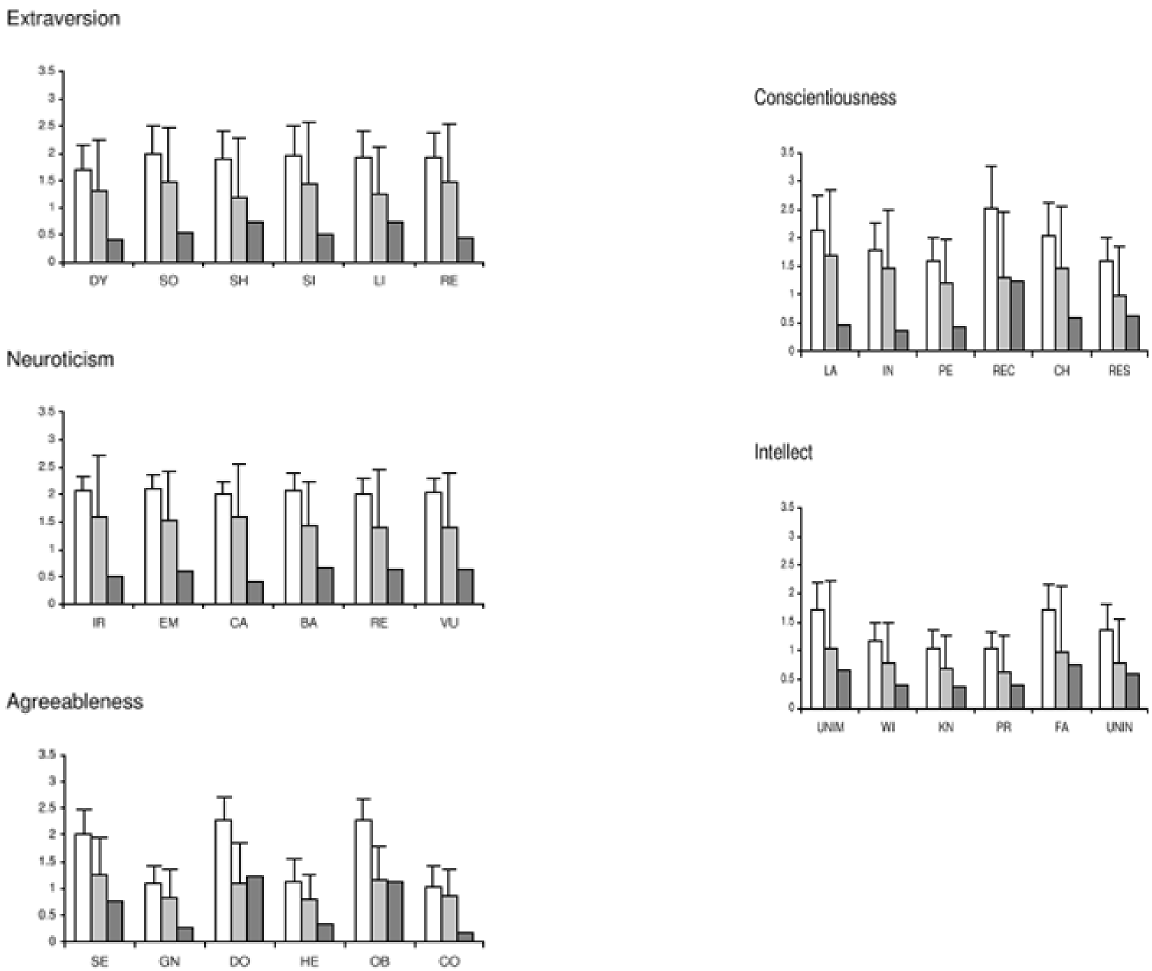

Figure 1 Decomposition of the total variation of personality test item scores in a replicated time series design into intra- and inter-individual components (see text for details)

3.2 The stationarity criterion for ergodicity The second condition for ergodicity of psychological processes in behavior space is that each trajectory should have constant statistical characteristics in time. Such a trajectory is called stationary. All processes with non-stationary trends and/or non-stationary sequential dependencies (autocorrelation) are non-ergodic. Prime examples of such processes are developmental processes, which almost by definition have time-varying statistical characteristics. But also learning, habituation, transient brain responses, and many more psychological processes are non-stationary. All these processes violate the stationarity criterion for ergodicity and hence their inter- and intra-individual structures of variation differ.

The proper way to analyze a non-stationary process is by means of special statistical time series analysis techniques, which until recently were unavailable in psychological science. One needs factor models for non-stationary multivariate time series in which not only the error variances, but also the factor loadings and the sequential dependencies of factor series can be arbitrary time-varying. Such models have become available only recently (Molenaar et al., 2006). In the next section an empirical is presented.

