

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Mäkikangas, Anne; Tolvanen, Asko; Aunola, Kaisa; Feldt, Taru; Mauno, Saija; Kinnunen, Ulla

Title: Multilevel Latent Profile Analysis With Covariates : Identifying Job Characteristics Profiles in Hierarchical Data as an Example

Year: 2018

Version: Accepted version (Final draft)

Copyright: © The Author(s) 2018

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Mäkikangas, A., Tolvanen, A., Aunola, K., Feldt, T., Mauno, S., & Kinnunen, U. (2018). Multilevel Latent Profile Analysis With Covariates : Identifying Job Characteristics Profiles in Hierarchical Data as an Example. *Organizational Research Methods*, 21(4), 931-954.
<https://doi.org/10.1177/1094428118760690>

Multilevel Latent Profile Analysis with Covariates:
Identifying Job Characteristics Profiles in Hierarchical Data as an Example

Anne Mäkikangas, Asko Tolvanen, Kaisa Aunola,
Taru Feldt, Saija Mauno, & Ulla Kinnunen

Address correspondence to:

Anne Mäkikangas, PhD, Adjunct Professor

Faculty of Social Sciences, Psychology, 33014 University of Tampere, Finland.

E-mail: anne.makikangas@uta.fi

Asko Tolvanen, PhD, Professor

Methodology Centre for Human Sciences, P.O. Box 35, 40014 University of Jyväskylä,
Finland.

E-mail: asko.tolvanen@jyu.fi

Kaisa Aunola, PhD, Professor

Department of Psychology, P.O. Box 35, 40014 University of Jyväskylä, Finland.

E-mail: kaisa.aunola@jyu.fi

Taru Feldt, PhD, Adjunct Professor

Department of Psychology, P.O. Box 35, 40014 University of Jyväskylä, Finland.

E-mail: taru.feldt@jyu.fi

Saija Mauno, PhD, Professor

Department of Psychology, P.O. Box 35, 40014 University of Jyväskylä, Finland
& Faculty of Social Sciences, Psychology, 33014 University of Tampere, Finland.

E-mail: saija.mauno@jyu.fi

Ulla Kinnunen, PhD, Professor

Faculty of Social Sciences, Psychology, 33014 University of Tampere, Finland.

E-mail: ulla.kinnunen@uta.fi

Biographical sketches

Anne Mäkikangas, PhD, adjunct professor in work psychology, is currently a university lecturer at the University of Tampere. Mäkikangas has substantial expertise in person-centered methodology, which she has successfully applied in occupational well-being research. She was awarded the EAOHP Andre Bussing Memorial Prize in 2014 for her high-quality person-centered research.

Asko Tolvanen, PhD (Statistics) is a professor in the Methodology Center for Human Sciences, University of Jyväskylä. He is knowledgeable on multilevel and mixture modeling which he has applied and developed for research purposes. Tolvanen has published extensively in the leading journals of I/O and developmental psychology.

Kaisa Aunola, PhD in psychology, is a professor of psychology at the University of Jyväskylä. Aunola is also an adjunct professor in correlative methods and has special expertise in applying multilevel and mixture modeling in longitudinal and diary studies focusing on family processes.

Taru Feldt, PhD, adjunct professor in work psychology, is a university researcher at the University of Jyväskylä. She has studied occupational well-being from different viewpoints for over twenty years. In recent years, she has specialized in latent profile analysis in longitudinal datasets and multilevel analysis in organizational research.

Saija Mauno, PhD, adjunct professor in work psychology, is a professor at the University of Tampere and a university researcher at the University of Jyväskylä. She has studied occupational well-being and work-family interface over twenty years. She has published both multilevel and person-centered studies of these topics.

Ulla Kinnunen, PhD, is a professor in work and organizational psychology at the University of Tampere. She has experience in longitudinal data analyses, including person-centered methods. Recently, she has focused on examining recovery from work stress.

Acknowledgements

This study was supported by a grant from the Academy of Finland to Anne Mäkikangas (Grant No. 258882), Ulla Kinnunen (Grant No. 124268) and Saija Mauno (Grant No. 124360). We wish to thank Professor Alexandre J.S. Morin for sharing his expertise and consulting with us on the statistical analyses used in this study.

Abstract

Latent Profile Analysis (LPA) is a person-centered method commonly used in organizational research to identify homogeneous subpopulations of employees within a heterogeneous population. However, in the case of nested data structures, e.g., employees nested in work departments, multilevel techniques are needed. Multilevel LPA (MLPA) enables adequate modeling of subpopulations in hierarchical datasets. MLPA enables investigation of variability in the proportions of Level 1 profiles across Level 2 units, and of Level 2 latent classes based on the proportions of Level 1 latent profiles and Level 1 ratings, and the extent to which covariates drawn from the different hierarchical levels of the data affect the probability of a membership of a particular profile. We demonstrate the use of MLPA by investigating job characteristics profiles based on the Job Demand-Control-Support (JDCS) model using data from 1,958 university employees clustered in 78 work departments. The implications of the results for organizational research are discussed, together with several issues related to the potential of MLPA for wider application.

Keywords: multilevel latent profile analysis, clustered data, hierarchical structure, Job Demand-Control-Support model

Multilevel Latent Profile Analysis with Covariates:

Identifying Job Characteristics Profiles in Hierarchical Data as an Example

The aim of this article is to demonstrate the use of multilevel latent profile analysis (MLPA) in the organizational sciences, an area in which it has been seldom utilized. A comprehensive description of this method is given by Henry and Muthén (2010) and Vermunt (2003, 2008, 2011; see also Finch & French, 2014), and thus, the statistical content covered is already familiar to methodological experts. For less statistically knowledgeable organizational researchers interested in applying person-centered methods in a multilevel context, this article may serve as a convenient introduction to how this can be done. To this end, we present and illustrate step by step how several types of MLPA can be applied to answer questions in organizational research. Each step is presented from a user perspective. Moreover, in our illustrative example, continuous variables rather than categorical variables are used, as they are more typically studied in organizational research.

We illustrate MLPA in the context of the well-known and widely used Job-Demand-Control-Support (JDCS) model (Karasek & Theorell, 1990). This job stress model has been widely utilized in occupational stress research (for reviews, see Häusser, Mojzisch, Niesel, & Schulz-Hardt, 2010; van der Doef & Maes, 1999), including recent longitudinal studies applying a person-centered perspective (Igic et al., 2017; Mauno, Mäkikangas, & Kinnunen, 2016). Despite decades of investigation, very little is known about the extent of the role played by individual and environmental factors in shaping JDCS profiles. Instead of estimating the quantitative amount of variance in job characteristics between work departments, as is typically done in multilevel modelling, MLPA enables qualitatively different working environments to be identified and further investigated. Consequently, from the stress theory perspective, MLPA has the potential to provide answers to various novel research questions, such as whether inter-group (e.g., team, work department, organization)

variability exists in within-person job characteristics profiles, whether qualitatively different work environments can be identified, or to what extent experiences of job characteristics are explained by individual (e.g., working hours) and environmental factors (e.g., leadership, climate, resource inadequacy in terms of unit size). Such knowledge could have important practical implications, regarding, for example, whether psychosocial work conditions can best be improved by redesigning jobs, or by treating individuals, or both.

First, we introduce the use of a person-centered approach in the multilevel context in general. We then describe three types of MLPA, namely, 1) Level 2 variations in the size of the Level 1 profiles (also known as the parametric model; Henry & Muthén, 2010); Level 2 profiles (henceforth ‘classes’) based on the relative frequency of the Level 1 profiles (also known as the nonparametric model; Henry & Muthén, 2010); and Level 2 classes based on Level 1 individual ratings. We then illustrate these MPLA models by investigating JDCS profiles in hierarchical data on university employees ($N = 1,958$) working in 78 departments.

Person-centered approach in the single versus multilevel context

A person-centered approach focuses on identifying subpopulations of individuals who show different patterns of values for certain criteria variables (Muthén & Muthén, 2000). In addition, the person-centered approach can also be used to identify profiles that differ from one another at the level of the relations between variables (i.e., mixture regression, factor mixture). Compared with a variable-centered approach, which applies a single model to the whole sample to estimate a single set of parameters, a person-centered approach is able to relax the assumption of homogeneity in the population through the estimation of a distinct set of parameters applicable to the various subpopulations present in the sample.

The concept of finite mixture models refers to statistical methods that probabilistically assign individuals into subpopulations (McLachlan & Peel, 2000). Latent Class Analysis (LCA) and Latent Profile Analysis (LPA) are special cases of finite mixture

models that aim to identify naturally occurring latent subpopulations that are homogeneous in relation to selected criteria variables (McLachlan & Peel, 2000). Although LCA is sometimes presented as a general statistical framework for mixture models, it typically refers to identifying subpopulations based on categorical variables, whereas LPA performs the same function using continuous variables (Oberski, 2016). In the present demonstration, we focus on LPA, which aims to identify latent subpopulations presenting different configurations on a set of indicators (McLachlan & Peel, 2000; Sterba, 2013). Consequently, the covariance structure of the observed variables is explained via differences in their mean values between the latent subpopulations rather than via the associations between continuous variables, as is done in variable-centered modeling.

