# CHAPTER 7

# APPENDIX

## 7.1 Appendix: STEM-PQA alignment

```
## Error: <text>:2:200: unexpected symbol
## 1: tibble::tribble(
## 2:   ~Work.With.Data.Codes.Originally.Proposed,
##
```

## 7.2 Appendix: Method additional materials

### 7.2.1 Statistical software developed

The functions in tidyLPA dynamically generate MPlus syntax, so that, for example, a user can simply provide a data frame with variables to be used in the analysis, the specification for one of six models, the number of profiles to be estimated as part of the analysis, and a number of fine-grained options concerning the estimation and the output generated. From these inputs, a data file for MPlus is prepared and saved, the model syntax is created and saved in a model input file, the model is run, and the output, including the "savedata", or the data with its associated posterior probabilities and profile assignments, is returned to R for use plots or in subsequent analyses.

Because of the considerable time that it takes to generate MPlus model syntax (i.e., when choosing to specify a model with different parameters or when changing the number of profiles to be estimated as part of the solution), this package makes it easier to carry out LPA in a flexible way, while retaining the power of the MPlus software. While this functionality makes it considerably easier to carry out LPA, it requires that MPlus be purchased and installed. Because of this, the R package I developed also includes wrapper

Table 7.1: Correlations among codes for instructional support for work with data (and composite of all codes)

| rowname | Asking.Questions | Making.Observations | Generating.Data | Data.Modeling |
|---|---|---|---|---|
| Asking Questions | | .38 | .28 | .43 |
| Making Observations | .38 | | .24 | .18 |
| Generating Data | .28 | .24 | | .30 |
| Data Modeling | .43 | .18 | .30 | |
| Communicating Findings | .73 | .55 | .65 | .67 |

functions to an open-source tool, mclust (Scrucca, Fop, Murphy, & Raftery, 2016). This is a very widely-used package for mixture modeling. While some authors have suggested that it can be used to carry out LPA (Oberski, 2016), a key challenge for analysts using it concerns specifying the models. This is because the models are described in terms of the geometric properties of the multivariate distributions being estimated (i.e., "spherical, equal volume"), rather than in terms of whether and how the means, variances, and covariances are estimated. This R package corresponds LPA models to the mclust models and provides the same functionality that the functions that use MPlus provide, namely, preparing data, running the model, and returning the output or use in subsequent analyses. As part of incorporating the mclust functionality, the functions that use MPlus and those that use mclust have been benchmarked (Rosenberg, 2018). Despite leading to identical results (in most cases) for small datasets, because of differences in how the E-M algorithm is initialized as well as other estimation-related differences, output will likely not be identical for many analyses.

### 7.2.2 Appendix: Descriptive statistics additional materials

The Spearman rank (because the data were dichotomous) correlations among the aspects of instructional support for work with data are presented. The variables were moderately correlated, with *rho* values between .18 and .50. These suggest that signals are assocaited

### 7.2.3 Appendix: Program descriptions

Table 7.2: STEM Enrichment Program Names and Their Descriptions

| Program.Name | Program.Description |
|---|---|
| Island Explorers | A science-focused program that aims to help youth develop expertise on one species found in the local ecosys |
| The Ecosphere | A science-focused program that aims to help youth to explore the marine life of Narragansett Bay. Efforts we |
| Zoology Partners | A science-focused program that aims to support youth's development of content knowledge related to the issu |
| Marine Investigators | A science-focused program that aims to provide youth with opportunities to learn about and experience Narr |
| Comunidad de Aprendizaje | A STEM-focused program that aims to help youth improve basic skills in mathematics and develop an intere |
| Jefferson House | A STEM-focused program that aims to support youth's development of basic math skills, the program was p |
| Uptown Architecture | An engineering-focused program that aims to support youth's participation in a process to design and build a |
| Building Mania | An engineering-focused program that aims to provide youth with the opportunity to experiment with designi |
| Adventures in Mathematics | A mathematics-focused program that aims to help youth to develop the basic math skills and prevent summe |

### 7.2.4 Appendix: Research Question #1 additional materials

### 7.2.5 Model specifications details

Here, the six models that are possible to specify in LPA are described in terms of how the variables used to create the profiles are estimated. Note that $p$ represents different profiles and each parameterization is represented by a 4 x 4 covariance matrix and therefore would represent the parameterization for a four-profile solution. In all of the models, the means are estimated freely in the different profiles. Imagine that each row and column represents a different variable, i.e., the first row (and column) represents broad interest, the second enjoyment, the third self-efficacy, and the fourth another variable, i.e., future goals and plans. Models 1 and 3 meet the assumption of independence, that is, that, after accounting for their relations with the profile, the variables used to estimate the profiles are independent (Collins & Lanza, 2010). They estimate variable variances but do not estimate covariances (i.e., as can be seen, the covariance matrices are "diagonal," without any off-diagonal parameters that are estimated). These models are estimated by default in MPlus, although these assumptions can be relaxed (Muthen & Muthen, 2017). Importantly, this does not mean the variables used to create the profile are assumed to be not related; as Collins and Lanza (2010) explain:

> The local independence assumption refers only to conditioning on the latent variable. It does not imply that in a data set that is to be analyzed, the observed variables are independent. In fact, it is the relations among the observed variables that are explained by the latent classes. An observed data set is a mixture of all the latent classes. Independence is assumed to hold only within each latent class, which is why it is called "local".