3.3 Non-stationary emotional experiences To illustrate the application of the newly developed model for non-stationary time series, data were taken from a study of the emotional experiences of sons (age range 14-18 years) as they interact with their fathers over time. For each participant a period of 6-8 weeks of data collection lasted until 80 interactions occurred (T=80).

In what follows the 3-variate of a single participant will be considered. The first series is interpreted as Involvement, the second as Anger, and the third as Anxiety. The time series model for Involvement describes the value of Involvement at the (t+1)-th interaction episode as function of Involvement, Anger and Anxiety at the preceding t-th interaction episode; t=1,…,80. The time series models for Anger and Anxiety are defined similarly. All parameters in these time series models are time-varying; their estimated time-varying values are plotted in Figure 2.

The uppermost panel in Figure 2 shows that the parameter $\beta_{11}(t)$ measuring the dependence of Involvement at t+1 on the value of Involvement at t declines from .44 at the start of the observation interval to .22 at t=45 and remains about constant thereafter. The parameter $\beta_{13}(t)$ measuring the dependence of Involvement at t+1 on Anxiety at t increases from -.20 at the start of the observation interval to .20 at about t=45 and then slightly decreases again. Hence at the start of the observation interval an increased value of Anxiety at t leads to a decrease in the value of Involvement at t+1, whereas in the second half of the observation interval an increased value of Anxiety at t leads to an increase in the value of Involvement at t+1. Noteworthy also is the steady increase in the value of the parameter $\beta_{32}(t)$ in the bottom panel of Figure 2, which measures the dependence of Anxiety at t+1 on Anger at t. It increases from a value of .08 at the start of the observation interval to .40 at t=80. Hence at the beginning of the observation interval the Anxiety process is almost unrelated to Anger, whereas at the end of the observation interval an increased value of Anger at t leads to an increase in the value of Anxiety at t+1.