In organizational research, LPA has previously been applied to identify, for example, different employee profiles in relation to the three symptoms of burnout, i.e., exhaustion, cynicism and reduced professional efficacy (for a review, see Mäkikangas & Kinnunen, 2016), or different mindsets (i.e., affective, normative and continuance) of workplace commitment (for a review, see Meyer & Morin, 2016). Because LPA assumes observations to be independent of each other, it is not a suitable method for use with hierarchical datasets where individuals are nested within groups (e.g., within units/departments/teams in an organization) (Asparouhov & Muthén, 2008; Henry & Muthén, 2010). In addition to violation of the independency assumption of observations, ignoring the presence of nested data structures may lead to inaccuracy in the classification of profiles (e.g., to inaccuracy in the number of latent classes and to biased standard errors; Asparouhov & Muthén, 2008; Chen, Kwok, Luo, & Willson, 2010; Luo & Kwok, 2009; Meyers & Beretvas, 2006; Moerbeek, 2004; Park & Yu, 2016; Vermunt, 2003, 2008). Recent evidence indicates that multilevel modeling is advised in the case of nested data, even if no group-level effects are expected or evident (Bliese, Maltarich, & Hendricks, 2017).

To overcome the obstacles caused by hierarchical data structures, multilevel LPA (MLPA) is needed. The benefit of MLPA is that it enables the different levels of the data to be taken into account in modeling the phenomena of interest, and thus answer research questions which are beyond the scope of single-level LPA. In the organizational sciences, for instance, MLPA enables investigation of the role played by qualitatively different work environments. Examples of such novel questions would be, how do burnout profiles vary across work departments or organizations? Or how can work departments be clustered according to the different proportions of individual burnout profiles? In multilevel modeling, the associations between variables are modeled at two or more levels (e.g., at the level of the individual and at the level of the work department). In variable-centered multilevel modeling the associations of variables are modeled at different levels (e.g., between job characteristics and turnover intentions at different hierarchical levels of the data; Mauno, De Cuyper, Tolvanen, Kinnunen, & Mäkikangas, 2014). In contrast, MLPA explains the observed covariance structure through latent profiles and relates the information to latent profiles at different levels of the data (e.g., to predict the proportions of profiles at different levels).

MLPA can be specified in many ways. Here, we focus on three specific types of MLPA. In the first model, *Level 2 variations in the size of Level 1 profiles*, the variation in the relative size of single-level profiles across the higher-level units (i.e., Level 2) is modeled. In the second model, *Level 2 classes based on the relative frequency of the Level 1 profiles*, latent classes of higher-level units (i.e., Level 2) based on the proportions of the single-level latent profiles are investigated. In the third model, *Level 2 classes based on Level 1 ratings*, latent classes of higher-level units are based on the variables observed at Level 1. In the following, we use the term ‘profile’ to refer to latent subpopulations at the level of individuals (Level 1) and the term ‘class’ to refer to subpopulations at the level of work

departments (Level 2). Next, we explain in detail the steps to follow in constructing these models.

MLPA: Steps and equations

Modeling MLPA follows six distinguishable steps: 1) definition of the problem and selection of the variables, 2) preliminary analysis of the data, 3) model specification, 4) model estimation for 1, 2,...,K latent profile (Level 1)/class (Level 2) solutions, 5) choosing the best-fitting model, and 6) investigating the role of potential covariates.

In the first step, the definition of the problem to be solved and selection of the variables are guided by the research target and questions. MLPA, like other person-centered methods, is typically conducted in an exploratory manner, that is, a priori hypotheses for the number or nature of the latent upper-level classes present in the data may not be given (Finch & French, 2014; see also Meyer & Morin, 2016; Morin & Wang, 2016). However, predictions of the characteristics of possible subpopulations can be theoretically anchored, keeping in mind that person-centered analysis is more model- than theory-driven (for exceptions, see Finch & Bronk, 2011; Wang, 2007). Thus, instead of a theory-driven definition of the problem, the theoretical rationale for person-centered studies is typically related to the following questions: Why is there a need for the identification of possible latent subpopulations in relation to the variables of interest? Why can we expect that the identification of these profiles will improve our understanding of the phenomenon under study? Do we expect these profiles to have a practical impact on various covariates?

After defining the problem and selecting the relevant variables, the second step, preliminary examination of the data, follows. This includes, for example, checking for possible outliers (e.g., by using Mahalanobis distance¹), evaluating the reliability of the variables, and conducting missing data analysis. In implementing MLPA, it is also necessary to calculate Intraclass Correlation Coefficients (ICCs) to see whether and how strongly the

responses given by the members of the same subpopulation are related (Bliese et al., 2017). A significant ICC(1), indicating the degree to which individuals within a designated unit (i.e., work department) are non-independent (Bliese, 2000), is often regarded as an essential precursor for multilevel modeling. However, recent evidence indicates that multilevel models are appropriate even in the case of a low or non-significant ICC(1) (Bliese et al., 2017). In our demonstration, ICCs are calculated by dividing the variance between work departments with the total variance (variance within + variance between work departments), which corresponds to ICC(1).

The third step, i.e. specification of the model, needs to be well-aligned with the research question and with current knowledge on best practices. In this stage, it is decided which of the model parameters (e.g., means, variances) are to be freely estimated and which are to be fixed/constrained at different levels of the data (Asparouhov & Muthén, 2008; Vermunt, 2011). Simulation studies conducted for different mixture models (i.e., LPA, growth mixture modeling, mixture regression modeling) suggest that to achieve a less biased and proper representation of the data, the variances across profiles should be freely estimated (Diallo, Morin, & Lu, 2016; Enders & Tofighi, 2008; Kim, Lamont et al., 2016; Peugh & Fan, 2013). However, despite this generic recommendation, it is not always possible to rely on the results of models in which both means and variances are freely estimated. The reason for this is that these models tend more frequently to converge on improper solutions, which is generally taken to indicate that the model might have been over-parameterized and that a more parsimonious model should be preferred (Bauer & Curran, 2003; Chen, Bollen, Paxton, Curran, & Kirby, 2001; Diallo et al., 2016). It should be also noted that LPA relies on a conditional independence assumption (i.e., variables within latent profile are uncorrelated). Consequently, covariances within profiles should only be estimated with great caution and on the basis of a strong theoretical justification (see Meyer & Morin, 2016). In MLPA, raw data

or grand-mean centering is recommended, as they both provide the same results. To aid interpretation of the results, either grand-mean centering could be applied or the variables could be standardized for the analysis. Group-mean centering, however, should not be used in MPLA because it can substantially change model interpretation (see also Bliese et al., 2017).

To conclude, model specification and tests of alternative models should be conducted after careful reasoning and described in detail. As stated earlier, we present three possible ways to specify MLPA, each of which, depending on the research question, can be implemented separately. However, two of these (Level 2 variations in the size of the Level 1 profiles and Level 2 classes based on the relative frequency of the Level 1 profiles) are based on single-level profiles, and therefore it is recommended to conduct single-level LPA before MLPA.

The model for Level 2 variations in the size of the Level 1 profiles is performed to investigate whether the proportional distributions of the single-level profiles vary across the upper-level units. The model equation with two levels in which latent profiles $j = 1, 2, \dots, J$ (across individuals i in $k = 1, \dots, K$ variables) appear at Level 1 and the proportions of the latent profiles vary randomly across Level 2 units u is

$$\text{Level 1: } y_{ijk_u} = \mu_{jk} + \varepsilon_{ik} \quad \varepsilon_{ik} \sim N(0, \sigma_k^2)$$

Probability of latent class j $P(y_{ijk_u}) = p_{ju}$ is associated to unit via latent variables

$$C1_u, C2_u, \dots, CJ_u;$$

$$\text{Level 2: } Cj_u \sim N(\mu_j, \sigma_j^2), j = 1, \dots, J-1.$$

Mean μ_j and variance σ_j^2 of the latent categorical profile variable Cj_u are estimated in Level

2. The mean proportion p_j of latent profile j is obtained via the value of Cj_u , where $p_j =$

$$\frac{e^{\mu_j}}{1 + \sum_{j=1}^{J-1} e^{\mu_j}}, j = 1, \dots, J-1 \text{ and } p_J = \frac{1}{1 + \sum_{j=1}^{J-1} e^{\mu_j}}. \text{ Consequently, this multilevel mixture model}$$

examines whether the probability that an employee belongs to a specific job characteristics profile varies significantly across work departments. In other words, whether in some work

departments there is a larger probability that an employee belongs to a certain profile than in other work departments. The model can be extended by regressing Cj_u on the Level 1 and/or Level 2 variables. The expected value of μ_{cj} of Cj_u is then dependent on the Level 1 and/or Level 2 variables and thus also explains part of the variance σ_{cj}^2 .

The model for Level 2 classes based on the relative frequency of the Level 1 profiles investigates whether Level 2 latent classes can be identified based on the proportions of the Level 1 latent profiles. Consequently, the equation for a two-level model in which latent profiles appear in Level 1 and the proportions of the Level 1 latent profiles are regressed on the Level 2 latent classes is

$$\text{Level 1: } y_{ijk_u} = \mu_{jku} + \varepsilon_{ijk_u} \quad \varepsilon_{ijk_u} \sim N(0, \sigma_{jk}^2)$$

Probability of latent class j $P(y_{ijk_u}) = p_{ju}$ is associated with unit via latent variables

$C1_u, C2_u, \dots, Cj_u$;

$$\text{Level 2: } Cj_u|b = \text{intercept}_j + \beta_{jb} \times CB_b, j = 1, \dots, J - 1, b = 1, \dots, B - 1.$$

The value of $Cj_u|b$ depends on which of the between latent profiles the CB_b unit u belongs to. The proportion of latent profiles in latent class CB_b is obtained via a regression equation

$$\text{where } p_{j|b} = \frac{e^{\text{intercept}_j + \beta_{jb} \times CB_b}}{1 + \sum_{j=1}^{J-1} e^{\text{intercept}_j + \beta_{jb} \times CB_b}}, j = 1, 2, \dots, J-1 \text{ and } p_{J|b} = \frac{1}{1 + \sum_{j=1}^{J-1} e^{\text{intercept}_j + \beta_{jb} \times CB_b}}.$$

Similarly, the proportion of CB_b in the whole data is $p_b = \frac{e^{CB_{jb}}}{1 + \sum_{k=1}^{K-1} e^{CB_{kb}}}$, $k = 1, \dots, K-1$ and

$$p_K = \frac{1}{1 + \sum_{k=1}^{K-1} e^{CB_{kb}}} \text{ when } b = B. \text{ To concretize, this multilevel mixture model extends the}$$

single-level LPA to the multilevel context (in our example, nesting employees in work

departments) by specifying latent profiles for both employees (Level 1) and work

departments (Level 2). The model can be extended by regressing $Cj_u|b$ on the Level 1 and/or

Level 2 variables and regressing CB_b on the Level 2 variables.

