



Revitalizing the typological approach: Some methods for finding types

Lars R. Bergman¹, András Vargha^{2,3}, and Zsuzsanna Kövi²

¹ Stockholm University, Stockholm

² Károli Gáspár University, Budapest

³ Eötvös Loránd University, Budapest

Address Correspondence to: Lars R. Bergman, Department of Psychology, Stockholm University, 106 91 Stockholm, Sweden

Email address:

lrb@psychology.su.se

To cite this article:

Bergman, L. R., Vargha, A., & Kövi, Z. (2017). Revitalizing the typological approach: Some methods for finding types. *Journal for Person-Oriented Research*, 3(1), 49-62. DOI: 10.17505/jpor.2017.04

Abstract

The purpose is to discuss and exemplify how a typological approach could be designed for studying phenomena believed to be best understood within a person-oriented theoretical framework. The focus is mainly restricted to the case of studying the typological structure of a sample at a single point in time, and restricted to analyzing variable profiles where each variable has a “negative” and “positive” endpoint. An artificial data set and an empirical data set were analyzed using two different methodological approaches, one more explorative (using LICUR, a cluster analysis-based procedure) and one more model-based (using the MCLUST procedure). For the artificial data set, the LICUR procedure was successful in finding the true classification structure but the MCLUST procedure performed surprisingly badly. For the empirical data set, both procedures produced rather similar solutions and they showed moderate validity. However, the LICUR solution appeared to be slightly superior. It was argued that applying a sound classification methodology and carefully validating the resulting classifications are extremely important, even more so in a developmental context. It was also argued that, in a number of situations, a more explorative approach could be more useful than a standard model-based one.

Keywords: person-oriented approach, classification, types, typology, cluster analysis, LICUR, model-based analysis

The study of types and typologies are old research areas in psychology that have been largely abandoned in modern psychology, in contrast to what is the case in, for instance, biology. Instead, a dimensional approach has become predominant, often focusing on models for studying relationships between variables. For this shift of approach there are a number of reasons, for instance the extreme subjectivity often involved in constructing older typologies and the emergence of elegant statistical models for analyzing dimensional data. In these models, the variables are treated as the fundamental analytic units. However, the theoretical framework of the modern person-oriented approach implies that a variable-oriented approach is in many research contexts incongruent with basic assumptions of the process under study (Bergman & Andersson, 2010; Bergman & Vargha, 2013). In contrast, an approach focusing on typical

patterns, akin to a typological approach, can in such contexts better match these basic assumptions. A revival of the typological approach in a more modern form is called for, based on methods of analysis that are to a large extent freed from the subjectivity of the old approaches.

The purpose of this paper is to discuss and exemplify how a typological approach could be designed for studying phenomena believed to be best understood within a person-oriented theoretical framework. The focus is rather narrow, being mainly restricted to the case of studying the typological structure of a sample at a single point in time. Examples of recent, broader presentations of the person-oriented approach, including many methodological issues, are given by Bergman and Lundh (2015) and by Wiedermann, Bergman, and von Eye (2016).

Definitions

No generally accepted definitions exist of the concepts “type” and “typology”. For the sake of clarity we will therefore provide definitions of them and of some other concepts, recognizing that researchers in other contexts might prefer other definitions. Admittedly, our definitions are rather restrictive and somewhat limit the scope of our paper.

1. With *data* is meant data in the sense of Cattell’s data box, and, in addition, it is assumed that all variables are interval scaled variables (this restriction does not apply to Cattell’s data box). The basic unit of analysis we are concerned with is the vector of variable values for a given individual at a given time point. This vector (*value profile*) is regarded as the natural basic unit that should be retained in the analyses (motivated in the next section).

2. With a *type* is meant a value profile that tends to occur frequently – either for the same person across time (intra-individual perspective) or occurs for many persons at the same time (inter-individual perspective). A type could have been found when analyzing a single sample or it could have emerged from a synthesis of the findings from analyzing many samples. A type could also refer to a theoretical expectation of a certain value profile being frequent. It should be pointed out that, in practice, perfect types are rarely observed. Therefore a less restrictive definition of type is usually applied in which data points with very similar value profiles are regarded as belonging to the same type.

3. With a *typology* is meant a set of types that together describe data. It could be with regard to a single sample or several samples, and it could also be a classification model derived from theory that needs to be empirically tested. Often the term “typology” is used for a sample of persons divided into mutually exclusive subsamples with each one characterized by approximately the same type (e.g. a classification based on cluster analysis of a sample of persons’ value profiles). There exist also more narrow definitions of the typology concept, stressing that the classification structure then is of theoretical relevance and/or has demonstrated generalizability.

Some basic tenets of the person-oriented theoretical framework and their implications for the choice of methodological approach

The modern person-oriented approach is an outgrowth from the holistic-interactionistic research paradigm developed by David Magnusson (e.g., Magnusson, 1988; Magnusson & Törestad, 1993). The approach has been presented and discussed in numerous papers, including some that have been concerned with the basic tenets and assumptions of the approach (e.g. Bergman & Magnusson, 1997; Bergman & Andersson, 2010). A selection of these tenets is presented below in a somewhat reformulated form

to provide a basis for a discussion of their methodological implications.

1. The person-oriented theoretical framework includes a dynamic systems view where the studied universe is best understood in terms of the operation of one or more dynamic systems normally characterized by emerging attractors. Often each attractor is a single specific system state (or a narrow region of phase space) that in some way is “optimal” for the survival of the system for a given set of start conditions and, hence, is often frequently observed (~type). This is consistent with a central tenet of the holistic-interactionistic research paradigm: Individuals function as whole organisms where the different parts work together and adjust to each other to achieve “good” functioning. It implies that the information the researcher has about the studied system should be regarded as “a whole” as far as it is possible in order to reflect basic system properties and individual functioning. From this standpoint, the method for data analysis should strive for conserving the totality of the information and analyze whole value patterns (value profiles). Hence, in many cases the value pattern and not the variable should be the basic unit of analysis. Vital information may be lost if this is violated (e.g., forming a variance-covariance matrix and using it as data in the analysis would be a violation). Of course, in some cases and to some extent this conservation can be accomplished by using a tailored variable-oriented methodological approach (e.g., applying a suitable dummy variable coding of some interesting specific value profile). From the above it should be clear that, if a person-oriented theoretical framework is accepted, a classificatory approach aiming at finding typical variable patterns (e.g., by using cluster analysis) matches this framework in a number of contexts because the holistic character of the information is retained in the analysis.

2. Another central tenet concerns the focus on understanding the *individual*: the findings should be interpretable for the single individual. This can normally not be accomplished by applying standard statistical methods that produce estimates of *group* parameters (Molenaar, 2004; von Eye & Bergman, 2003). For instance, a correlation coefficient computed for a sample of persons is a group statistic that is not informative of a specific individual’s position in the two variables and it is not even very informative of an individual’s value in one variable conditional on the value in the other variable, except when the correlation is very large. This second tenet implies that in some contexts a typological approach can also fall short of presenting findings that are interpretable at the level of the individual (for instance, a classification with classes that are heterogeneous).