Despite the assumption of independence, as Collins and Lanza (2010), Muthen and Muthen (2017), and others (i.e., Pastor et al., 2007; Vermunt & Magidson, 2002) note, it can be lifted to improve model fit, though these models without the assumption of independence may be better described as general or Gaussian mixture models (Fraley et al., 2017).

#### 7.2.5.1 Varying means, equal variances, and covariances fixed to 0 (model 1)

In this model, which corresponds to the mclust model wit the name “EEI”, the variances are estimated to be equal across profiles, indicated by the absence of a p subscript for any of the diagonal elements of the matrix. The covariances are constrained to be zero, as indicated by the 0’s between every combination of the variables. Thus, this model is highly constrained but also parsimonious: the profiles are estimated in such a way that the variables’ variances are identical for each of the profiles, and the relationships between the variables are not estimated. In this way, less degrees of freedom are taken used to explain the observations that make up the data. However, estimating more parameters–as in the other models–may better explain the data, justifying the addition in complexity that their addition involves (and their reduction in degrees of freedom).

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

#### 7.2.5.2 Varying means, equal variances, and equal covariances (model 2)

This model corresponds to the mclust model “EEE”. In this model, the variances are still constrained to be the same across the profiles, although now the covariances are estimated (but like the variances, are constrained to be the same across profiles). Thus, this model is the first to estimate the covariance (or correlations) of the variables used to create the profiles, thus adding more information that can be used to better understand the characteristics of the profiles (and, potentially, better explain the data).

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{12} & \sigma_3^2 & \sigma_{33} \\ \sigma_{14} & \sigma_{12} & \sigma_{12} & \sigma_4^2 \end{bmatrix}$$

#### 7.2.5.3 Varying means, varying variances, and covariances fixed to 0 (model 3)

This model corresponds to the mclust model "VVI" and allows for the variances to be freely estimated across profiles. The covariances are constrained to zero. Thus, it is more flexible (and less parsimonious) than model 1, but in terms of the covariances, is more constrained than model 2.

$$\begin{bmatrix} \sigma_{1p}^2 & 0 & 0 & 0 \\ 0 & \sigma_{2p}^2 & 0 & 0 \\ 0 & 0 & \sigma_{3p}^2 & 0 \\ 0 & 0 & 0 & \sigma_{4p}^2 \end{bmatrix}$$

#### 7.2.5.4 Varying means, varying variances, and equal covariances (model 4)

This model, which specifies for the variances to be freely estimated across the profiles and for the covariances to be estimated to be equal across profiles, extends model 3. Unfortunately, this model cannot be specified with mclust, though it can be with MPlus; this model *can* be used with the functions to interface to MPlus described below.

$$\begin{bmatrix} \sigma_{1p}^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{12} & \sigma_{2p}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{12} & \sigma_{3p}^2 & \sigma_{33} \\ \sigma_{14} & \sigma_{12} & \sigma_{12} & \sigma_{4p}^2 \end{bmatrix}$$

#### 7.2.5.5 Varying means, equal variances, and varying covariances (model 5)

This model specifies the variances to be equal across the profiles, but allows the covariances to be freely estimated across the profiles. Like model 4, this model cannot be specified with mclust, though it can be with MPlus. Again, this model *can* be used with the functions to interface to MPlus described below.

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21p} & \sigma_{31p} & \sigma_{41p} \\ \sigma_{12p} & \sigma_2^2 & \sigma_{23p} & \sigma_{24p} \\ \sigma_{13p} & \sigma_{12p} & \sigma_3^2 & \sigma_{33p} \\ \sigma_{14p} & \sigma_{12p} & \sigma_{12p} & \sigma_4^2 \end{bmatrix}$$

#### 7.2.5.6 Varying means, varying variances, and varying covariances (model 6)

This model corresponds to the mclust model "VVV". It allows the variances and the covariances to be freely estimated across profiles. Thus, it is the most complex model, with the potential to allow for understanding many aspects of the variables that are used to estimate the profiles and how they are related. However, it is less parsimonious than all of the other models, and the added parameters should be considered in light of how preferred this model is relative to those with more simple specifications.

$$\begin{bmatrix} \sigma_{1p}^2 & \sigma_{21p} & \sigma_{31p} & \sigma_{41p} \\ \sigma_{12p} & \sigma_{2p}^2 & \sigma_{23p} & \sigma_{24p} \\ \sigma_{13p} & \sigma_{12p} & \sigma_{3p}^2 & \sigma_{33p} \\ \sigma_{14p} & \sigma_{12p} & \sigma_{12p} & \sigma_{4p}^2 \end{bmatrix}$$