The equation for the Level 2 classes based on Level 1 ratings with two levels in which latent classes $j = 1, 2, \dots, J$ (across individuals i in $k = 1, \dots, K$ variables) appear at Level 2 is

$$\text{Level 1: } y_{ijk} = \mu_{jk} + \varepsilon_{ijk} \quad \varepsilon_{ijk} \sim N(0, \sigma_{jk}^2)$$

$$\text{Level 2: } \mu_{jk} = \alpha_{jk}$$

Probability p_j of latent class j is received via latent variable $CB_j = \mu_j$

$$p_j = \frac{e^{\mu_j}}{1 + \sum_{j=1}^{J-1} e^{\mu_j}}, j = 1, \dots, J-1 \text{ and } p_J = \frac{1}{1 + \sum_{j=1}^{J-1} e^{\mu_j}}. \text{ Consequently, this model}$$

identifies the higher-level classes based on the Level 1 ratings of certain criteria variables.

To sum up, both models, i.e., *Level 2 classes based on the relative frequency of the Level 1 profiles* and *Level 2 classes based on Level 1 ratings*, can be used where the interest is in classifying upper-level units (in this case, work departments). The main difference between these models is that *Level 2 classes based on the relative frequency of the Level 1 profiles* investigates latent classes based on within-person profiles, whereas *Level 2 classes based on Level 1 ratings* examines upper-level classes on the basis of variables observed on the single level. Consequently, model selection depends on whether the research target is to investigate latent subpopulations simultaneously at both levels of the data, or focus solely on upper-level latent classes. In turn, in the model *Level 2 variations in the size of the Level 1 profiles*, upper-level latent subpopulations are not investigated. Instead, the main research question is to investigate inter-contextual variability in the Level 1 profiles, i.e., whether the rate of individual (i.e., single-level) experiences vary by higher-level units.

In the fourth step, the estimation of latent profiles (Level 1)/classes (Level 2) for alternative models is conducted, starting from one profile/class solution and continuing to K/B latent profile/class solutions. Because mixed normal distributions of observations are assumed, and typically data appear to be missing, the full information maximum likelihood estimator with robust standard error estimation is needed (Asparouhov & Muthén, 2008).

While trying to find the maximum likelihood value, the algorithm may prematurely stop and return a sub-optimal set of parameter values; that is, lead to local maxima. To overcome the local maxima problem and find the global maximum, i.e., optimal solution, it is necessary to use several sets of starting values for parameters (Asparouhov & Muthén, 2008; McLachlan & Peel, 2000; see also Li, Harring, & Macready, 2014). It should also be noted that different MLPA models may require different starting value sets.

In the fifth step, the final model is selected. Selection of the optimal number of profiles should be based on the inspection of several statistical information criteria (e.g., AIC, consistent AIC, BIC, and the sample-size adjusted BIC) and tests (e.g., Lo-Mendell-Rubin LRT and the Bootstrap LRT), the statistical adequacy of the solution (e.g., convergence and absence of negative variance estimates) and inspection of the substantive meaning and theoretical conformity of the solution (McLachlan & Peel, 2000; see also Morin, Morizot, Boudrias, & Madore, 2011). Simulation studies indicate that the BIC, adjusted BIC, consistent AIC and BLRT test appear to be particularly effective in choosing the model which best covers the sample's true parameters on the single-level (Henson, Reise, & Kim, 2007; Nylund, Asparouhov, & Muthén, 2007; Tein, Coxe, & Cham, 2013; Tofighi & Enders, 2008; Tolvanen, 2007; Yu & Park, 2014). For deciding the number of upper-level latent classes, the BIC has proved to be the most efficient criterion (Finch & French, 2014; Zhang, Zhang, Zhang, & Jiou, 2014; Yu & Park, 2014). Graphical presentations of alternative solutions also aid the selection of the final model. Moreover, as recommended by Meyer and Morin (2016), the final model results, based on both standardized and raw scores, are also presented graphically.

In the sixth and final step, it is possible to extend MLPA to include single- and upper-level covariates and to investigate their cross-level interaction. According to the latest knowledge, the covariates should not be incorporated into the models until after the final

unconditional solution has been identified (Diallo & Lu, 2017; Diallo, Morin, & Lu, 2017; Kim, Vermunt, Bakk, Jaki, & Van Horn, 2016; Nylund-Gibson & Masyn, 2016).

Furthermore, to avoid the problem of profile switching the starts values from the final unconditional MLPA model should be used.

Having provided an overview of the steps of MLPA, we now demonstrate the method using a study example in which job characteristics profiles based on the JDCS model are analyzed with hierarchical data on university employees.

Data used in the demonstration

Participants and procedure

Participants were employees from two multidisciplinary equal-sized universities located in Central Finland. The data were collected in autumn 2008 with an electronic questionnaire which was sent to the e-mail addresses of 4,508 university employees working at least 20 hours per week. In total, 2,137 responses were received after two reminders, yielding a response rate of 48%. Of the participants, 67.1% were women, and the mean sample age was 42.9 ($SD = 11.1$). As expected, the sample was highly educated: 76% had a master's degree or higher. The majority (93.8%) worked full-time and, owing to the project nature of research work, 57% had a temporary employment contract. The sample was representative of the population of university workers in these two universities in occupational grouping (i.e., teachers, researchers, administrative and technical staff). However, women and temporary workers were overrepresented among the participants compared to the population from which the sample was drawn (for sample representativeness, see Kinnunen, Mäkikangas, Mauno, Siponen, & Nätti, 2011).

Altogether, 78 work departments, covering 15 faculties and two administrative units, were represented in our data. These 78 departments served as clustering units in the multilevel analyses. Mean departmental size was 24.8 ($SD = 20.1$, Range = 3–115)

employees. Information on their unit was missing for 179 participants, and thus the number of participants (Level 1) whose data were used in the analysis was 1,958.

Measures

To assess the job characteristics of JDCS model, we used the QPS Nordic Questionnaire, which has been developed and validated by Lindström et al. (2000). All the items were rated on a 5-point scale (1 = rarely or never, 5 = very often or always). Job demands were operationalized by *workload* and measured with three items (e.g., “Do you have too much to do?”; Cronbach’s alpha = .83). *Job control* was also measured by three items (e.g., “Can you influence decisions that are important for your work?”; Cronbach’s alpha = .73). *Social support from supervisors* (e.g., “If needed, is your nearest superior willing to listen to your work-related problems?”; Pearson correlation = .84) and *colleagues* (e.g., “If needed, can you get support and help with your work from your co-workers?”; Pearson correlation = .73) were each evaluated by two items.

Steps and results of the MPLA demonstration

Step 1: Definition of the problem and selection of the variables

The JDCS model (Karasek, 1979; Karasek & Theorell, 1990) suggests that eight types of job can be identified based on three job characteristics (job demands, job control and social support). The combination of job demands and control is assumed to form four job types, namely, active (high job demands, high job control), passive (low job demands, low job control), high-strain (high job demands and low job control) and low-strain (low job demands and high job control) (Karasek, 1979). The level of social support further specifies each of these four job types as either isolated or collective (Karasek & Theorell, 1990).

Most of the previous studies have focused on the combinations of job demands and job control (i.e., JDC model) by using mean or median splits (for a meta-analysis, see Kivimäki et al., 2012). Therefore, knowledge of the combinations of all three job

characteristics without using any cut-off values is lacking. Two studies investigating the JDC model have applied a person-centered approach (Ilgic et al., 2017; Mauno et al., 2016). However, both studies failed to find all the constellations posited in the JDC model. In one of these, a Finnish two-year follow-up study, low- (80%) and high-strain (10%) profiles were identified (together with two change profiles in which the level of job control either increased or decreased over time). In the other, a Swiss 10-year longitudinal study, 57% of the employees studied belonged to the low-strain profile and 17% to the active job profile (Ilgic et al., 2017). Approximately a third of the employees belonged to the profile in which changes were evident in workload and job control. Based on the findings of these two studies, it seems likely that in the present cross-sectional sample of university employees two or three profiles (i.e., low-strain, high-strain, active job) will be identified. However, in the absence of previous person-centered studies utilizing the JDCS model, it is unknown how the facets of social support are likely to combine together with workload and job control. Via LPA it is possible to identify these naturally occurring latent subpopulations.