3. Taken together, several tenets of the person-oriented research paradigm point to that, in most nonexperimental settings, the studied systems are complex, imperfectly understood, and they exhibit variation across individuals and age. From this perspective, the commonly used “big”

models, claiming to explain multivariate variation, are often premature and unrealistic (Richters, 1997). In such cases, a more modest ambition level is called for in which the focus is on theory guided exploration and building “smaller” models that, at a later stage, can be building blocks for “bigger” models. Within typological research, this suggests that many forms of model-based classification analyses can be premature. Normally, it is not expected that there exists only one “true” typology that divides a sample into a small set of classes with each class sharing the same value profile, except for random noise. In fact, usually several different but similar classifications have comparable (limited) explanatory power of the data and, in addition, it is often to be expected that there exists a residue of data points that are unclassifiable (i.e., clearly do not belong to any of a small set of types/classes, see Bergman, 1988).

Researchers engaged in model-based classification sometimes claim that this approach is superior to other more exploratory approaches like cluster analysis, mainly on the grounds that model-based findings “explain” data by a simple elegant model with confidence intervals of parameter estimates and with the strength of being able to test model fit. Considering what was said in the previous paragraph, this claim seems overstated because the assumptions made in such models are often unrealistic. The counter argument is that, if the model fits the data, it is shown to be a good model. However, when testing such models, model fit is usually *not* based on discrepancies between data points estimated by the model and the actual data points. Instead, standard tests compare *group* parameters estimated by the model to the corresponding sample group statistics (e.g., actual intra class correlations are compared to the zero correlations expected from the model). In addition, for moderate sample sizes the power to reject such models may be low in many cases (see the Discussion section). Often a well-designed explorative classification approach is more compatible with the person-oriented tenets presented above. In such an approach, the aim is to find one or more classification structures that reasonably well summarize the data structure, and the classification of residue objects (\approx outliers) is avoided to prevent them from distorting the classification structure. Of course, what has been stated above should not be interpreted as a critique of the usefulness of *all* model-based approaches in the context discussed here. Their appropriateness in a given context depends on the assumptions they make about the data structure, and many types of such models exist.

Finally, we add an admittedly subjective comment: Some of us who have extensive experience of classification analyses and are active in different research areas have used model-based classification analysis and compared its findings to those produced by some more explorative method. The findings from the former often seemed doubtful in relation to what was expected from our knowledge of the substantive field under study. For

instance, using Latent Profile Analysis the analysis frequently indicated a three-class solution with one class being generally high, one being generally low, and one being generally intermediate. That a complex multivariate system is well described by such a simple model would seem to be an exceptional case, and not one that is often observed. In comparison, the explorative method often indicated that more classes were needed to describe the data structure, with some indicating types with more complex value profiles that also (partly) fitted theoretical expectations.

Comparing findings using some different types of classification analysis

Methods for classification are just tools and no method is generally superior to any other reasonable method – it all depends on the scientific problem, properties of the studied data set, and the assumptions about the data generating process one can make in the specific research context. Nevertheless, it is useful to examine examples of how different methods perform in well-defined contexts, as long as the findings are not over-interpreted. In this article, a cluster analysis-based partly explorative method and a more model-based standard method are compared. Perhaps the most useful conclusion that may result from a method comparison based on a few examples is that if some method performs badly, there is some ground for doubting its general usefulness.

In the following, two data sets are analyzed, one data set based on artificial data with a typological structure that can be expected to be not uncommon, and one empirical data set where theoretical expectations exist about the nature of the classification structure and some possibilities exist to empirically validate an obtained classification by using external variables. The method of comparison will concern three aspects of how well a classification method performed: (1) the degree to which the classes produced are homogenous and distinctive; (2) whether the description of the classes show internal and external validity, and (3) the extent to which the objects that were classified belonged to the “appropriate” class.

Data sets

Artificial data set, medium sample size (n=400)

As an example, we have constructed an artificial data set that is rather typical of a number of real data sets that have been subjected to classification analysis. First, we assumed all variables are scaled from “bad” (coded 1) to “good” (coded 5), like when studying adjustment, and that the value profile is constituted by four such variables. Second, we assumed there is one small “bad” class (10% of sample, Type A), and there are two large classes, one “generally good” (40% of sample, Type B), and one “generally medium” (40% of sample, Type C). There is also one small

additional class, “rather bad” in some variables, and “rather good” in some other variables (10% of sample, Type D). Further, all subjects in a class (i.e., of a type) are characterized by exactly the same value profile in the four variables. This theoretical “true” data set (TEO data set) is presented in Figure 1.

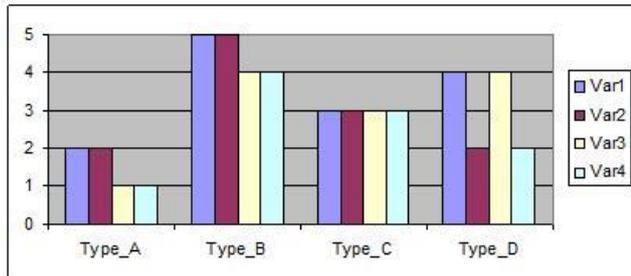


Figure 1. Theoretical type structure (all cases in a type have exactly the same value profile): Design with four types (A to D) and four variables (Var1 to Var4)

Based on the TEO data set, a new data set was then constructed with errors of measurement added to the true values to reflect the imperfections in measurement that almost always exist. This data set, called the MEA data set, will be analyzed in the Results section. The MEA data set was constructed in the following way. To each variable in the TEO data set we added to the true value a random normal variable with a mean of zero and a variance that corresponded to a reliability of 0.80, which is a rather common level of measurement precision. More specifically, independently for each original true variable value in the TEO data set, a new value was created by adding a rounded independent random $N(0; 0.5)$ variate. For an $X = N(0; 0.5)$ variable, $P(-0.5 < X < 0.5) = 0.68$. Data values less than 1 or greater than 5 were set to 1 or 5, respectively. For TEO data set values 1 to 5, the percentages retaining the TEO score in the MEA data set after the data generations were 83.0, 70.1, 68.1, 67.1, and 83.2, respectively (average = 74.3). This means that in 74.3% of cases, the generated rounded data value (TEO value plus random error) did not differ from the corresponding TEO value. This is greater than the initially expected 68%, because the most extreme values (1 and 5) could not freely change due to the ceiling and bottom effect, in difference to the values that were nearer to the center of the 5-point scale (2, 3, or 4).

Empirical data set ($n=541$)

We also analyzed an empirical data set in the Results section. The data set was taken from the Swedish longitudinal program Individual Development and Adaptation (IDA; Magnusson, 1988). The data concerned teacher ratings of boys' extrinsic school adjustment at age 13, and three variables constituted the value profile to be analyzed: Aggression (Aggr), Motor Restlessness (Motr), and Concentration difficulties (Concd). All variables were coded by 1 (no adjustment problem), 2 (“normal” level of adjustment

problem), 3 (tendency to adjustment problem), 4 (a clear adjustment problem), and 5 (an extreme adjustment problem). All variables were positively skewed. The number of boys in the analyzed data set was 541 and they can be considered fairly representative of the general population of Swedish boys aged 13 at the time period of the original data collection (1968), see Bergman, Corovic, Ferrer Wreder, & Modig (2014). The analyzed data set is called the EMP data set.