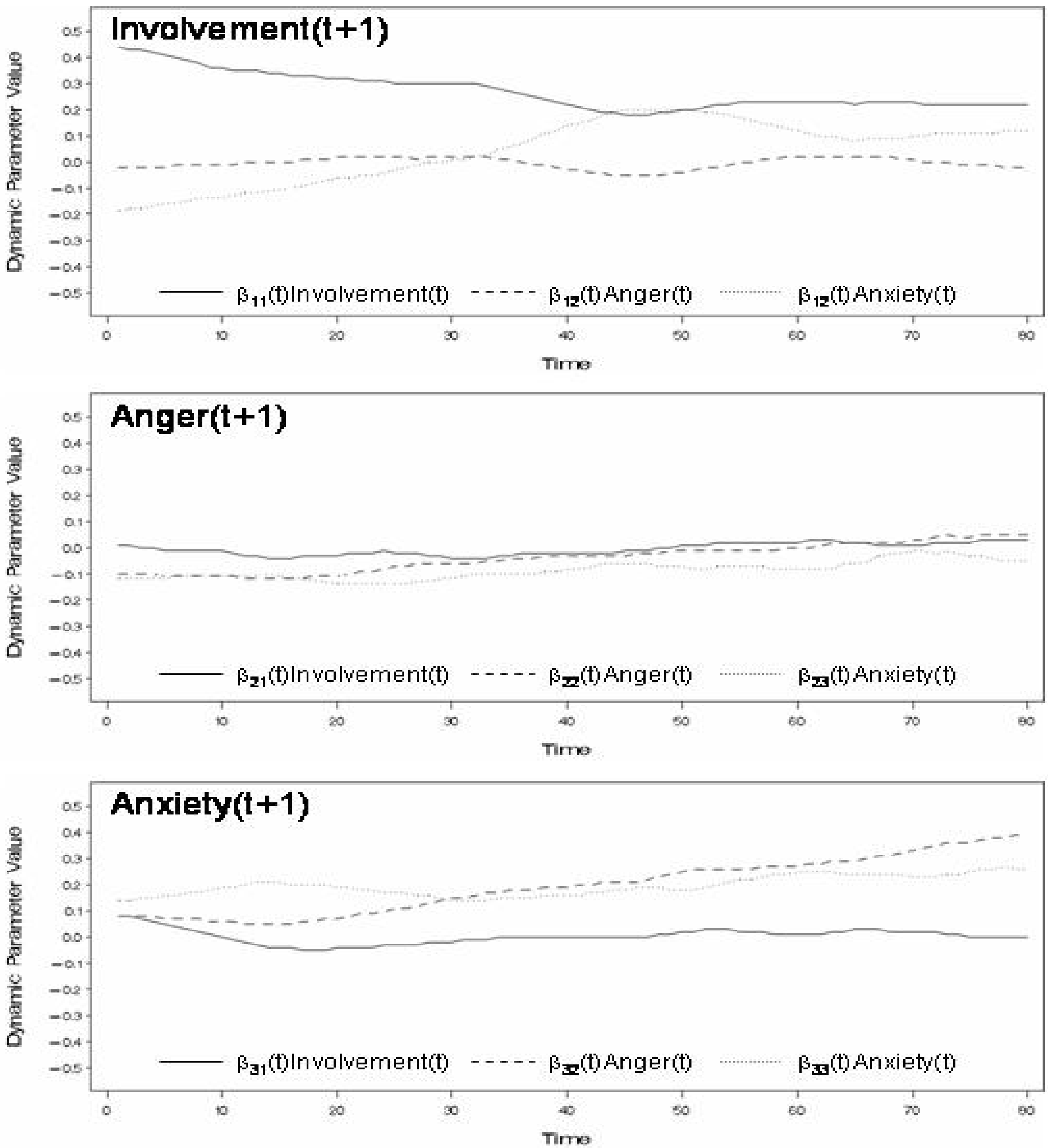

Figure 2 Time-varying parameter estimates in the model describing a son's emotional experiences interacting with his father (see text for details)

4. Discussion and conclusion

Psychological processes like cognitive information processing, perception, emotion, and motor behavior occur in real time in concrete individual persons. In so far as these processes are non-ergodic, i.e., obey person-specific dynamic laws and/or have non-stationary statistical characteristics, their analysis should be based on intra-individual variation. That is the direct consequence of the classical ergodic theorems. A heuristic description was given of the two conditions which are required for a psychological process to be ergodic. The first condition, ensemble homogeneity, requires that the trajectories of all subjects in behavior space obey the same dynamic law (in particular, the same statistical model). An empirical example involving replicated time series of Big Five personality scores was presented where this condition is violated. It was found that while the inter-individual covariance matrix of these personality scores obeys the nominal Big Five factor model, the intra-individual covariance matrices of different subjects obey factor models that differ in various ways from each other. Hence the homogeneity condition for ergodicity is violated by these data, and therefore it is no surprise that none of the single-subject dynamic factor models conforms to the nominal Big Five factor structure.

The second condition for ergodicity is that each trajectory in behavior space should have constant statistical characteristics in time. Many psychological processes, in particular developmental and learning processes, violate the stationarity criterion and therefore are non-ergodic. Scientific study of such processes has to be based on analysis of intra-individual variation. This requires the application of special statistical modeling techniques which are appropriate for the analysis of non-stationary psychological time series. These techniques have become available only recently, and we presented an illustrative application to a non-stationary time series of emotional experiences.

One worthwhile spin-off of the necessary focus on intra-individual variation at the level of individual subjects involves the possibility to optimally guide the psychological processes concerned. The time series models obtained in analyses of intra-individual variation can be used to carry out feedback-feedforward guidance in real time. For instance, we are involved in new research projects focusing on person-specific optimal guidance of daily treatment of diabetes and asthma. In principle the modeling of intra-individual variation associated with any psychological process can be used to implement optimization techniques in order to guide that process to desirable levels and keep it there for the continuous benefit of each individual person in his/her time-varying specific life situation.

In conclusion, we are at the brink of a major reorientation in psychological methodology in which the focus is on the intra-individual variation characterizing time-dependent psychological processes occurring in the individual human subject. It will require substantial efforts from the community of psychological scientists to effectuate this reorientation. At present there is hardly any literature on multivariate time series designs and analysis techniques that is tailored to the special needs of psychological science. Neither are there established curricula to teach students of psychology in state-of-the-art statistical techniques and methodology for the analysis of intra-individual variation. Several research centers collaborating within the Developmental Systems Group[2] are in the process of initiating new teaching and research projects focused on intra-

[2] Additional information about the Developmental Systems Group can be obtained from the author

individual variation of cognitive, emotional and personality processes, because the dedicated study of intra-individual variation is, in view of the classical ergodic theorems, no longer an option, but a necessity.

References

Borkenau, P., & F. Ostendorf, (1998). The Big Five as states: How useful is the five-factor model to describe intra-individual variations over time? *Journal of Personality Research*, *32*, 202-221.

Cattell, R.B. (1952). The three basic factor-analytic designs – Their interrelations and derivatives. *Psychological Bulletin*, *49*, 499-520.

Hamaker, E.J., Nesselroade, J.R., & Molenaar, P.C.M. (2007). The integrated state-space model. *Journal of Research in Personality*, *41*, 295-315.

Molenaar, P.C.M., 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 life-span psychology*. Dordrecht: Kluwer, pp. 339-360.

Molenaar, P.C.M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*, 201-218.

Molenaar, P.C.M., Rovine, M.J., Ram, N., & Corneal, S.E. (2006). Analysis of developmental processes based on intra-individual variation by means of non-stationary time series modeling (submitted).

.

.