Previous studies have only investigated the JDC model at the individual level, thus neglecting the possibility that these job characteristics may also be socially shared, signaling that employees who share the same working environment may also share a similar interpretation of these job characteristics (see Mauno et al., 2014; van Yperen & Snijders, 2000). Sharing (also termed emotional contagion; Barsade, 2002) occurs because employees who share the same work context tend to interact with each other as well as share the same organizational policies and practices, implying that their thinking and behavior in relation to work and the organization tend to converge over time (see Peiro, 2001).

Instead of simply estimating the amount of variance in job characteristics between work departments, as in traditional multilevel modeling, we first investigated whether individual JDCS *profiles* vary across work departments. This inter-group (e.g., team, work

department, organization) variability in the within-person job characteristics profiles reveals whether certain, e.g., high- or low-strain, constellations are more typical in certain work departments than others. Moreover, it also offers a possibility to identify higher level classes based on individual JDCS experiences (based either on profiles or observed ratings).

Therefore, our second MLPA aims to contribute to the stress literature by identifying and investigating the proportions of healthy vs. unhealthy work environments in light of the JDCS model. Investigating upper-level latent classes should also make it possible to answer the generalization problem encountered in person-centered research, i.e., the extent to which the profiles found are context-specific or more universal (see Meyer & Morin, 2016; Mäkikangas & Kinnunen, 2016).

It is well known that objective working hours are associated with increased levels of job stress (including stressors such as excessive job demands, insufficient job control, and inadequate social support; Lee, Suh, Kim, & Park, 2017). Also, long working hours at the work department level might serve as a risk factor associated with high-strain jobs and high-strain work departments. Therefore, working hours at the individual and at the department levels were investigated as covariates. Moreover, the average size of the work department was investigated as a covariate, since it is well established that interaction and emotional contagion are more prevalent in smaller than larger work units (Kelly & Barsade, 2001). Emotional contagion may have both positive and negative outcomes, in that average work department size may predict either high- or low-strain constellations.

Step 2: Preliminary analysis of the data

Using Mahalanobis distance and corresponding p -values, the preliminary analyses revealed the presence of deviant cases. At the $p < .05$ probability level, 5% of the cases were expected to be statistically significant at random. In our case, the percentage of statistically significant cases was 5.9%. However, in the cases with the smallest p -values, further

inspection revealed that the observed values of job characteristics were highly similar, suggesting the existence of a subpopulation. Therefore, all these deviant cases were included in the analysis (see also Footnote 1). The preliminary analyses showed further that the reliabilities of the observed variables were acceptably high (i.e., $> .73$; see Measures section). ICCs (together with descriptive information on the variables) are presented in Table 1. As shown, the ICCs were significant for each of the four job characteristics variables, thus demonstrating the existence of meaningful variation between the Level 2 units (i.e., work departments). Significant ICCs mean that as well as varying from one individual to another, perceived job characteristics also varied from one work department to another.

Step 3: Model specification

We expected to find job characteristics profiles at Level 1 that would further either vary randomly across work departments or form latent classes at the work department level based on either the Level 1 profiles or observed ratings. Consequently, the model specification would correspond to the above-described MLPA models. For this demonstration, we first estimated a single-level LPA, which was further extended to form alternative MLPA models. This enables us to demonstrate the difference between single and multilevel LPAs. We estimated the single-level LPA in two alternative ways: with unequal means and variances, and with unequal means only, across profiles. According to the common practice, covariances between the observed variables were not freely estimated, but instead were fixed to zero according to the assumptions of LPA. In the model of Level 2 variations in the size of the Level 1 profiles, in addition to mean profile size, the variances associated with the variation in profile size were freely estimated at Level 2. In the model of Level 2 classes based on the relative frequency of the Level 1 profiles, the latent class variable at Level 2 was freely estimated. Additionally, in this model, the Level 1 profile was regressed on the Level 2 classes, thus estimating coefficients for multinomial regression, that

is, statistically significant regression coefficients associated with the Level 1 latent profile proportions for the Level 2 latent classes. In the model of Level 2 classes based on Level 1 ratings, the variances in the Level 1 and Level 2 were freely estimated and only the Level 2 mean values were allowed to differ between the latent classes.

Step 4: Model estimation

All the analyses were performed using Mplus (version 7.3). The model parameters were estimated using full information maximum likelihood (FIML) estimation with standard errors that are robust to non-normality (MLR estimator in Mplus) (Muthén & Muthén, 1998–2014). The MLR estimator is recommended for use in multilevel mixture models, as it is also robust to the nesting of observations into higher units. This method allowed use of all the observations in the data to estimate the model parameters. Starting value sets of 1,000 were used. Based on 50 iterations, the 50 best-fitting models were further iterated to find the likelihood maxima. If the log-likelihood is not replicated in several starting values or if convergence problems appear, the starting value sets should be increased.

Step 5: Choosing and interpreting the best-fitting model

Single-level LPA. The LPAs were first estimated with unequal means and variances across the profiles as recommended in the literature (Diallo et al., 2016; Enders & Tofighi, 2008; Kim, Lamont et al., 2016; Peugh & Fan, 2013) (see Appendix A for Mplus script; Model 1). This analysis produced a two-profile solution. The solutions with three or more profiles did not converge. Therefore, we next estimated the more parsimonious model, i.e., a model where the latent profiles differed in the mean values of job characteristics (see Appendix A for Mplus script; Model 1). Figure 1 presents the elbow plot of the information criteria (BIC, aBIC, AIC and CAIC) for the different profile solutions of the single-level LPA based solely on the mean differences between the profiles. As illustrated, between the one- and two-profile solutions, all the information criteria decreased strongly, but remained at the

same level up to the six-profile solution. After a six-profile solution, the information criteria again decreased compared with the previous solutions. When the k and $k-1$ profile solutions were compared, the p -value of the BLRT test was $< .001$; consequently, in all the comparisons with the $k-1$ profile solutions, the null hypothesis was rejected.

Typically, the change (i.e., decrease) in ICs depicted via an elbow plot should be U-shaped, while the point after which the slope flattens out indicates the optimal number of profiles in the data (see e.g., Morin, Meyer, McInerney, Marsh, & Ganotice, 2015; Wang, Morin, Ryan, & Liu, 2016). The shape of the elbow plot obtained here may reflect the fact that the observed variables were not strictly continuous. Inspection of the profile solutions indicated that the six-profile solution and solutions with more than six profiles contained profiles with very small memberships (i.e., 1-2% of cases). We further investigated the two- to five-solution profiles graphically. It became evident that in the two-profile solution, the profiles differed from each other qualitatively (i.e., the two profiles had distinct shapes). The three- to five-profile solutions contained similarly shaped profiles with small level differences in workload, job control and facets of social support. Thus, these profile-solutions did not reveal new job types, as depicted in the JDCS model. Also, the classification quality of the two-profile solution was high: the average latent profile probabilities for the most likely latent profile membership (AvePP) were .90 and .94. In the three- to five-profile solutions, the AvePPs were much lower, the lowest varying between .74-.83.

The two-profile solution is shown in Figures 2 (standardized scores) and 3 (raw scores). The first profile contained 64% of the participants, and was characterized by high social support, high levels of job control and low levels of workload. Consequently, this profile was labeled “Collective low-strain job”. The second profile comprised 36% of the participants, and was characterized by high levels of workload, low levels of job control and low levels of social support from both supervisor and colleagues. Thus, using the terminology

of the JDCS model, this profile was labeled “Isolated high-strain job” (Karasek & Theorell, 1990). Previous person-centered research has also identified these low- and high-strain profiles (Mauno et al., 2016), and facets of social support further specified them as collective and isolated respectively.

Simulation studies suggest that LPAs with unequal variances lead to a less biased and a more accurate representation than models with only freely estimated means (Diallo et al., 2016; Enders & Tofighi, 2008; Kim, Lamont et al., 2016; Peugh & Fan, 2013). As the estimation with equal variances resulted in a two-profile solution, as shown above, we also interpreted a two-profile solution based on unequal variances across the profiles. According to the BIC, the model with unequal variances (19151.664) was statistically better than the model with equal variances (21210.085). The two-profile solution based on unequal variances is shown in Figure 4. The form of the profiles is the same as in the two-profile solution with equal variances across profiles, but the size of the profiles is noticeably different (64% vs. 41% and 36% vs. 59%). Below, the MLPAs are illustrated based on the two-profile solution with unequal variances across the profiles².

MLPA. Of the alternative MLPA models, Level 2 variations in the size of the Level 1 profiles was first tested (see Appendix A for Mplus script; Model 2). This was done by setting the starting values according to the single-level LPA and the random starting values as zero, thereby ensuring that the Level 1 profiles would remain unchanged (see Mplus script for Model 2). The results of the random effect MLPA showed that the size of the latent profile varied significantly between work departments (variance estimate = 0.322, s.e. = 0.162, $p = .047$), suggesting that the typicality of the “Collective low-strain job” and “Isolated high-strain job” profiles varied from one department to another. Overall, these results indicate that, depending on the university department, the proportional distributions of the two individual job characteristics profiles are different.