From previous research within IDA, and also from findings in other studies, we had expectations about three types that should emerge from the analyses: 1. One characterized by generalized adjustment problems. 2. One characterized by hyperactivity (i.e., high in both Motor Restlessness and Concentration difficulties), and 3. A type characterized by generalized good adjustment.

Results from analyses of the artificial data set with measurement errors added (MEA data set)

An explorative cluster analysis-based approach

In all analyses, the ROPstat statistical package was used (Vargha & Bergman, 2015). Ward's (1963) hierarchical cluster analysis was applied, in some cases followed by a k-means cluster analysis using a Ward solution as the start classification. This was done following the LICUR rationale that also includes procedures for choosing the number of clusters and for handling outliers (Bergman, Magnusson, & El-Khoury, 2003). LICUR was developed in the context of classifying adjustment problems, and the procedure is suitable for the kind of data we have. Of course, cluster analysis is not a method but rather a label for a large class of different methods used for the classification of objects and that have been applied for many decades in many sciences (see, for instance, Milligan, 1980, for an overview and discussion of many types of cluster analysis). Other sound methods of cluster analysis could have been applied, and they would probably have produced similar but not identical findings to the ones we present. This does not mean that cluster analysis methods are inferior and should be avoided – it is instead a consequence of the extreme difficulty in obtaining a single adequate summary of the structure of a complex multivariate data set. Normally, this structure cannot be summarized by a single clear-cut model without resulting in a distorted representation of the profile structure in the data. We would have liked to also present findings from some other sound method for cluster analysis but this was not possible within the constraints of a short article. In the Discussion section, some more general issues of classification analysis are presented that relate to what has been said above.

Before cluster analysis, an analysis was made to identify multivariate outliers, following the LICUR rationale (Bergman, Magnusson, & El-Khoury, 2003). Standard LICUR criteria were used and no outliers were found.

Hence, all objects belonging to the MEA data set were included in the cluster analysis. The variables were not standardized because they were already scaled in the same way (values 1-5 with the same anchors). Then the data were cluster analyzed using Ward's method.

The following quality coefficients (QCs) were used to evaluate a cluster solution (see more detailed descriptions in Vargha & Bergman, 2015, and Vargha, Bergman, & Takács, 2016).

1. The homogeneity coefficient (HC) of a cluster is the average of the pairwise within-cluster distances of its cases. To evaluate a cluster solution, HCmean can be used as a QC. It is the weighted mean of the cluster HC values (weights are cluster sizes).

2. Explained error sum of square percentage (EESS%), a multivariate generalization of eta-squared, known in ANOVA:

$$\text{EESS}\% = 100 * (\text{SS}_{\text{total}} - \text{SS}_{\text{cluster}}) / \text{SS}_{\text{total}}, \quad (1)$$

where SS_{total} is the sum over the whole sample of each case's sum of squared deviations between each variable value and the mean for the whole sample in that variable, and $\text{SS}_{\text{cluster}}$ is the sum over clusters of the within cluster sums of squared deviations between cases and variable centroids.

3. Cluster point-biserial correlation (PB), a Pearson correlation computed in the sample of all pairs of cases between the binary variable of belonging to the same cluster (0) or not (1), and the distance between the two paired cases. A well-known formula of PB (see, e.g., Glass & Hopkins, 1996):

$$PB = \frac{M_1 - M_0}{s_{n-1}} \sqrt{\frac{n_1 n_0}{n(n-1)}} \quad (2)$$

Here M_0 is the average pairwise within-cluster case distance, M_1 is the average pairwise between-cluster case distance, $n = N(N-1)/2$ is the number of pairs of cases in the total sample of size N , and n_0 and n_1 are the number of pairs of cases that belong to the same (n_0) or to different (n_1) clusters; s_{n-1} is the SD of the pairwise differences between cases in the total sample of size n .

4. Considering that PB depends primarily on the $M_1 - M_0$ difference, the first component in formula (2) that is a kind of standardized difference of $M_1 - M_0$ can also be used as a QC, called CLdelta. It can be explained analogously to the well-known Cohen delta effect size measure (Cohen, 1977). CLdelta indicates the extent to which cases are closer to their own cluster members than to cases from other clusters.

5. A simplified version of the Silhouette coefficient (SC) was defined as follows. First, compute SC_i for each case i in the sample, using formula (3):

$$SC_i = (B-A) / \max(A, B), \quad (3)$$

where A is the distance from the case to the centroid of the cluster that the case belongs to, and B is the minimal distance from the case to the centroid of every other cluster. SC is the average of all cases' SC_i values. A high SC value indicates that on the average, cases are substantially closer to their own cluster centers than to the nearest of other cluster centers.

6. XBmod, a modified version of the Xie-Beni index (Xie & Beni, 1991) was defined as follows:

$$\text{XBmod} = (D - W) / \max(D, W), \quad (4)$$

where W is the average distance of cases from their own cluster centers, whereas D is the distance of the two closest cluster centroids. The meaning of XBmod is similar to that of SC.

7. The GDI24 index is a special case of the family of generalized Dunn indices and it can be defined as follows (Desgraupes, 2013):

$$\text{GDI24} = \frac{D}{\max_k (HC_k)} \quad (5)$$

where D is the same as in (4), and $\max_k (HC_k)$ is the HC value of the most heterogeneous cluster.

Some QCs of cluster solutions with different numbers of clusters are presented in Table 1. Based on the size of the QCs, $k=4$ seems to be the best cluster number, since before $k = 4$ there is a low stepwise decrease in EESS%, but at $k = 3$ there is a sudden drop (from 77.20 to 70.77), and this is the case for HCmean and the HC range as well. For $k > 4$ many QC coefficients (PB, XBmod, SC) worsen (XBmod radically from .780 to .067), whereas EESS%, HCmean, and the HC range do not change substantially. We tried to improve the cluster solution via relocation analysis but could not do so to any appreciable extent, so the hierarchical 4-cluster solution was accepted (bold face type in Table 1). In addition, other things being equal, a hierarchical solution is also preferable in that it retains a straightforward classificatory relationship to hierarchical solutions with more or fewer clusters.

To obtain further evidence about the quality of a cluster solution of real data, it is important to also show that it is significantly and in a measurable way substantially better than a solution obtained on a random data set of the same size, with the same number of variables, and same number of clusters. For this purpose, Vargha et al. (2016) developed the MORI coefficient. MORI measures the relative improvement of a cluster structure (as measured by a QC) obtained for real data as compared to that obtained for the cluster structures resulting from analyzing several types of random data sets with the same general properties as the real data set. In our study, we first chose as random controls the independent random permutations of the values of the input variables, and the independent random normal variables.