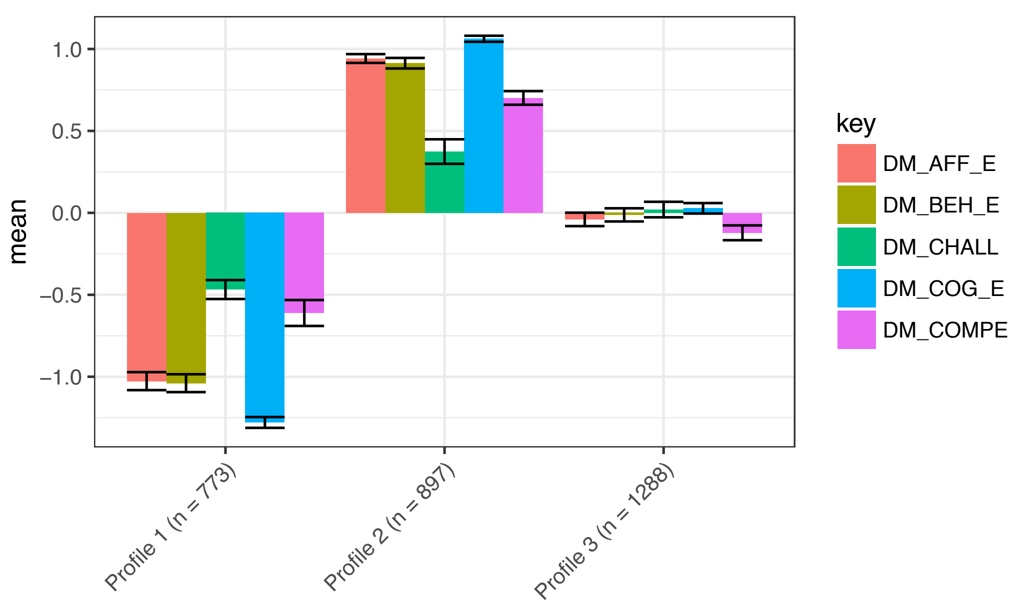
### 7.2.6 Model 1 candidate solutions

#### 7.2.6.1 Model: 1, Profiles: 3

This solution is characterized by:

- a **full** profile, profile 2 (though with more modestly high levels of challenge)
- a **universally low** profile, profile 1 (again with more modestly - in this case low - levels of challenge)
- an **all moderate** profile, profile 3, characterized by levels of all of the variables close to the mean, profile 3

The number of observations associated with each of the profiles is somewhat balanced, with the all moderate profile demonstrating a higher number of observations ($n$ = 1,288) than the full ($n$ = 897) and universally low ($n$ = 773) profiles. The log-likelihood was replicated many (more than 10) times. Because the profiles associated with this solution all demonstrated the same overall pattern (i.e., all five variables are high, low, or moderate), on the basis of interpretability, this particular solution may not be useful in terms of understanding how youth experience engagement and its conditions.



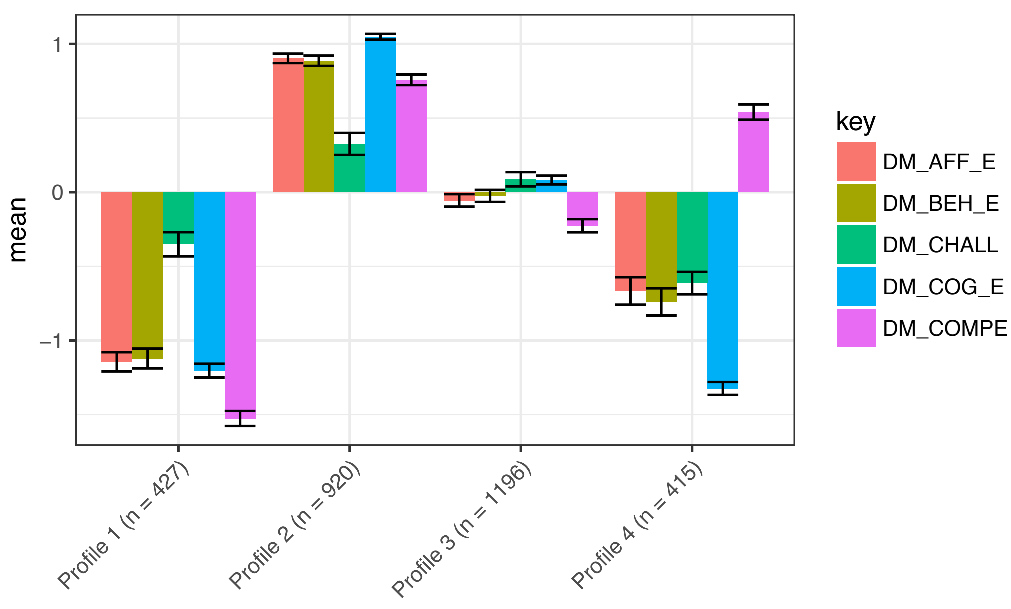
#### 7.2.6.2 Model: 1, Profiles: 4

This solution is characterized by:

- a **full** profile, profile 2
- a **universally low** profile, profile 1
- an **all moderate** profile, profile 3.
- a **competent but not engaged or challenged** profile, with high levels of competence and low levels of engagement and challenge

Most profiles are in the all moderate profile ($n$ = 1,288), with a large number in the full ($n$ = 920) profile, and fewer in the universally low and competent ($n$ n = 427) but not engaged or challenged profiles ($n$ = 415). With somewhat more purchase in terms of its interpretability than the solution for model 1 with three profiles, like that solution, this one may not be as useful as more complex models for understanding youth's experiences.

The log-likelihood was replicated many (more than 10) times.