The model of Level 2 classes based on the relative frequency of the Level 1 profiles was then estimated (see Appendix A for Mplus script; Model 3). We compared the different Level 2 latent class solutions using the BIC criterion (Finch & French, 2014; Zhang et al., 2014; Yu & Park, 2014). The corresponding values for the one- to three-class solutions were 19139.26, 19137.84 and 19473.53. Consequently, the two-class solution was deemed sufficient. These classes are presented graphically in Figure 5. The results showed that the first Level 2 latent class included departments in which the great majority of employees were characterized by the profile “Isolated high-strain job” (8% of all work departments); these departments were thus labeled “Isolated high-strain departments”. The second Level 2 latent class contained work departments which had relatively equal proportions of “Isolated high-strain job” and “Collective low-strain job” profiles (92% of all work departments); the Latent class 2 was thus labeled as “Mixed strain departments”. In this class solution (two Level 1 profiles, two Level 2 latent classes), the multinomial regression coefficient obtained from regressing the Level 1 latent profiles on the Level 2 latent classes was 2.314 (s.e. = 0.815), $p = .002$. This regression coefficient indicates that in Level 2 latent class 1, the probability for the “Isolated high strain job” profile is .936 and for the “Collective low strain job” profile .064. In Level 2 latent class 2, the corresponding probabilities are .590 and .410.

Finally, a model of Level 2 classes based on Level 1 ratings was estimated (see Appendix A for Mplus script; Model 4). In the one-class solution, the BIC was value 22028.68 and in the two-class solution 22037.56. Thus, it seems that no Level 2 latent classes were present in the data when the MLPA was based on Level 1 observations.

Step 6: Investigating the role of potential covariates

In this step, the effects of the covariates and their cross-level interactions were tested in the two of the MLPA models, namely Level 2 variations in the size of the Level 1 profiles and Level 2 classes based on the relative frequency of the Level 1 profiles. As no latent

classes were evident in the model of Level 2 classes based on Level 1 ratings, the role of covariates in this model were not investigated. In the tested models, working hours (Level 1), the size of the department, and average departmental working hours (Level 2) were used as covariates (see Appendix A for Mplus script; Model 5 and Model 6). To avoid the problem of profile switching, the starting values of the model without covariates were used to ensure that the nature of the Level 1 profiles remained unchanged. The random starts function is only needed for the identification of the profiles themselves, and no longer needed thereafter.

First, covariates were added to the model of Level 2 variations in the size of the Level 1 profiles. The results showed that working hours at Level 1 (standardized estimate = 0.042, s.e. = 0.007, $p < .001$) predicted the likelihood of membership of the profile “Isolated high-strain job”. Consequently, the longer the hours worked weekly by the individual, the more likely he/she was to be characterized by the “Isolated high-strain job” profile. At Level 2, neither the size of the department (standardized estimate = -0.002, s.e. = 0.003, $p = .403$) nor average working hours (standardized estimate = -0.049, s.e. = 0.047, $p = .303$) predicted the random mean variation. No cross-level interactions were evident for working hours.

In the model of Level 2 classes based on the relative frequency of the Level 1 profiles, the results showed that working hours at Level 1 (standardized estimate = 0.040, s.e. = 0.007, $p < .001$) predicted the likelihood of the Level 1 profile “Isolated high-strain job”. Neither of the Level 2 covariates predicted membership of the Level 2 latent classes (standardized estimate = 0.045, s.e. = 0.037, $p = .225$ for department size; standardized estimate = 0.075, s.e. = 0.173, $p = .666$ for average working hours). No evidence was found for cross-level interactions.

Discussion

This study presented and illustrated the use of MLPA, a method which offers an interesting extension of single-level LPA as it allows estimation of the distribution of Level 1

profiles across Level 2 units and the estimation of Level 2 classes based on either single-level profiles or variables. These methods have been successfully applied in various disciplines. For instance, they have been used to analyze categorical variables in the school context in the educational sciences (e.g., Allison, Adlaf, Irving, Schoueri-Mychasiw, & Rehm, 2016), and regional health behaviors in the health sciences (e.g., Erickson, Lenk, Toomey, Nelson, & Jones-Webb, 2016), and even for clustering scientific disciplines according to research output profiles (Mutz, Bornmann, & Daniel, 2013). However, in the organizational sciences the use of multilevel mixture modeling continues to be uncommon, despite datasets that are often hierarchical, e.g., employees are nested in work units and teams are nested in organizations. Moreover, in the organizational sciences the variables analyzed are typically continuous rather than categorical, and thus MLPA is a highly advantageous modeling technique in this field of inquiry.

The consequences of not taking the nested structure of the data into account, e.g., inaccuracy in parameter estimation and biased standard errors due to using the Level 1 model alone, have been widely discussed and demonstrated (Asparouhov & Muthén, 2008; Chen, Kwok et al., 2010; Luo & Kwok, 2009; Meyers & Beretvas, 2006; Moerbeek, 2004; Park & Yu, 2016; Vermunt, 2003, 2008). Thus, it is essential to recognize whether the data have been collected in a nested context, and to use multilevel modeling even in the case of low ICCs (Bliese et al., 2017). Conversely, aside from avoiding inaccuracy in profile estimation, the existence of a higher-level nested structure offers interesting novel analytical possibilities, as this study has demonstrated.

First, we showed in our demonstration that the distribution of profiles of individual latent job characteristics varied between university departments. That is, in some work departments the “Collective low-strain job” profile was more typical, whereas in others the “Isolated-high strain job” profile was more typical. Second, we were able to demonstrate in

some detail how work departments were clustered according to the proportions of single-level profiles. In our data demonstration, two Level 2 latent classes were identified: “Collective low-strain” work departments, comprising most (92%) of the departments studied and “Isolated high-strain” departments, accounting for the remaining 8%. To conclude, although job characteristics describe individual experiences of job control, demands and support, a high-strain job profile was more typical in some work departments than others, and the identification of these departments was achieved through the application of MLPA.

Overall, our results indicate that the sample of a university employees studied typically belonged to a high-strain job profile. A novel finding is that the facets of social support did not reveal the additional job types described in the JDCS model (Karasek & Theorell, 1990). Thus, it seems that the original JDC model could be sufficient to characterize and differentiate types of jobs (Karasek, 1979), at least among highly educated university employees. Moreover, and perhaps more interestingly, through the application of a multilevel mixture approach, we were able to identify latent classes of ‘healthy’ and ‘unhealthy’ work departments. The unhealthy “Isolated high-strain” departments contained nearly all the employees with a high-strain job. This suggests that psychosocial job characteristics are not only individually experienced, but may also be linked to collective stress and thus shared with colleagues.

Implications and recommendations

MLPA importantly extends latent profile identification beyond the level of the individual employee, thereby enlarging the applicability of a person-centered methodology. Moreover, MLPA also enriches the outcomes of multilevel modeling. For example, in our case example, rather than simply estimating the amount of variance in job characteristics between work departments, it was possible to identify qualitatively different work departments. There were no valid hypotheses for Level 2 latent classes and thus they were

identified purely on a data-driven way. In the context of the multilevel person-centered research it would be advisable to use and combine both individual and organizational theories. Use of multiple theoretical frameworks in person-centered research also enables the possibility to conduct studies in a theory-driven way, as Wang (2007) has demonstrated. Multilevel mixture modeling also has the potential to answer the generalization problem of person-centered research (see Meyer & Morin, 2016; Mäkikangas & Kinnunen, 2016), i.e., whether the profiles found are unit-specific or more universal. Via the identification of Level 2 latent classes, it is possible to discover which profile type is predominant (versus underrepresented) across the higher-level units.

Besides offering theoretically innovative and interesting novel research targets, multilevel mixture modeling also provides plenty of modeling extension possibilities. In our MLPA demonstration, it was assumed that the Level 1 profiles are the same across the work departments, and vary only in their relative proportions. However, it is recommended that the Level 1 profile solution be verified when estimating the alternative MLPA models. In the longitudinal context, multilevel mixture models can be used to investigate Level 1 and Level 2 growth trajectories simultaneously (Chen, Kwok et al., 2010). It is also worth noticing that all the multilevel mixture models can be based on latent factors instead of the sum scores used in this demonstration. However, too much complexity in the model easily induces convergence problems, and hence it might be reasonable to start the modeling procedure with a simpler model and increase the level of complexity within the boundaries of the tested model and sample size.

Overall, multilevel mixture models are highly recommended for analyzing nested data sets where the research interest is in studying latent classes in within-person profiles or their inter-contextual variability (see also, Henry & Muthén, 2010; Vermunt, 2003, 2008, 2011). One challenge in LPA - as in other types of mixture modeling - is to find the best fit

solution, i.e., the one showing the highest likelihood. It is typical of this kind of analysis that the iteration process stops at the local maximum likelihood value. Therefore, given the aim of finding the correct number of latent profiles, the use of several random sets of starting values is recommended (see also Meyer & Morin, 2016).

Although MLPA models offer indisputable benefits in modeling latent subpopulations in nested datasets, they also have some limitations. First, as models of these kinds are relatively complex and computationally heavy, they require a large sample size at both levels: a large number of observations at Level 1 and a large number of Level 2 clusters. Large sample sizes are especially required when the latent subpopulations are multifaceted and less distinct (Park & Yu, 2017). For example, the present data included 78 departments, of which only 8% were in the “Isolated high-strain” class. An insufficient sample size in mixture modeling can lead to convergence issues, improper solutions, and the inability to identify small but meaningful subpopulations (Berlin, Williams, & Parra, 2014). However, a sufficient sample size is difficult to determine, as it depends, for example, on the size of the model, the distribution of the variables, the amount of missing data, the reliability of the variables, and the strength of the relations among the variables (Muthén & Muthén, 2002). Consequently, further statistical research is needed on the calculation of statistical power and adequate sample size for different MLPA models.

Second, a simplification of the model may be needed, e.g., in the case of a large number of Level 1 latent profiles whose random means correlate highly at Level 2. In this case, using a common factor to model the different random means at Level 2 may be a solution, as recommended by Henry and Muthén (2010). Third, modeling the Level 2 latent subpopulations and investigating the effects of the Level 2 covariates requires careful planning and theoretical reasoning prior to the data collection. For example, in the present

demonstration, the Level 2 covariates were nonsignificant and their number was rather limited.