Table 1

Quality coefficients (QCs) of cluster solutions for cluster numbers (k) between 2 and 10, MEA data set.

k	Quality coefficients					
	EESS%	PB	XBmod	SC	HCmean	HC range
10	85.48	0.362	0.022	0.639	0.376	0.25-0.57
9	84.45	0.366	-0.045	0.615	0.402	0.25-0.60
8	83.42	0.429	0.283	0.639	0.427	0.26-0.60
7	82.23	0.441	0.122	0.610	0.456	0.26-0.76
6	80.94	0.473	0.134	0.631	0.488	0.26-0.76
5	79.44	0.520	0.067	0.705	0.525	0.45-0.76
4	77.20	0.607	0.780	0.799	0.580	0.45-0.73
3	70.77	0.622	0.830	0.822	0.741	0.45-1.02
2	56.48	0.579	0.815	0.832	1.101	0.45-1.54

Table 2

MORI validation coefficients for the 4-cluster HCA solution for seven QCs and three types of random control data sets (number of independent random replications = 100), MEA data set.

Type of random control data set	MORI coefficients for different QCs						
	EESS%	PB	XBmod	SC	HCmean	CLdelta	GDI24
Random permutation	0.63	0.38	0.73	0.64	0.63	0.57	2.59
Independent normal	0.67	0.46	0.79	0.69	0.67	1.08	3.38
Correlated normal	0.41	0.37	0.70	0.61	0.41	0.54	2.01

Note: all MORI coefficients are significantly larger than 0 at the $p < .001$ level

Table 3

The match of theoretical types (TEO data set, Figure 1) to cluster centroids (MEA data set, Figure 2).

Index	Theoretical Type	HCA4	ASED
1	B	CL2	0
2	A	CL1	0
3	C	CL3	0.002
4	D	CL4	0.013

Table 4

Cross-tabulation of the original types in the TEO data set (Type_A to Type_D) and cluster membership in the analysis of the MEA data set (CL1 to CL4).

Type	Cluster				Total
	CL1	CL2	CL3	CL4	
Type_A	160	0	0	0	160
Type_B	0	40	0	0	40
Type_C	1	0	159	0	160
Type_D	0	0	11	29	40
Total	161	40	170	29	400

Using the ROPstat statistical package, we performed 100 independent random replications (see Vargha et al., 2016), and the validation results are summarized in the first two lines of MORI coefficients in Table 2 for the 4-cluster HCA solution.

The obtained MORI values indicate that, as measured by the seven QCs, the 4-cluster solution based on the MEA data set was significantly and substantially better than those resulting from the analyses of random data, supporting the internal validity of that cluster solution for all QCs (see Vargha et al., 2016). The MORI-values are somewhat smaller for the random permutation data set than for the independent random normal data set, due to the combined effect of non-normality and the correlation pattern of the input variables.

An additional, more stringent, random control procedure was also applied and is described below. If the joint distribution of the four input variables were multidimensional normal, there would be one single center in the distribution, which in turn would exclude a multi-cluster structure, and would render senseless the search for a natural clustering. For this reason, it is of interest to confirm that the cluster solution is significantly and substantially better than a cluster solution that is based on a random multivariate normal data set where the intercorrelations are the same as for the MEA data set variables. This is also an (overly) strict test that a real clustering structure has been found, in that it reflects a pattern structure that cannot be explained by the pairwise relationships between the variables. The latest development of the ROPstat software includes such a test. To create the appropriate data set for the test, ROPstat first generates the required number of independent random normal variables and then transforms these data with an orthogonal rotated factor loading matrix based on the MEA data set. In this way, the intercorrelations of the transformed random variables will equal those of the MEA data set variables (these correlations are given in Table 5). Then we performed the MORI analyses based on the correlated random data and the results are summarized in the last row of Table 2, confirming that our obtained 4-cluster solution is substantially better – in terms of all QCs – than what would be expected in a random data set of correlated normally distributed variables.

Our clustering method may be claimed to be inferior because we do not have a clear-cut method for identifying the optimal cluster number. To address this issue, we carried out a series of validating simulations to compute MORI values also for cluster numbers 3 and 5. For all QCs, $k = 4$ was the best solution, sometimes (GDI24) with a very high advantage.

An indication of a successful CA is that the correlations between the variables within the clusters are close to zero (\approx local independence). This was the case for the $k = 4$ solution. Despite the high intercorrelations of the input variables (see Table 5), the correlations computed

separately for the different clusters were not significantly different from zero. The corresponding p -values obtained by Bartlett's chi-square statistic for testing the independence of the variables were .683, .363, .303, and .069, respectively, showing that local independence of the obtained cluster solution cannot be rejected.

Table 5

Pearson correlation matrix of the variables in the MEA data set.

Variables	Var1	Var2	Var3	Var4
Ranvar1	1	0.547**	0.739**	0.600**
Ranvar2	0.547**	1	0.555**	0.701**
Ranvar3	0.739**	0.555**	1	0.669**
Ranvar4	0.600**	0.701**	0.669**	1

Notation: **: $p < .001$

The situation when studying the validity of the classification of the MEA data set is unusual in that we know what the theoretical true classification and types are (they are given in the TEO data set, see Figure 1). This enables us to study the external validity of the obtained 4-cluster solution with regard to whether the theoretical types are reproduced in the analysis of the MEA data set. In Figure 2, the cluster centroids obtained in the analysis of the MEA data set are presented and they are very similar to the theoretical types. More formally, this is confirmed by a pairwise matching of the theoretical types to the cluster centroids, using averaged squared Euclidian distances (ASEDs) between them as inverted measures of similarity (Table 3). The ASEds are very small, indicating a very good match. We can conclude that the true types have been successfully reproduced by the 4-cluster structure, despite the substantial proportion ($100 - 74.3 = 25.7\%$) of measurement error introduced in the MEA data set.

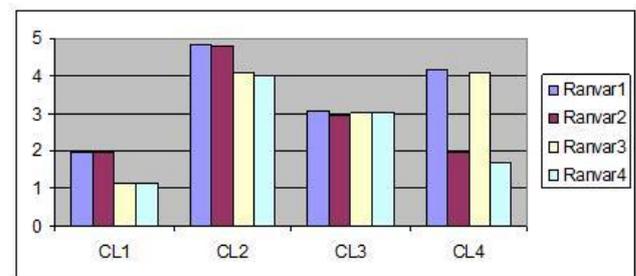


Figure 2. The obtained 4-cluster centroids in the analysis of the MEA data set (raw means).

A second aspect of the validity of the MEA data set 4-cluster solution is the degree to which its objects have been classified into groups in the same way as in the theoretical, true TEO data set. To a high degree this is the case, as shown by the following high validity coefficients: Cramér's contingency $V = .943$, Jaccard index = .917, Rand

index = .970, and Adjusted Rand index = .934. The high degree of correspondence between TEO type membership and MEA cluster membership is also apparent in Table 4.

A model-based approach

If it is believed that the investigated population might constitute a mixture of several subpopulations, each characterized by its specific variable profile (type), a model-based clustering (MBC) approach is often used. It is then usually assumed that the theoretical distribution of the set of observed input variables form a mixture of simpler unimodal (most commonly normal) multidimensional component distributions. From this perspective, the basic aims of MBC is

- (a) to identify the number of components (k);
- (b) to estimate the densities of the k components.