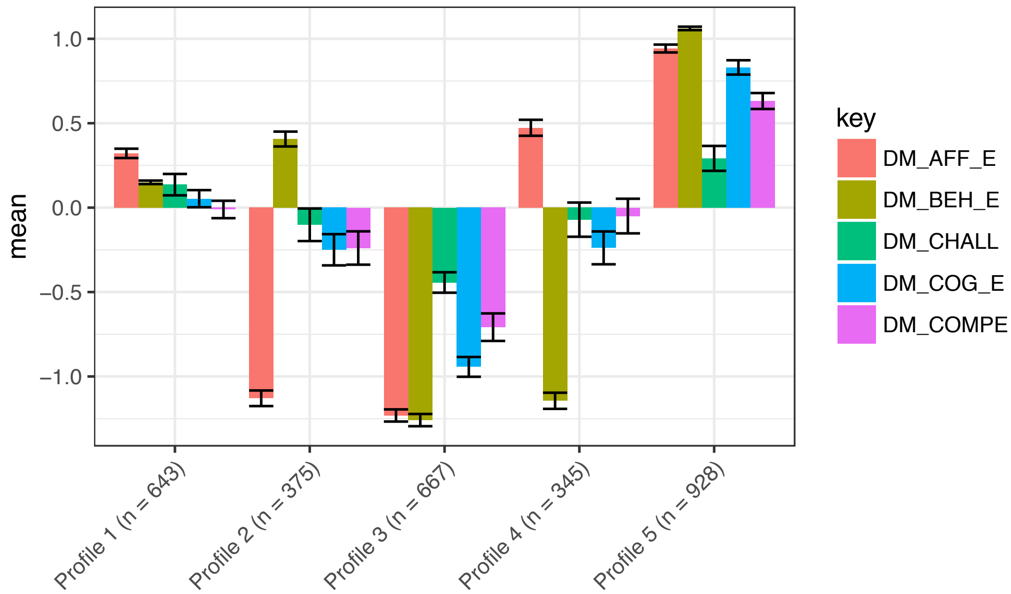



#### 7.2.6.3 Model: 1, Profiles: 5

This solution is characterized by:

- a **full** profile, profile 5
- a **universally low** profile, profile 3
- an **all moderate** profile, profile 3, though with moderate levels of affective engagement than in similar profiles associated with the four and five profile solutions, perhaps suggesting that a different profile than in those solutions
- an **only behavioral** profile, profile 2, with moderate levels of behavioral engagement, very low affective engagement, and moderately (low) levels of cognitive engagement and challenge and competence
- an **only affective** profile, profile 4, with moderate levels of affective engagement, low levels of behavioral engagement, and moderately (low) levels of cognitive engagement and challenge and competence

The number of observations associated with each of the profiles is somewhat balanced, with a large number in the full profile ($n = 928$), a moderate number of observations in the universally low ($n = 667$) and all moderate ($n = 643$) profiles, and fewer observations in the only behaviorally engaged ($n = 375$) and only affective engaged ($n = 345$) profiles. This solution primarily distinguishes between affective and behavioral engagement; unlike the solution for model 1 with four profiles, there is not a competent but not engaged or challenged profile. This may suggest that solutions with a greater number of profiles represents both the distinction between behavioral and affective engagement highlighted by profiles in this solution as well as profiles that are characterized by higher or lower levels of the conditions for engagement (i.e., competence). The log-likelihood was replicated four times.

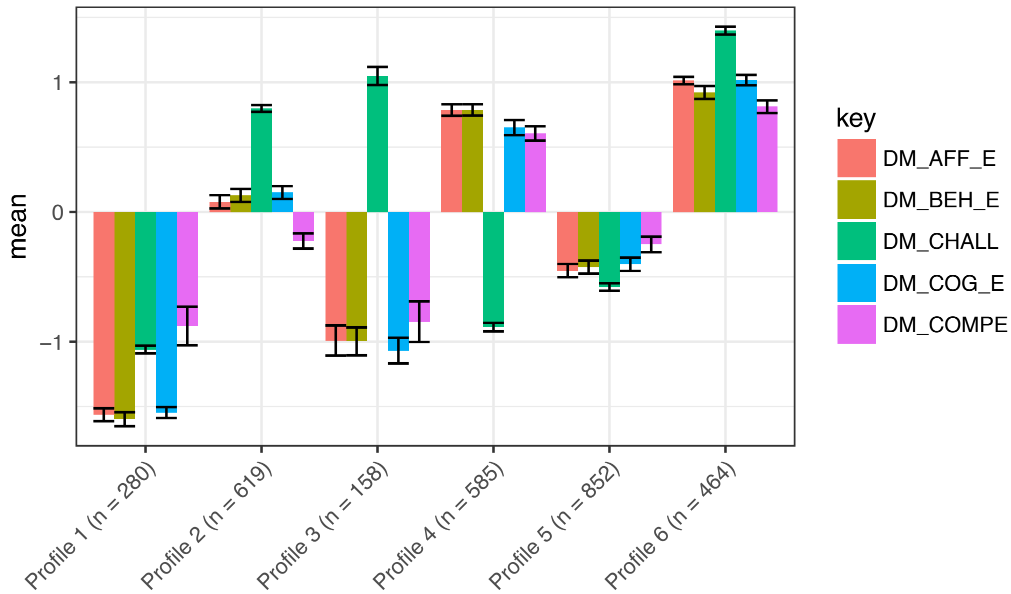



#### 7.2.6.4 Model: 1, Profiles: 6 (alternate)

This solution is characterized by:

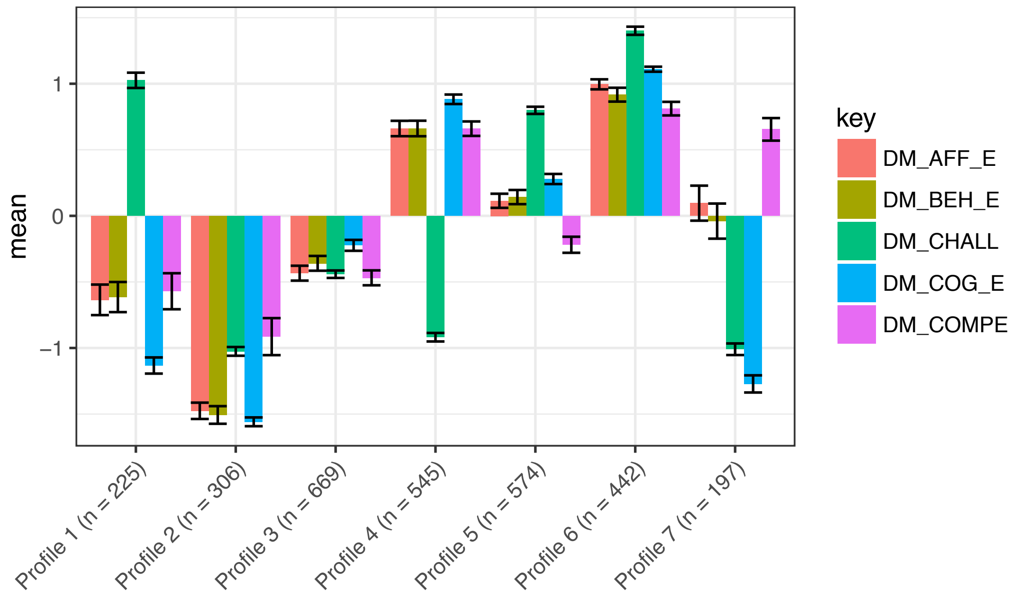
- a **full** profile, profile 6
- a **universally low** profile, profile 1
- an **engaged and competent but not challenged** profile, profile 3
- a **challenged** profile, profile 2
- a **highly challenged** profile, profile 3
- a **moderately low** profile, profile 5

The number of observations are not very balanced, with the moderately low profile with a large number of observations ($n = 852$) and the challenged, engaged and competent but not challenged, and full profiles with moderate numbers of observations (from 464 to 619 observations), and low numbers of observations exhibited by universally low ($n = 280$) and highly challenged ($n = 158$) profiles. This–and, critically, the lower log-likelihood of the other model 1, six profile solution–suggests that this solution is not preferred. However, the very different profiles that emerge for this solution suggest that there might not be a somewhat under-identified solution associated with model 1 and six profiles.




#### 7.2.6.5 Model: 1, Profiles: 7 (alternate)

When investigating an alternate solution (associated with the second lowest log-likelihood) for the model 1, seven profile solution, we can see that even for the solutions associated with other log-likelihoods, the profiles that can be identified are very similar. One minor distinction concerns the **competent but not engaged or challenged** profile, which in the alternate solution is associated with neutral levels of affective engagement, compared to moderately low levels of affective engagement in the solution with the lowest log-likelihood. Because five of the seven profiles associated with both of these model 1, seven profile solutions seem to be distinct from those identified from simpler model 1 solutions, investigation of this alternate solution provides additional evidence that these profiles are not associated with an under-identified model and that simpler models may be preferred over these seven profile solutions.



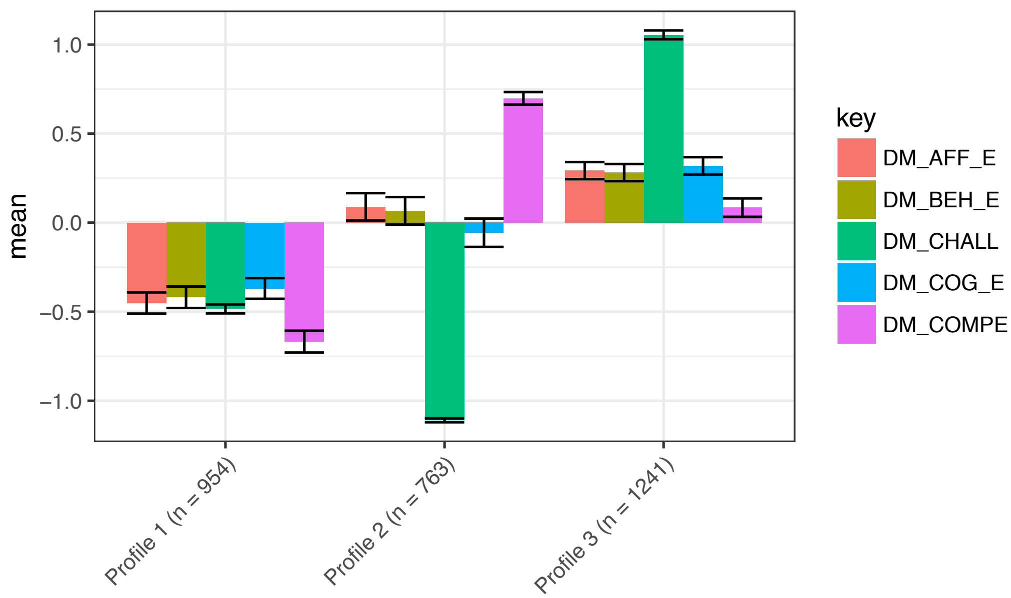

### 7.2.7 Model 2 candidate solutions

#### 7.2.7.1 Model: 2, Profiles: 3

This solution is characterized by:

- a **universally low** profile, profile 1, associated with moderate (low) and low levels of all of the variables; this profile is similar to the universally low profile identified as part of other solutions, although with more moderate values for some of the variables (especially cognitive engagement)
- a **competent but not challenged** profile, profile 2, characterized by high competence and low challenge
- a **challenged** profile, profile 3, characterized by very high challenge and moderate (high) levels of the other variables, similar to the challenged profile found as part of the model 1, four profile solution, but with higher levels of competence, which are moderately high in this solution but moderately low for the other solution.

The number of observations associated with each solution is fairly balanced, with the most in the challenged profile ($n$ = 1,241), followed by the universally low ($n$ = 954 observations) and competent but not challenged ($n$ = 763) profiles. This solution is very different than the three profile solution that was interpreted for model 1. Model 2 differs from model 1 in that covariances between the variables are estimated (they are constrained to be the same are across the profiles). The log-likelihood was replicated (at least) ten times. Thus, this and other solutions associated with model 2 include information about how the variables relate. Including this information seems to be associated with profiles that differentiate the groups on the basis of the levels of each of the variables in more distinct ways: the model 1, three profile solution was characterized by high, moderate, or low levels of all variables for each of the three profiles.

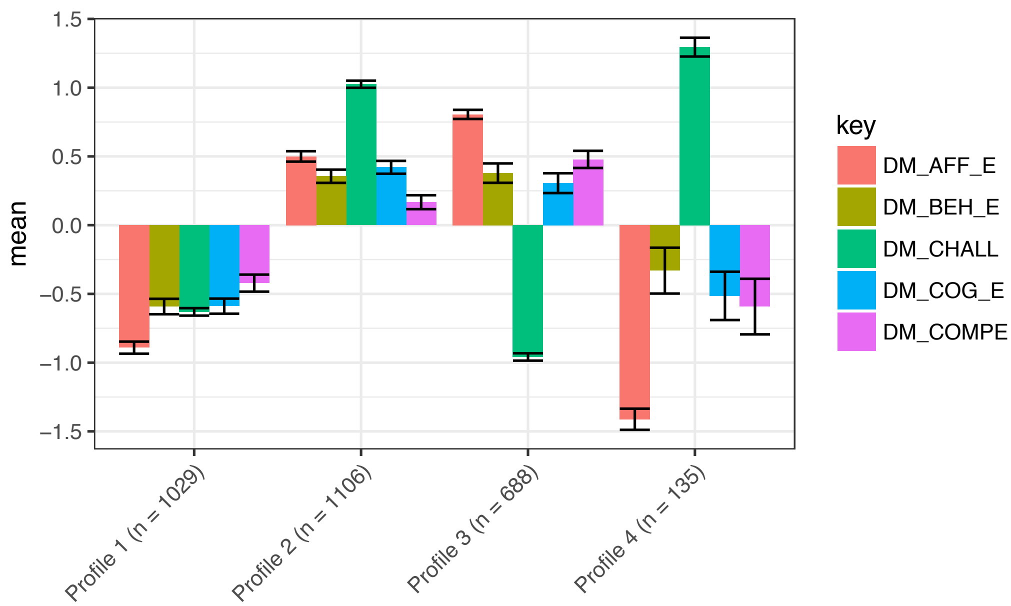



#### 7.2.7.2 Model: 2, Profiles: 4

This solution is characterized by:

- a **universally low** profile, profile 1
- a **challenged** profile, profile 2
- a **highly challenged** profile, profile 4
- an **engaged and competent but not challenged** profile, profile 3

The number of observations in each of the profiles is not very balanced, with more than 1,000 observations in both the universally low ($n$ = 1,029) and challenged ($n$ = 1,106) profiles, a moderate number if the engaged and competent but not challenged profile ($n$ = 688), and very few in the highly challenged ($n$ = 135) profile. The log-likelihood was replicated three times. While each of these profiles has been identified in another solution, the small number of observations in the highly challenged profile suggests that this solution be interpreted with some skepticism because of the potentially limited utility (and statistical power associated with the use) of the profiles in subsequent analyses.