In sum, multilevel mixture modeling offers a fruitful approach to the analysis of hierarchical data in person-centered research, which is nowadays more commonly undertaken in the organizational sciences. Given the emphasis in organizational research to understand human behavior in organizational settings, this method allows to assess latent typologies in different organizational hierarchies (e.g., individual, team and organizational) together disentangling the varied effects of covariates expressed at different levels. Via studying the effects of various contextual factors on lower and upper-level typologies, this method have potential to consider a broader context needed in organizational research. Especially being suitable for studying contextual-level predictors on individual typologies of behavior, we believe that these advanced statistical methods will encourage the formulation of new research questions, lead to new theorizing, and eventually also to new answers to issues pertaining to the employee-organization relationship.

References

- Allison, K. R., Adlaf, E. M., Irving, H. M., Schoueri-Mychasiw, N., & Rehm, J. (2016). The search for healthy schools: A multilevel latent class analysis of schools and their students. *Preventive Medicine Reports, 4*, 331–337. doi: 10.1016/j.pmedr.2016.06.016
- Asparouhov, T., & Muthén, B. (2008). Multilevel mixture models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 27–51). Charlotte, NC: Information Age.
- Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47*, 644–675. doi: 10.2307/3094912

- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*, 338–363. doi: 10.1037/1082-989X.8.3.338
- Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology, 39*, 174–187. doi: 10.1093/jpepsy/jst084
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Kline & S. W. J. Kozlowski (Eds), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco: Jossey-Bass.
- Bliese, P. D., Maltarich, M. A., & Hendricks, J. L. (2017). Back to basics with mixed-effects models: Nine take-away points. *Journal of Business and Psychology*. Advance online publication. doi: 10.1007/s10869-017-9491-z
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural models: Causes, consequences, strategies. *Sociological Methods and Research, 29*, 468–508. doi: 10.1177/0049124101029004003
- Chen, Q., Kwok, O-M., Luo, W., & Willson, V. L. (2010). The impact of ignoring a level of nesting structure in multilevel growth mixture models: A Monte Carlo study. *Structural Equation Modeling, 17*, 570–589. doi: 10.1080/10705511.2010.510046
- Diallo, T. M. O., & Lu, H. (2017). On the application of the three-step approach to growth mixture models. *Structural Equation Modeling, 24*, 714–732. doi: 10.1080/10705511.2017.1322516
- Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2016). Impact of misspecifications of the latent variance-covariance and residual matrices on the class enumeration accuracy of growth

- mixture models. *Structural Equation Modeling*, *23*, 507–531. doi: 10.1080/10705511.2016.1169188
- Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2017). The impact of total and partial inclusion or exclusion of active and inactive time invariant covariates in growth mixture models. *Psychological Methods*, *22*, 166–190. doi: 10.1037/met0000084
- Enders, C. K., & Tofighi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling*, *15*, 75–95. doi: 10.1080/10705510701758281
- Erickson, D. J., Lenk, K. M., Toomey, T. L., Nelson, T. F., & Jones-Webb, R. (2016). The alcohol policy environment, enforcement and consumption in the United States. *Drug and Alcohol Review*, *35*, 6–12. doi: 10.1111/dar.12339
- Finch, W. H., & Bronk, K. C. (2011). Conducting confirmatory latent class analysis using Mplus. *Structural Equation Modeling*, *18*, 132–151. doi: 10.1080/10705511.2011.532732
- Finch, W. H., & French, B. F. (2014). Multilevel latent class analysis: Parametric and nonparametric models. *The Journal of Experimental Education*, *82*, 307–333. doi: 10.1080/00220973.2013.813361
- Henry, K. L., & Muthén, B. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling*, *17*, 193–215. doi: 10.1080/10705511003659342
- Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling*, *14*, 202–226. doi: 10.1080/10705510709336744
- Häusser, J. A., Mojzisch, A., Niesel, M., & Schulz-Hardt, S. (2010). Ten years: A review of

- recent research on the Job Demand-Control (-Support) model and psychological well-being. *Work and Stress*, *24*, 1–35. doi: 10.1080/02678371003683747
- Igic, I., Keller, A. C., Elfering, A., Tschan, F., Kälin, W., & Semmer, N. K. (2017). Ten-year trajectories of stressors and resources at work. Cumulative and chronic effects on health and well-being. *Journal of Applied Psychology*, *102*, 1317–1343. doi: 10.1037/apl0000225
- Karasek, R. (1979). Job demands, job decision latitude and mental strain: implications for job redesign. *Administrative Science Quarterly*, *24*, 285–308. doi: 10.2307/2392498
- Karasek, R., & Theorell, T. (1990). *Healthy work: Stress, productivity and the reconstruction of working life*. New York: Basic Books.
- Kelly, J. R., & Barsade, S. G. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes*, *86*, 99–130. doi: 10.1006/obhd.2001.2974
- Kim, M., Lamont, A. E., Jaki, T., Feaster, D., Howe, G., van Horn, M. L. (2016). Impact of an equality constraint on the class-specific residual variances in regression mixtures: A Monte Carlo simulation study. *Behavior Research Methods*, *48*, 813–826. doi: 10.3758/s13428-015-0618-8
- Kim, M., Vermunt, J., Bakk, Z., Jaki, T., & Van Horn, M. L. (2016). Modeling predictors of latent classes in regression mixture models. *Structural Equation Modeling*, *23*, 601–614. doi: 10.1080/10705511.2016.1158655
- Kinnunen, U., Mäkikangas, A., Mauno, S., Siponen, K., & Nätti, J. (2011). Perceived employability: Investigating outcomes among involuntary and voluntary temporary employees compared to permanent employees. *Career Development International*, *16*, 140–160. doi: 10.1108/13620431111115604

Kivimäki, M., Nyberg, S. T., Batty, G. D., Fransson, E. I., Heikkilä, K., Alfredson, L. ...

Theorell, T. (2012). Job strain as a risk factor for coronary heart disease: A collaborative meta-analysis of individual participant data. *Lancet*, *380*, 1491–1497. doi: 10.1016/S0140-6736(12)60994-5

Lee, K., Suh, C., Kim, J-E., & Park, J. O. (2017). The impact of long working hours on psychosocial stress response among white-collar workers. *Industrial Health*, *55*, 46–53. doi: 10.2486/indhealth.2015-0173

Li, M., Harring, J. R., & Macready, G. B. (2014). Investigating the feasibility of using Mplus in the estimation of growth mixture models. *Journal of Modern Applied Statistical Methods*, *13*, 484–513. doi: 10.22237/jmasm/1398918600

Lindström, K., Elo, A-L., Skogstad, A., Dallner, M., Gamberale, F., Hottinen, V., Knardahl, S., & Orhede, E. (2000). *User's guide for the QPS Nordic. General Nordic Questionnaire for psychosocial and social factors at work*. Copenhagen, Denmark: Nordic Council of Ministers.

Luo, W., & Kwok, O-M. (2009). The impacts of ignoring a crossed factor in analyzing cross-classified data. *Multivariate Behavioral Research*, *44*, 182–212. doi: 10.1080/00273170902794214

Mauno, S., De Cuyper, N., Tolvanen, A., Kinnunen, U., & Mäkikangas, A. (2014). Occupational well-being as a mediator between job insecurity and turnover intention: Findings at the individual and work department levels. *European Journal of Work and Organizational Psychology*, *23*, 381–393. doi: 10.1080/1359432X.2012.752896

Mauno, S., Mäkikangas, A., & Kinnunen, U. (2016). A longitudinal person-centred approach to the job demands-control model. *European Journal of Work and Organizational Psychology*, *25*, 914–927. doi: 10.1080/1359432X.2016.1187135

- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons, Inc.
- Meyer, J. P., & Morin, A. J. S. (2016). A person-centered approach to commitment research: Theory, research, and methodology. *Journal of Organizational Behavior, 37*, 584–612. doi: 10.1002/job.2085
- Meyers, J. L., & Beretvas, S. N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research, 41*, 473–497. doi: 10.1207/s15327906mbr4104_3
- Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research, 39*, 129–149. doi: 10.1207/s15327906mbr3901_5
- Morin, A. J. S., Meyer, J. P., McInerney, D. M., Marsh, H. W., & Ganotice Jr, F. A. (2015). Profiles of dual commitment to the occupation and organization: Relations to well-being and turnover intentions. *Asia Pacific Journal of Management, 32*, 717–744. doi: 10.1007/s10490-015-9411-6
- Morin, A. J. S., Morizot, J., Boudrias, J.-S., & Madore, I. (2011). A multifoci person-centered perspective on workplace affective commitment: A latent profile/factor mixture analysis. *Organizational Research Methods, 14*, 58–90. doi: 10.1177/1094428109356476
- Morin, A. J. S., & Wang, J. C. K. (2016). A gentle introduction to mixture modeling using physical fitness data. In N. Ntoumanis, & N. Myers (Eds.), *An introduction to intermediate and advanced statistical analyses for sport and exercise scientists* (pp. 183–210). London, UK: Wiley.
- Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*, 882–891. doi: 10.1111/j.1530-0277.2000.tb02070.x

- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599–620. doi: 10.1207/S15328007SEM0904_8
- Muthén, L. K., & Muthén, B. O. (1998–2014). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.
- Mutz, R., Bornmann, L., & Daniel, H.-D. (2013). Types of research output profiles: A multilevel latent class analysis of the Austrian Science Fund's final project report data. *Research Evaluation, 22*, 118–133. doi: 10.1093/reseval/rvs038
- Mäkikangas, A., & Kinnunen, U. (2016). The person-oriented approach to burnout: A systematic review. *Burnout Research, 3*, 11–23. doi:10.1016/j.burn.2015.12.002.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*, 535–569. doi: 10.1080/10705510701575396
- Nylund-Gibson, K., & Masyn, K. E. (2016). Covariates and mixture modeling: Results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling, 23*, 782–797. doi: 10.1080/10705511.2016.1221313
- Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson, & M. Kaptein (Eds.), *Modern statistical methods for HCI: A modern look at data analysis for HCI research* (pp. 275–287). Switzerland: Springer.
- Park, J., & Yu, H-T. (2016). The impact of ignoring the level of nesting structure in nonparametric multilevel latent class models. *Educational and Psychological Measurement, 76*, 824–847. doi: 10.1080/10705511.2010.510046
- Park, J., & Yu, H-T. (2017). Recommendations on the sample sizes for multilevel latent class models. *Educational and Psychological Measurement*. Advanced online publication. doi: 10.1177/0013164417719111

- Peiro, J. M. (2001). Stressed teams in organizations: A multilevel approach to the study of stress in work units. In J. Pryce, C. Weilkert, & E. Torkelson (Eds.), *Organizational health psychology: Europe 2001* (pp. 9–13). Nottingham, UK: European Academy of Occupational Health.
- Peugh, J. & Fan, X. (2013). Modeling unobserved heterogeneity using latent profile analysis: A Monte Carlo simulation. *Structural Equation Modeling, 20*, 616–639. doi: 10.1080/10705511.2013.824780
- Sterba, S. K. (2013). Understanding linkages among mixture models. *Multivariate Behavioral Research, 48*, 775–815. doi:10.1080/00273171.2013.827564
- 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*, 640–657. doi: 10.1080/10705511.2013.824781
- Tofghi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 317–341). Charlotte, NC: Information Age Publishing.
- Tolvanen, A. (2007). *Latent growth mixture modeling: A simulation study*. University of Jyväskylä, Department of Mathematics and Statistics, Report 111. Jyväskylä: University Printing House.
- van der Doef, M., & Maes, S. (1999). The job Demand-Control(-Support) model and psychological well-being: A review of 20 years of empirical research. *Work & Stress, 13*, 87–114. doi: 10.1080/026783799296084
- van Yperen, N. W., & Snijders, T. A. (2000). A multilevel analysis of the Demands-Control Model: Is stress at work determined by factors at the group level or the individual level? *Journal of Occupational Health Psychology, 5*, 182–190. doi: 10.1037//1076-8998.5.1.182

- Vermunt, J. K. (2003). Multilevel latent class models. *Sociological Methodology, 33*, 219–239. doi: 10.1111/j.0081-1750.2003.t01-1-00131.x
- Vermunt, J. K. (2008). Latent class and finite mixture models for multilevel data sets. *Statistical Methods in Medical Research, 17*, 33–51. doi: 10.1177/0962280207081238
- Vermunt, J. K. (2011). Mixture models for multilevel data sets. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 59–84). New York: Routledge.
- Wang, M. (2007). Profiling retirees in the retirement transition and adjustment process: Examining the longitudinal change patterns of retirees' psychological well-being. *Journal of Applied Psychology, 92*, 455–474. doi: 10.1037/0021-9010.92.2.455
- Wang, J. C. K., Morin, A. J. S., Ryan, R. M., & Liu, W. C. (2016). Students' motivational profiles in the physical education context. *Journal of Sport & Exercise Psychology, 38*, 612–630. doi: 10.1123/jsep.2016-0153
- Zhang, J., Zhang, M., Zhang, W., & Jiou, C. (2014). Model selection for complex multilevel latent class model. *Communications in Statistics – Simulation and Computation, 43*, 838–850. doi: 10.1080/03610918.2012.718836
- Yu, H-T., & Park, J. (2014) Simultaneous decision on the number of latent clusters and classes for multilevel latent class models. *Multivariate Behavioral Research, 49*, 232–244. doi: 10.1080/00273171.2014.900431

Figure Captions

Figure 1. Elbow plot of information criteria for different Level 1 profile solutions

Figure 2. Level 1 two-profile solution with equal variances across profiles based on standardized scores

Figure 3. Level 1 two-profile solution with equal variances across profiles based on raw scores

Figure 4. Level 1 two-profile solution with unequal variances across profiles based on raw scores

Figure 5. Level 2 classes based on the relative frequency of the Level 1 profiles

Footnotes

¹ By using Mahalanobis distance, it is possible to detect single outliers. However, deviant observations may also indicate the existence of latent subpopulation(s). Therefore, this method should be applied deliberately in person-centered studies.

² MLPAs based on the Level 1 profile solution with equal variances across profiles failed to identify Level 2 latent classes. Also, the relative size of the Level 1 profiles did not vary across work departments.

Appendix A

Mplus scripts for the tested models

Model 1. Single-level latent profile analysis: Fixed effects model

TITLE: Latent Profile Analysis – Single-Level Analysis

DATA: FILE IS ORM.dat;

VARIABLE: NAMES ARE

ID unit size wkhour wload jobc supps suppc;

USEVARIABLES ARE wload jobc supps suppc;

MISSING ARE ALL (-999);

CLASSES = c(2);

ANALYSIS: TYPE=MIXTURE;

ESTIMATOR=MLR;

STARTS=1000 50;

STITERATIONS=50;

LRTSTARTS=1000 50 1000 50;

PROCESSORS=4;

MODEL:

%OVERALL%

%C#1%

[wload jobc supps suppc];

wload jobc supps suppc; (This row is deleted estimating LPA with equal variances)

%C#2%

[wload jobc supps suppc];

wload jobc supps suppc; (This row is deleted estimating LPA with equal variances)

OUTPUT TECH14;

Model 2. Multilevel latent profile analysis: Level 2 variations in the size of the Level 1 profiles

TITLE: Multilevel LPA

DATA: FILE IS ORM.dat;

VARIABLE: NAMES ARE

ID unit size wkhour wload jobc supps suppc;

USEVARIABLES ARE wload jobc supps suppc;

MISSING ARE ALL (-999);

CLASSES = c(2);

CLUSTER IS unit;

ANALYSIS: TYPE=MIXTURE TWOLEVEL;

STARTS=0;

MODEL:

%BETWEEN%

%OVERALL%

c#1;

%WITHIN%

%C#1%

[wload*3.202 jobc*2.355 supps*3.247 suppc*3.665];

Wload*0.495 jobc*0.547 supps*0.567 suppc*0.586;

%C#2%

[wload*3.666 jobc*4.227 supps*4.287 suppc*3.305];

Wload*0.495 jobc*0.547 supps*0.567 suppc*0.586;

OUTPUT;

Model 3. Multilevel latent profile analysis: Level 2 classes based on the relative frequency of the Level 1 profiles

TITLE: Multilevel LPA
 DATA: FILE IS ORM.dat;
 VARIABLE: NAMES ARE
 ID unit size wkhour wload jobc supps suppc;
 USEVARIABLES ARE wload jobc supps suppc;
 MISSING ARE ALL (-999);
 CLASSES = BC(2) c(2);
 CLUSTER IS unit;
 WITHIN ARE wload jobc supps suppc;
 BETWEEN ARE BC;
 ANALYSIS: TYPE=MIXTURE TWOLEVEL;

MODEL:

% WITHIN%
 % OVERALL%
 % BETWEEN%
 % OVERALL%
 C ON BC;
 MODEL C:
 % WITHIN%
 % C#1%
 [wload jobc supps suppc];
 % C#2%
 [wload jobc supps suppc];
 OUTPUT;

Model 4. Multilevel latent profile analysis: Level 2 classes based on Level 1 ratings

TITLE: Multilevel LPA
 DATA: FILE IS ORM.dat;
 VARIABLE: NAMES ARE
 ID unit size wkhour wload jobc supps suppc;
 USEVARIABLES ARE wload jobc supps suppc;
 MISSING ARE ALL (-999);
 CLASSES = BC(2);
 CLUSTER IS unit;
 BETWEEN ARE BC;
 ANALYSIS: TYPE=MIXTURE TWOLEVEL;

MODEL:

% WITHIN%
 % OVERALL%
 wload jobc supps suppc;
 % BETWEEN%
 % OVERALL%
 wload jobc supps suppc;
 % BC#1%
 [wload jobc supps suppc];
 % BC#2%
 [wload jobc supps suppc];
 OUTPUT;

Model 5. Level 2 variations in the size of the Level 1 profiles with Level 1 and Level 2 covariates

```
TITLE: Multilevel LPA with covariates
DATA: FILE IS ORM.dat;
VARIABLE: NAMES ARE
ID unit size wkhour wload jobc supps suppc wkhourME;
USEVARIABLES ARE wload jobc supps suppc wkhour size;
MISSING ARE ALL (-999);
CLASSES = c(2);
CLUSTER IS unit;
WITHIN ARE wload jobc supps suppc wkhour;
BETWEEN ARE size;
DEFINE: CENTER wkhour (GROUPMEAN);
ANALYSIS: TYPE=MIXTURE TWOLEVEL;
STARTS=0;
```

```
MODEL:
% WITHIN%
% OVERALL%
C ON wkhour;
% C#1%
[wload*3.202 jobc*2.355 supps*3.247 suppc*3.665];
Wload*0.495 jobc*0.547 supps*0.567 suppc*0.586;
% C#2%
[wload*3.666 jobc*4.227 supps*4.287 suppc*3.305];
Wload*0.495 jobc*0.547 supps*0.567 suppc*0.586;
% BETWEEN%
% OVERALL%
C ON size wkhourME;
c#1;
OUTPUT;
```