Several types of MBCs are easy to run in well-known statistical software. The most comprehensive among them is the MCLUST program in R (Fraley & Raftery, 2002), by means of which 10 different models allowing for variations in volume (cluster size proportion), shape (shape of cluster distributions determined by their covariance structures), and orientation for a series of cluster numbers (see Fraley & Raftery, 2003, Table 1). This set of models includes also the equal-volume spherical variance model underlying Ward's HCA method if multidimensional normal component distributions are assumed (Fraley & Raftery, 2003).

Technically the MCLUST assumes the normality of the components and performs a HCA with a special classification likelihood measure (see Fraley & Raftery, 2003, formula (10)) followed by an Expectation-Maximization (EM) relocation, for each of the 10 possible models for all cluster numbers below a certain upper limit. The suggested cluster model (including also the specification of k) is the one for which the value of a Bayesian-type information criterion (BIC) is the largest.

We note that the latent profile analysis (LPA) of the Mplus program is practically a special model type of MCLUST where the cluster sizes are allowed to differ but all cluster distributions are normal with a diagonal covariance matrix (local independence; see Muthén, 2001).

In order to test the usefulness of the model-based approach for analyzing the MEA data set we performed an MCLUST analysis of these data. The generation of our random data set was based on independent equal random normal errors, which – taking into account the variable proportions of the four types – corresponds to a spherical, variable volume, equal shape (VII) normal model. However, the rounding and the truncations due to ceiling and bottom effects yielded a biased normal model, which can be regarded as rather typical in practice. Running MCLUST the results indicated that the best model was an ellipsoidal, equal volume and shape (EEV) model with 6 components.

The second best model was ellipsoidal, variable volume, equal shape (VEV) with 2 components, and the third a spherical, equal volume (EII) model with 5 components. We have to conclude that none of them equals to the true VII model with 4 components, which is a disappointing result from using the model-based approach. It is clear that, for these artificial data, the cluster analysis-based procedure was much more successful in finding the true typological structure.

Results from analyses of the empirical data set (EMP data set)

Cluster analysis-based approach

As a pre-analysis of the EMP data set, a residue analysis of outliers was carried out, following the LICUR rationale, and with a standard cut-off point based on the closest neighbor (see Bergman, 1988; Bergman et al, 2003). The results showed that there were no outliers to be dropped. The cluster analysis of the EMP data set was carried out using the same procedure that was used and described in the preceding section. The three variables in the value profile were not standardized because they are on comparable scales. The results of the hierarchical cluster analysis in terms of QCs are presented in Table 6. It appears that k = 5 to 8 are promising cluster numbers. Comparing the cluster centroids of the different solutions, we accepted the 8-cluster solution (bold type face in Table 6). The theoretically expected high Motor Restlessness – Lack of Concentration type emerges first when moving from the 6-cluster to the 7-cluster solution but the worsening of HCmax from k = 8 (0.98) to k = 7 (1.45) is hardly acceptable (it indicates that, reducing the number of clusters from eight to seven, two reasonably homogeneous and theoretically meaningful clusters were merged into one heterogeneous cluster). We could further improve the QCs via relocation allowing EESS% to increase from 81.32 to 83.07 (see the last row of Table 8). Therefore, we finally retained the 8-cluster k-means cluster solution, and its centroids are presented in Figure 3. Among the eight clusters, the three expected types did emerge, namely one cluster with generalized adjustment problems (CLE8), one cluster high only in Motor restlessness and Concentration difficulties (CLE6), and one cluster with generalized good adjustment (CLE1).

We then computed MORI coefficients for the 7 QCs (see previous section) to find out whether our cluster solution is significantly and substantially better than a solution obtained on a random data set of the same size, with the same number of variables, and same number of clusters. One hundred independent random replications were performed (see Vargha, Bergman, & Takács 2016). The obtained MORI values indicate an acceptable but moderate internal validity level of the 8-cluster k-means solution for almost all QCs, even for the most demanding control (correlated random normal variables, see Table 7).

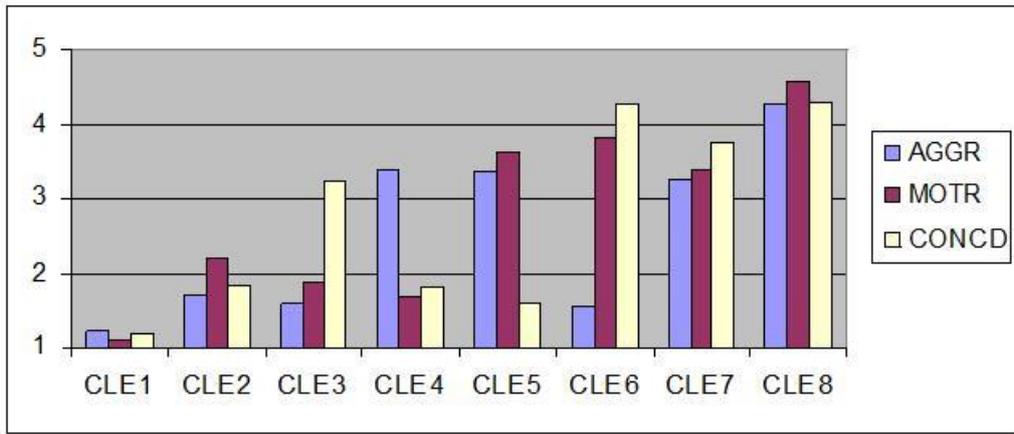


Figure 3. The 8-cluster k-means solution centroids in the analysis of the EMP data set (raw means)

Table 6

Results of hierarchical cluster analyses for cluster numbers between $k = 2$ and $k = 10$ carried out on the EMP data set. Six cluster quality coefficients are presented.

k	EESS%	Quality coefficient				HCmean	HC range
		PB	XBmod	SC	SC		
10	84.36	0.382	0.507	0.657	0.514	0.24-0.98	
9	82.96	0.387	0.463	0.660	0.558	0.24-0.98	
8	81.32	0.407	0.606	0.660	0.610	0.24-0.98	
7	79.20	0.407	0.562	0.650	0.678	0.24-1.45	
6	76.98	0.406	0.515	0.623	0.749	0.24-1.55	
5	73.46	0.435	0.473	0.586	0.862	0.24-1.55	
4	69.20	0.437	0.388	0.635	0.997	0.24-1.75	
3	62.66	0.443	0.402	0.633	1.206	0.24-1.75	
2	51.48	0.659	0.821	0.782	1.564	1.50-1.75	

Table 7

MORI coefficients of 7 QCs for the chosen 8-cluster k-means solution using three types of random control data sets (first three rows), and for the best model-based solution (last row) (number of independent random replications = 100).

Type of random control data set	EESS%	PB	XBmod	SC	HCmean	CLdelta	GDI24
Random permutation	0.29	-0.02	0.23	0.12	0.28	-0.13	0.15
Independent normal	0.49	0.14	0.11	0.31	0.49	0.12	-0.24
Correlated normal	0.22	0.17	0.20	0.27	0.21	0.24	0.07
Correlated normal, MCLUST EII, $k = 8$	0.18	0.22	0.30	0.19	0.17	0.28	0.20

Note: all coefficients above 0.07 are significantly greater than 0 at the $p < .001$ level.