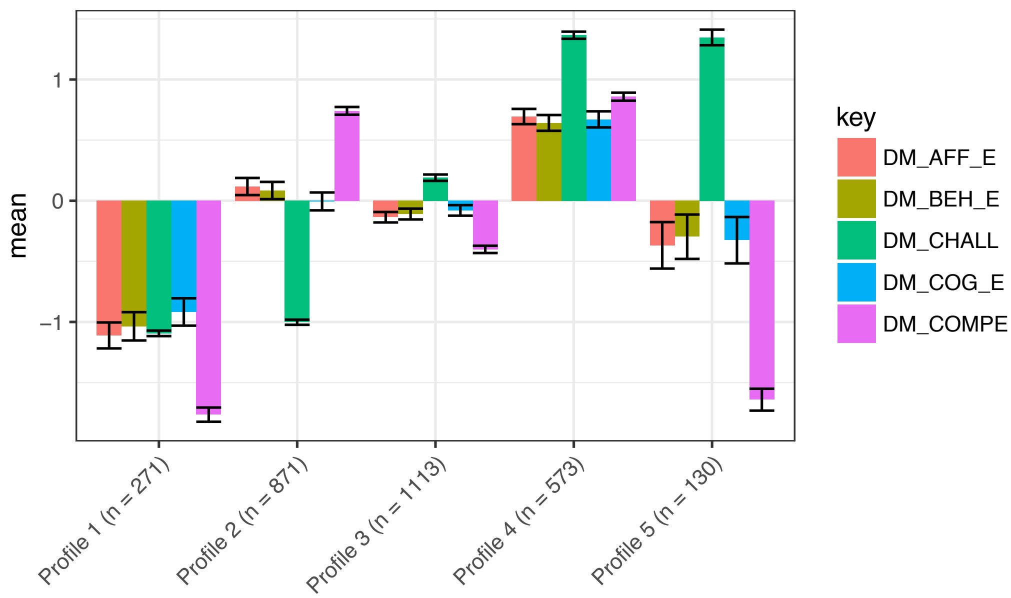

#### 7.2.7.3 Model: 2, Profiles: 5

This solution is characterized by:

- a **universally low** profile, profile 1
- a **full** profile, profile 4, although with very high levels of challenged (in addition to high levels of all of the other variables), making this profile similar to that (challenged) profile
- a **highly challenged** profile, profile 5

- an **all moderate** profile, profile 3, although with moderately lower levels of competence than is found in profiles associated with other solutions
- a **competent but not challenged** profile, profile 2, similar to the competent but not challenged or engaged profile, but with neutral, rather than low, levels of the engagement variables

The number of observations associated with each of the profiles is not very balanced, with a very large number of observations in the all moderate profile ($n$ = 1,113) and a large number in the competent but not challenged profile ($n$ = 871), a moderate number in the full profile ($n$ = 573), and very few in the universally low ($n$ = 271) and challenged but not competent ($n$ = 130) profiles. The log-likelihood was replicated four times. Like for the model 2, four profile solution, the small number of observations associated with two of the profiles suggests that this solution should be interpreted with some caution.




## 7.3 Appendix: Models for research question #2 and #3 with the seven-profile solution

Table 7.3: Results of mixed effects models for instructional support for work with data as separate variables

| model | intercept | dm_ask | dm_obs | dm_gen | dm_mod | dm_com | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|---|---|---|---|
| Universally low | 0.072 (0.01) (p < .001) | -0.011 (0.011) (p = 0.839) | 0.001 (0.011) (p = 0.457) | -0.005 (0.01) (p = 0.686) | -0.002 (0.011) (p = 0.58) | -0.002 (0.01) (p = 0.583) | 0.008 | 0.176 | 0.002 |
| Competent but not engaged or challenged | 0.111 (0.019) (p < .001) | 0.002 (0.013) (p = 0.428) | -0.007 (0.013) (p = 0.696) | -0.01 (0.012) (p = 0.78) | -0.025 (0.014) (p = 0.969) | -0.002 (0.013) (p = 0.549) | 0.014 | 0.222 | 0.021 |
| Moderately low | 0.186 (0.029) (p < .001) | 0.024 (0.017) (p = 0.077) | 0.024 (0.017) (p = 0.074) | 0.009 (0.016) (p = 0.281) | -0.002 (0.018) (p = 0.544) | 0.027 (0.017) (p = 0.053) | 0.016 | 0.279 | 0.033 |
| Challenged | 0.197 (0.023) (p < .001) | 0.011 (0.016) (p = 0.252) | -0.013 (0.016) (p = 0.799) | 0.034 (0.015) (p = 0.015) | -0.005 (0.017) (p = 0.626) | -0.002 (0.016) (p = 0.561) | 0.007 | 0.261 | 0.015 |
| Highly challenged | 0.083 (0.01) (p < .001) | -0.007 (0.012) (p = 0.714) | 0 (0.012) (p = 0.504) | -0.016 (0.012) (p = 0.913) | 0.017 (0.013) (p = 0.1) | -0.013 (0.012) (p = 0.859) | 0.023 | 0.104 | 0.000 |
| Engaged and competent but not challenged | 0.199 (0.018) (p < .001) | -0.014 (0.017) (p = 0.787) | 0.019 (0.017) (p = 0.129) | -0.033 (0.016) (p = 0.979) | -0.013 (0.018) (p = 0.77) | 0.019 (0.017) (p = 0.127) | 0.015 | 0.286 | 0.000 |
| Full | 0.154 (0.023) (p < .001) | -0.009 (0.014) (p = 0.748) | -0.024 (0.014) (p = 0.961) | 0.022 (0.013) (p = 0.048) | 0.032 (0.014) (p = 0.012) | -0.025 (0.013) (p = 0.97) | 0.018 | 0.513 | 0.012 |