Model 6. Level 2 classes based on the relative frequency of the Level 1 profiles with Level 1 and Level 2 covariates

```
TITLE: Multilevel LPA with covariates
DATA: FILE IS ORM.dat;
VARIABLE: NAMES ARE
ID unit size wkhour wload jobc supps suppc;
USEVARIABLES ARE wload jobc supps suppc;
MISSING ARE ALL (-999);
CLASSES = BC(2) c(2);
CLUSTER IS unit;
WITHIN ARE wload jobc supps suppc;
BETWEEN ARE BC;
ANALYSIS: TYPE=MIXTURE TWOLEVEL;
STARTS=0;
```

```
MODEL:
% WITHIN%
% OVERALL%
% BETWEEN%
% OVERALL%
C ON BC;
BC ON wkhourme size; (note! wkhourme is an aggregated variable)
MODEL C:
% WITHIN%
% C#1%
```

```
[wload*3.288 jobc*2.937 supps*3.501 suppc*3.591];  
wload*0.562 jobc*1.060 supps*0.749 suppc*0.589;  
%C#2%  
[wload*3.836 jobc*4.536 supps*4.579 suppc*3.184];  
wload*0.331 jobc*0.253 supps*0.220 suppc*0.557;  
OUTPUT;
```

Table 1

Means (M), Standard Deviations (SD), Intraclass Correlations (ICC) and Correlations between Study Variables at Level 1 (below the diagonal, N = 1,958 employees) and at Level 2 (above the diagonal, N = 78 departments)

<i>Variables</i>	M	SD	ICC	1	2	3	4	5	6	7	8
1. Workload	3.44	0.18	.05**	-	-.59	-.48	-.07				
2. Job control	3.52	0.27	.14**		-	.11	-.05				
3. Supervisory support	3.54	0.19	.03*			-	.73				
4. Colleague support	3.90	0.19	.04**				-				
5. Workload ^a	-	0.76						-			
6. Job control ^a	-	0.68						-.33	-		
7. Supervisory support ^a	-	1.15						-.18	.30	-	
8. Colleague support ^a	-	0.89						-.17	.27	.53	-

Note. ^a = Level 1 variable.

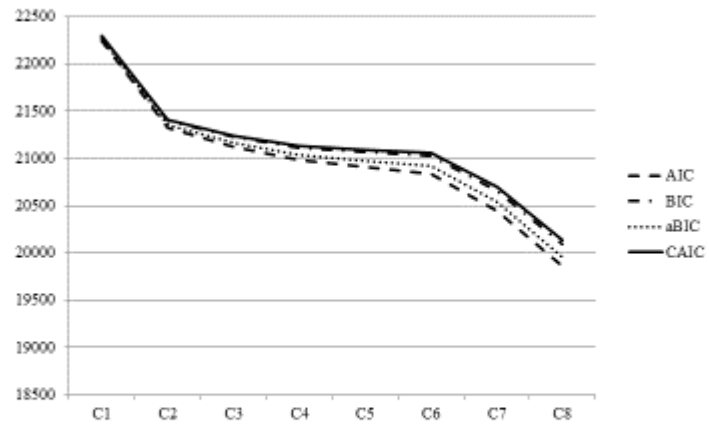


Figure 1.
Elbow plot of information criteria for different Level 1 profile solutions

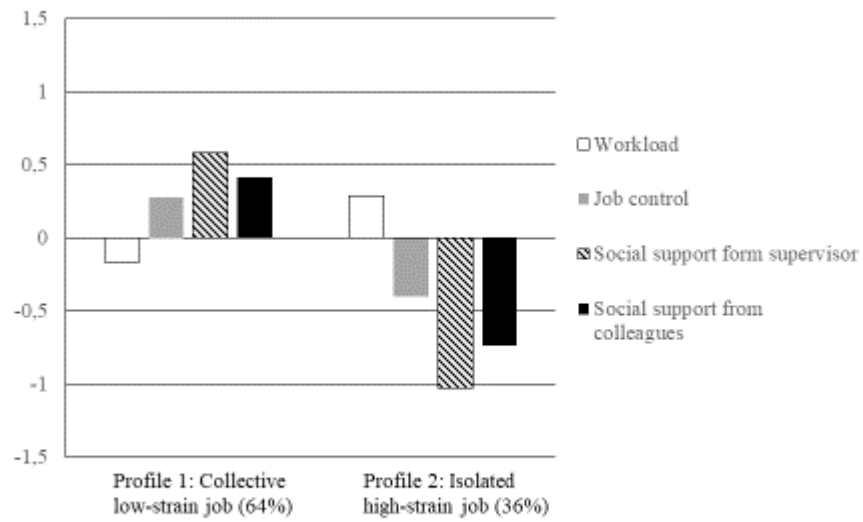


Figure 2.
Level 1 two-profile solution with equal variances across profiles based on standardized scores

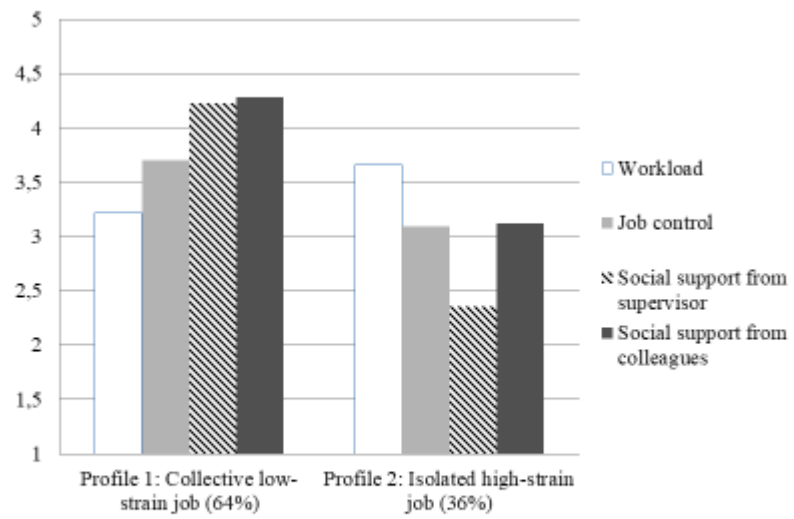


Figure 3.
Level 1 two-profile solution with equal variances across profiles based on raw scores

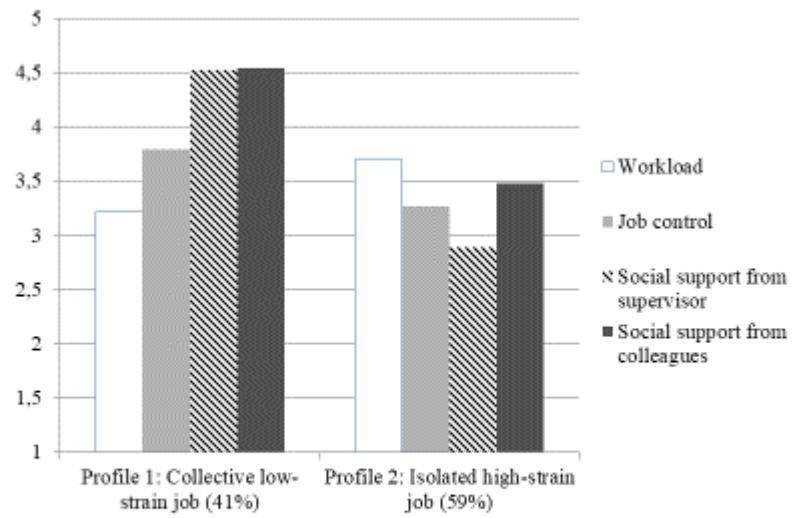


Figure 4.
Level 1 two-profile solution with unequal variances across profiles based on raw scores

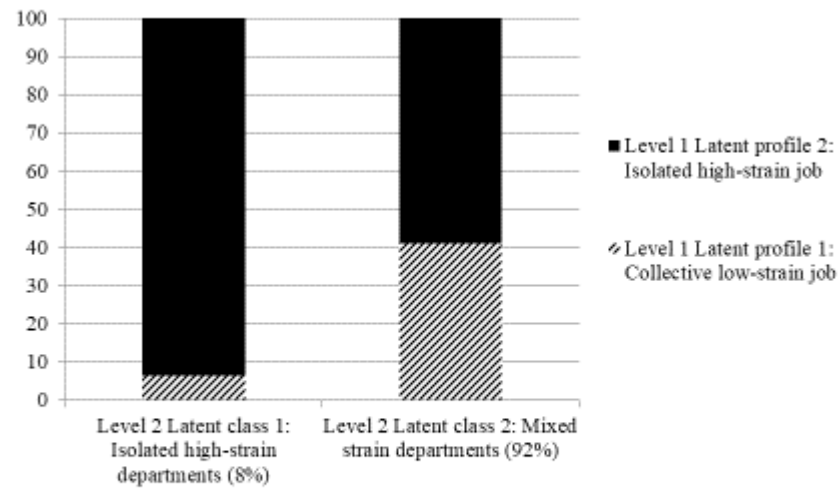


Figure 5.
Level 2 classes based on the relative frequency of the Level 1 profiles