Table 8

Basic quality characteristics of four MBC solutions (rows 1-4), and the 8-cluster k-means solution (last row)

Clustering method	EESS%	PB	XBmod	SC	HCmean	CLdelta	GDI24	HC range
MBEEV9	66.18	0.377	-0.389	0.345	1.104	0.899	0.189	0.00-1.75
MBEEV8	73.11	0.374	0.240	0.498	0.877	0.923	0.462	0.62-1.23
MBEII5	76.00	0.476	0.538	0.646	0.779	1.077	0.580	0.44-1.44
MBEII8	82.20	0.446	0.628	0.658	0.582	1.079	0.688	0.39-1.12
kCA8	83.07	0.409	0.573	0.692	0.554	1.052	0.617	0.28-1.03

Table 9

Stepwise LRA on dichotomized crime with the 8 cluster indicator variables as independent variables.

Depvar: YDCrime		B	Sig.	Exp(B)
Nagelkerke R ² = 0.174	CLE1I	-1.745	.001	.175
	CLE6I	1.277	.010	3.587
	CLE8I	1.421	.000	4.139
Depvar: ADCrime		B	Sig.	Exp(B)
Nagelkerke R ² = 0.097	CLE1I	-1.132	.014	.322
	CLE8I	1.223	.000	3.398

Table 10

Stepwise LRA on dichotomized crime with the 3 continuous IDA variables as independent variables.

Depvar: YDCrime		B	Sig.	Exp(B)
Nagelkerke R ² = 0.178	MOTR	.420	.002	1.522
	CONCD	.363	.008	1.438
Depvar: ADCrime		B	Sig.	Exp(B)
Nagelkerke R ² = 0.097	MOTR	.539	.000	1.714

An additional indication that the cluster analysis was rather successful is that, despite the high (.49-.68) pairwise correlations of the three EMP data set variables, except for the first two clusters (CLE1 and CLE2) the Bartlett test of overall independence was not significant or just marginally significant, whereas in the total sample the test was highly significant ($p < .0001$).

To find evidence of external validity, cluster membership was related to registered crimes below 21 years (YDCrime, coded "0" for 0-2 crimes and coded "1" for more than 2 crimes) and related to registered crimes above 20 years (ADCrime, coded "0" for 0-2 crimes and coded "1" for more than 2 crimes). It was expected that membership in the cluster with generalized adjustment problems or in the cluster high in both Motor Restlessness and Concentration difficulties would be related to high criminality. The external validity was studied by using logistic regression analysis (LRA) with a dichotomized crime variable as the dependent variable and the eight dichotomized cluster indicator variables (denoted CLE1I, CLE2I, etc.) entered

stepwise in the regression equation using the forward LR method (level of entry was $p=.05$). The results are presented in Table 9. The Nagelkerke R² values show that young dichotomized crime (YDCrime) was better predicted by means of the cluster indicator variables than adult dichotomized crime (ADCrime). In both cases, cluster indicator variables CLE1I and CLE8I were significant, and when predicting YDCrime also CLE6I. Hence, not only did the three theoretically expected types emerge as clusters, these clusters were also related to high criminality.

To compare the predictive power of the information about cluster membership to the predictive power of the three original continuous variables, additional LRAs were carried out with a crime variable as the dependent variable and the continuous variables entered stepwise as independent variables (level of entry was $p = .05$), see Table 10. Again, the Nagelkerke R² values show that young dichotomized crime (YDCrime) was better predicted than adult dichotomized crime (ADCrime). Motor restlessness emerged as the best predictor. Comparing the Nagelkerke

R^2 values of Table 9 and Table 10 we can conclude that both YDCrime and ADCrime were similarly well predicted by means of the binary cluster indicator variables and the continuous IDA variables.

An interesting question is whether the variable-oriented or person-oriented information adds to the predictive power achieved by the other type of information. We therefore performed two additional analyses, one where the three continuous variables were added stepwise to the best equation achieved by the cluster membership variables (i.e. those retained in the final step of the stepwise analysis of the eight cluster variables), and another analysis where the eight cluster membership variables were added stepwise to the best equation achieved by the continuous variables (i.e., those retained in the final step of the stepwise analysis of the three continuous variables). The results indicated that the LRA model obtained first, using one type of variables, was not significantly improved by allowing the other type of variables entering the regression model. Our conclusion is that the person-oriented summary of the data in the form of a typology retains the predictive power of the original variables - but it does not significantly improve it.

A model-based approach

In order to test the usefulness of a model-based approach for analyzing the EMP data set we also performed an MCLUST analysis of these data. Running MCLUST the results first indicated that the best model – having the largest BIC values – was an ellipsoidal, equal volume and shape (EEV) model with 9 components. However, inspecting closely this solution it turned out that for 3 out of the 9 clusters the cluster sizes were small (2, 4, and 11), and the ellipsoidal forms with varying orientations proved to be hardly interpretable. In addition, the BIC values of the different EEV solutions between $k = 2$ and $k = 12$ were unstable, strongly oscillating, indicating that the EEV type is unstable, which was also confirmed by the unacceptably low QC levels of the 9- and 8-cluster EEV solutions (see MBEEV9 and MBEEV8 in Table 8). This evaluation of the MBC solutions was performed using the Validation module of ROPstat. We then turned to the second best solution type, which was the spherical, equal volume and shape (EII) model, where the $k = 9$ and $k = 5$ solutions had the largest BIC values (see Figure 4). However, in the $k = 9$ solution one cluster was empty, thus it was in fact an 8-cluster structure that was similar in many respects to the 5-cluster solution. Most importantly, all clusters of the $k = 5$ solution appeared also in the $k = 8$ solution – the differences of the corresponding cluster centroids were .001, .007, .029, .043, and .056, respectively. Since the HC-range of the $k = 5$ solution (0.44-1.44) indicated the existence of at least one very heterogeneous cluster, we preferred the $k = 8$ solution. This solution is similar to our 8-cluster solution presented in the previous section, and it is reflected by the following high validity coefficients: Cramér's contingency coefficient

$V = .704$, Jaccard index = .580, Rand index = .893, and Adjusted Rand index = .602. Comparing the centroid structures of our 8-cluster solution and this 8-cluster MCLUST solution, two centroid pairs (including our theoretically important CLE1 and CLE8 clusters) were very similar (with $ASED \leq .01$), and three others moderately similar (with $ASED \leq .08$). Computing the QCs (see the MBEII8 row in Table 8) and the MORI coefficients (see the last row in Table 7) for the 7 QCs of the MCLUST solution, we obtained somewhat poorer values for the cohesion QCs (EESS% and HCmean) and SC, and somewhat better values for the other QCs (XBmod and GDI24) that give more weight to the distance of the closest two clusters.

To sum up, considering that our previously described 8-cluster k-means solution has more homogeneous classes, and acknowledging that a moderate similarity of some clusters (like CLE1 and CLE2, or CLE2 and CLE3, see Figure 3) is not incompatible to our person-oriented model, it might be regarded as marginally superior to the 8-cluster MCLUST solution (MBEII8).

Discussion

Consider a research context in which a person-oriented theoretical framework is appropriate. In the presentation of this framework, we argued that a typological methodological approach is then often a natural choice because it preserves the information contained in the whole profile of values in the variables. This profile reflects “whole system” properties that are lost if a standard variable-oriented methodological approach is applied where the profile information is “atomized” and the pieces (variables) are studied as separate units in the analyses, disregarding the information provided only by the profile as a whole. A call was made for the revival of a typological approach, using newer methods of analysis that are less subjective than those used in the heyday of the typological approach in Psychology. The purpose of the present paper was to discuss and exemplify the use of some such methods and to examine the findings they produced for two data sets, one artificial “prototype” data set and one empirical data set. In doing this, we also compared the usefulness of the findings produced by a sound partly explorative classificatory analysis using cluster analysis to the findings produced by a standard model-based method.

In the Results section, we first analyzed an artificial data set ($n=400$), constructed so that it exemplifies a typological structure that may not be uncommon in real life (each individual truly belongs to one of four distinct value profiles/types but the variables forming the profile contain errors of measurement with $r_{it}=.80$). The strength of using such an artificial data set is that we know the true typological structure and can examine how well the different methods for classification analysis succeed in finding this structure when moderate errors of measurement are present.

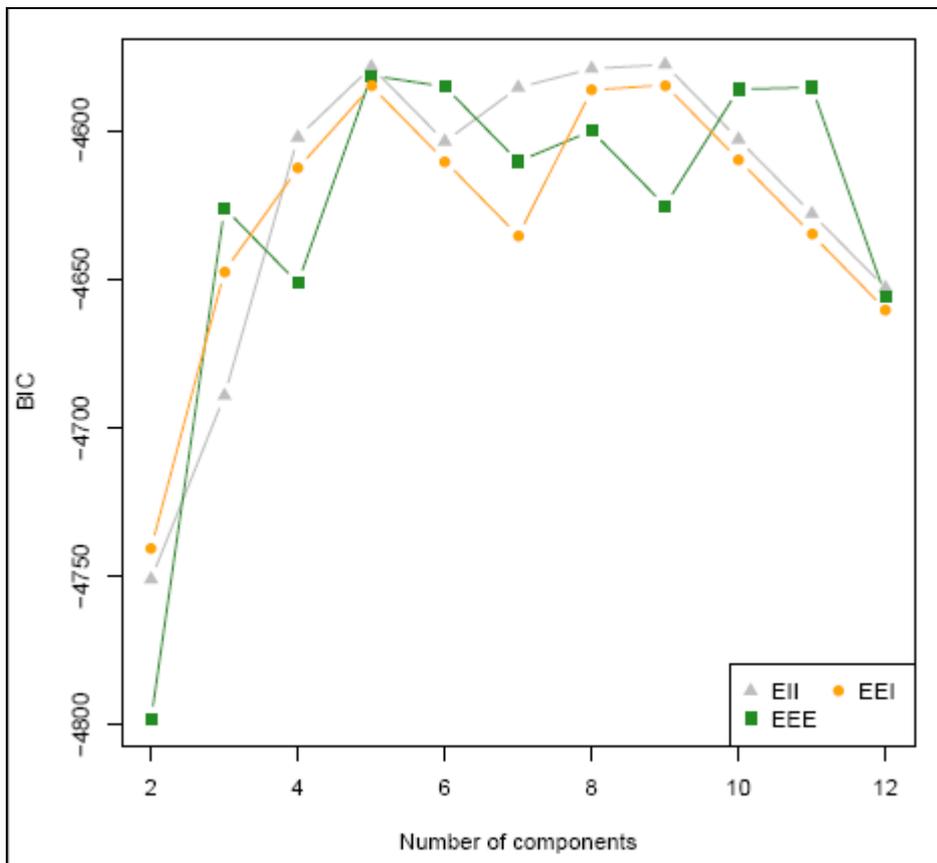


Figure 4. The BIC values for cluster numbers between 2 and 12, for three types (EII, EEI, EEE) of MBC solutions in the analysis of the EMP data set

The basic finding was that the cluster analysis-based method did quite well, both in identifying the four types and in ascribing the correct type membership to the studied individual objects. The model-based method (MCLUST, Fraley and Raftery, 2003) performed surprisingly badly and did not identify the four types; instead the solution that according to the method was best was a clearly suboptimal 6-class solution. This finding was unexpected and it may have been caused by that the four variables in the profile were discrete, taking only the values 1, 2, 3, 4, or 5, both for original true scores and for the analyzed scores after errors of measurement were added. Hence, the measurement errors were not really independently normal. However, deviations from normality are common for real data (see e.g. Micceri, 1989) and our finding exemplifies a possible limitation in the usefulness of model-based methods that are based on standard assumptions. Of course, the artificial data set was only a single simulation but we performed an additional one and similar results were found. A nice property of the artificial data example we used is that the interested readers can easily create their own data set of the

same type and examine how their preferred method performs. It is also possible to omit the rounding to integers we used and instead analyze continuous data with independently normally distributed measurement errors to verify that a method performs well in this ideal case.

In the Results section, we also analyzed an empirical data set with just three variables measuring different aspects of externalizing adjustment problems ($n = 541$). All variables were coded 1 (no adjustment problem) to 5 (extreme adjustment problem). The cluster analysis-based method suggested that an 8-cluster solution was preferable. It had acceptable values in a number of internal validity coefficients (quality coefficients), the three theoretically expected clusters were reproduced, the solution was significantly better than what would have been found, had random data of the same type been analyzed, and the solution showed a moderate degree of external validity in predicting registered criminality. The predictive power of cluster membership was also about the same as that when the three variables in the profile were used as continuous variables. This suggests that no information of predictive

value that the three “continuous” variables contained was lost by replacing them with the information contained in just one categorical cluster variable taking eight values. For this data set, the model-based method with an automatic BIC-value based decision led again to an obviously poor model (EEV with $k = 9$), but carefully analyzing the BIC value patterns for several model types (see Figure 4) we finally found a solution (EII with $k = 8$) that appeared satisfactory according to different criteria. It happened to be in most respects quite similar to our 8-cluster solution. However, our conclusion is that the 8-cluster solution was marginally better.

As pointed out above, variables that are non-normally distributed are quite common in empirical data, and are often characterized by skewed distributions (Micceri, 1989). We believe that the simple examples of findings we presented are sufficient to suggest that the restrictive assumptions that model-based classification analysis usually is based on may produce biased findings in a number of practical applications. In many cases, it might be informative to check the findings from a model-based classification by applying in parallel another method of classification that makes less assumptions about the properties of the data to be analyzed. This highlights the issue of the limited usefulness of “big” classification models in situations characterized by complex multivariate data that do not match standard assumptions about model properties. As pointed out in the presentation of the person-oriented theoretical framework, such an ambitious modeling approach might often be premature; a caveat that receives some support by that the power of rejecting a false classification model appears to be rather low for many model-based methods under fairly normal conditions (e.g. moderate sample sizes, not very distinct classes, see Dziak, Lanza, & Tan, 2014; Tein, Coxe, & Cham, 2013). It should also be pointed out that model-based clustering, as it is usually applied, is not a fully automatic process. Like when applying a sound clustering procedure, some subjective decisions are often involved in selecting the best classificatory structure.

In all classification analyses based on profile data, the issue of how to measure (dis)similarity arises and there is no generally best way to do it. Our choice of the averaged squared Euclidean distance was a sound choice for the type of data we cluster analyzed; for many other types of data, other types of (dis)similarity measures are more appropriate. For instance, the correlation coefficient could be used if the relevant characteristic of the profile is profile form, not profile level. Of course, the choice of classification method should also be aligned to the specific problem and data that are analyzed, and for a thoughtful discussion of these issues the reader is referred to, for instance, von Eye, Mun, and Indurkha (2004).

It is a limitation that both data sets analyzed in this paper contained discrete data with variables taking just five values. An alternative strategy for analyzing these data is to

apply the powerful tool of configural frequency analysis (von Eye & Pena, 2004) that is well adapted to analyzing discrete data of this type. Space limitations prevented us from adding such analyses but it would be of interest to do so and compare the findings to those we presented.

In this paper, we discussed only the comparatively simple case of classifying individuals based on a value profile from a single time point of measurement. A very important case that was not treated is the more complex issue of the identification of *developmental* types. Due to the frequently high complexity of developmental profiles (which often tend to contain very many variables), it is usually not suitable to include all the variables in a single “big” profile analyzed in a single classification analysis. The resulting classes/clusters then often emerge as highly heterogeneous groups, without much explanatory value and that are hard to interpret. It is often a more sound strategy to first perform classifications at each age separately and then connect class memberships across time (the LICUR approach, see Bergman et al, 2003). However, for this approach to produce clear and interpretable findings it is important that each classification has been successful in the sense that the major “real” types have been identified at each age, and it is also important that, to a large extent, each individual has been ascribed to the appropriate cluster/class. If not, “true” individual development in the form of changing/constant type membership will be obscured by two types of errors (errors in *general* type identification and errors in *individual* type identification). Hence, in a developmental context, applying a sound classification methodology and carefully validating the resulting classifications is even more important than it is in the cross-sectional case discussed in this paper. Unfortunately, in our experience, this is far from always the case.

Acknowledgment

The preparation of the present article was supported by the National Research, Development and Innovation Office of Hungary (Grant No. K 116965).

References

- Bergman, L.R. (1988). You can't classify all of the people all of the time. *Multivariate Behavioral Research*, 23, 425-441.
- Bergman, L. R. & Andersson, H. (2010). The person and the variable in developmental psychology. *Journal of Psychology*, 218 (3), 155-165. DOI: 10.1027/0044-3409/a000025
- Bergman, L. R., Corovic, J., Ferrer-Wreder, L., & Modig, K. (2014). High IQ in early adolescence and career success in adulthood: Findings from a Swedish longitudinal study. *Research in Human Development*, 11, No. 3, 165-185. DOI: 10.1007/s12124-009-9102-2

- Bergman, L. R. & Lundh, L.-G., Eds. (2015). The person-oriented approach: Roots and Roads to the Future. *Journal for Person-Oriented Research, 1, Special issue 1-2*, 1-109. DOI: 10.17505/jpor.2015.01
- Bergman, L. R. & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9*, 291-319. DOI: 10.1017/S095457949700206X
- Bergman, L. R., Magnusson, D., & El-Khoury, B. M. (2003). *Studying individual development in an interindividual context. A Person-oriented approach*. Mahwa, New Jersey, London: Lawrence-Erlbaum Associates.
- Bergman, L. R. & Vargha, A. (2013). Matching method to problem: A developmental science perspective. *European Journal of Developmental Psychology, 10, No. 1*, 9-28. DOI: 10.1080/17405629.2012.732920
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (rev. ed.)*. New York: Academic Press.
- Desgraupes, B. (2013). Clustering Indices. University Paris Ouest, Lab Modal'X, April 2013. <https://cran.r-project.org/web/packages/clusterCrit/vignettes/clusterCrit.pdf>. Downloaded: August 28, 2015.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical Methods in Education and Psychology* (3rd edition). Boston: Allyn & Bacon.
- Dziak, J. J., Lanza, S. T., & Tan, X. (2014). Effect size, statistical power and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. *Structural Equation Modeling, 21 (4)*, 534-552. DOI: 10.1080/10705511.2014.919819
- Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis and density estimation. *Journal of the American Statistical Association, 97*, 611-631. DOI: 10.1198/016214502760047131
- Fraley, C., & Raftery, A. E. (2003). Enhanced software for model-based clustering, density estimation, and discriminant analysis: MCLUST. *Journal of Classification, 20*, 263-286. DOI: 10.1007/s00357-003-0015-3
- Magnusson, D. (1988). *Individual development from an interactional perspective: A longitudinal study*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Magnusson, D., & Törestad, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology, 44*, 427-452.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*, 156-166.
- Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika, 45*, 325-342.
- 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 and Perspectives, 2(4)*, 201-218. DOI: 10.1207/s15366359mea0204_1
- Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (eds.), *New Developments and Techniques in Structural Equation Modeling* (pp. 1-33). Lawrence Erlbaum Associates.
- Richters, J. E. (1997). The hubble hypothesis and the developmentalist's dilemma. *Development and Psychopathology, 9(2)*, 193-229.
- Tein, J.-Y., Coxe, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling, 20(4)*, 640-657. DOI: 10.1080/10705511.2013.824781
- Vargha, A., & Bergman, L. R. (2015). Finding typical patterns in person-oriented research within a cluster-analytic framework using ROPstat. *Conference on Person-Oriented Research. May 8 and 9, 2015*, Vienna, Austria.
- Vargha, A., Bergman, L. R. & Takács, Sz. (2016). Performing cluster analysis within a person-oriented context: Some methods for evaluating the quality of cluster solutions. *Journal for Person-Oriented Research, 2 (1-2)*, 78-86. DOI: 10.17505/jpor.2016.08.
- Vargha, A., Torma, B. & Bergman, L. R. (2015). ROPstat: a general statistical package useful for conducting person-oriented analyses. *Journal for Person-Oriented Research, 1 (1-2)*, 87-98. DOI: 10.17505/jpor.2015.09
- von Eye, A., & Bergman, L.R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15*, 553-580.
- von Eye, A., Mun, E. Y., & Indurkha, A. (2004). Classifying developmental trajectories – a decision making perspective. *Psychology Science, 46*, 65-98.
- von Eye, A., & Pena, G. E. (2004). Configural frequency analysis: The Search for Extreme Cells. *Journal of Applied Statistics, 31(8)*, 981-997.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*, 236-244.
- Wiedermann, W., Bergman, L. R., & von Eye, A., Eds. (2016). Development in methods for person-oriented analysis. *Journal for Person-Oriented Research, 2, Special issue 1-2*, 1-12. DOI: 10.17505/jpor.2016.01
- Xie, X. L., Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13 (4)*, 841-846.