Table 7.4: Results of mixed effects models for the composite

| model | intercept | dm_composite | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|
| Universally low | 0.071 (0.01) (p < .001) | -0.004 (0.003) (p = 0.937) | 0.007 | 0.176 | 0.002 |
| Competent but not engaged or challenged | 0.113 (0.019) (p < .001) | -0.008 (0.003) (p = 0.993) | 0.013 | 0.222 | 0.022 |
| Moderately low | 0.188 (0.029) (p < .001) | 0.016 (0.004) (p < .001) | 0.015 | 0.279 | 0.032 |
| Challenged | 0.2 (0.023) (p < .001) | 0.005 (0.004) (p = 0.09) | 0.007 | 0.260 | 0.015 |
| Highly challenged | 0.08 (0.01) (p < .001) | -0.004 (0.003) (p = 0.917) | 0.022 | 0.104 | 0.000 |
| Engaged and competent but not challenged | 0.2 (0.018) (p < .001) | -0.005 (0.004) (p = 0.887) | 0.015 | 0.285 | 0.000 |
| Full | 0.151 (0.023) (p < .001) | 0 (0.003) (p = 0.527) | 0.019 | 0.510 | 0.013 |

Table 7.5: Results of mixed effects models for the composite

| model | intercept | dm_composite_di | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---:|---:|---:|
| Universally low | 0.078 (0.012) (p < .001) | -0.02 (0.01) (p = 0.98) | 0.006 | 0.176 | 0.003 |
| Competent but not engaged or challenged | 0.118 (0.019) (p < .001) | -0.03 (0.012) (p = 0.993) | 0.014 | 0.223 | 0.021 |
| Moderately low | 0.175 (0.029) (p < .001) | 0.063 (0.016) (p < .001) | 0.015 | 0.279 | 0.031 |
| Challenged | 0.203 (0.024) (p < .001) | 0.01 (0.015) (p = 0.261) | 0.008 | 0.260 | 0.014 |
| Highly challenged | 0.083 (0.011) (p < .001) | -0.016 (0.011) (p = 0.922) | 0.023 | 0.104 | 0.000 |
| Engaged and competent but not challenged | 0.198 (0.019) (p < .001) | -0.011 (0.016) (p = 0.763) | 0.015 | 0.285 | 0.000 |
| Full | 0.146 (0.024) (p < .001) | 0.006 (0.013) (p = 0.308) | 0.019 | 0.510 | 0.014 |

Table 7.6: Results of mixed effects models with interest and other characteristics

| model | intercept | overall_pre_interest | gender_female | urm | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|---|---|
| Universally low | 0.069 (0.04) (p = 0.042) | -0.002 (0.01) (p = 0.569) | 0.017 (0.018) (p = 0.177) | -0.005 (0.025) (p = 0.576) | 0.016 | 0.194 | 0.000 |
| Competent but not engaged or challenged | 0.126 (0.049) (p = 0.006) | -0.011 (0.013) (p = 0.808) | 0.021 (0.022) (p = 0.165) | -0.007 (0.031) (p = 0.591) | 0.017 | 0.211 | 0.003 |
| Moderately low | 0.301 (0.082) (p < .001) | -0.023 (0.021) (p = 0.869) | 0.02 (0.034) (p = 0.275) | -0.018 (0.048) (p = 0.642) | 0.017 | 0.290 | 0.030 |
| Challenged | 0.297 (0.073) (p < .001) | -0.013 (0.019) (p = 0.757) | -0.04 (0.031) (p = 0.9) | -0.032 (0.044) (p = 0.768) | 0.013 | 0.254 | 0.016 |
| Highly challenged | 0.108 (0.034) (p = 0.001) | -0.014 (0.009) (p = 0.946) | -0.013 (0.015) (p = 0.803) | 0.016 (0.021) (p = 0.221) | 0.026 | 0.112 | 0.001 |
| Engaged and competent but not challenged | 0.065 (0.07) (p = 0.177) | 0.033 (0.018) (p = 0.034) | 0.042 (0.032) (p = 0.094) | -0.002 (0.045) (p = 0.518) | 0.013 | 0.275 | 0.000 |
| Full | 0.048 (0.082) (p = 0.28) | 0.022 (0.021) (p = 0.142) | -0.043 (0.037) (p = 0.878) | 0.06 (0.053) (p = 0.129) | 0.023 | 0.510 | 0.000 |

Table 7.7: Results of mixed effects models with interest and other characteristics and the aspects of work with data

| model | intercept | dm_ask | dm_obs | dm_gen | dm_mod | dm_com | overall_pre_interest | gender_female | urm | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Universally low | 0.077 (0.041) (p = 0.031) | -0.01 (0.011) (p = 0.822) | 0.002 (0.011) (p = 0.429) | -0.003 (0.011) (p = 0.612) | -0.003 (0.012) (p = 0.616) | -0.002 (0.011) (p = 0.569) | -0.003 (0.01) (p = 0.63) | 0.017 (0.018) (p = 0.166) | -0.002 (0.025) (p = 0.535) | 0.008 | 0.189 | 0.004 |
| Competent but not engaged or challenged | 0.141 (0.05) (p = 0.003) | 0.002 (0.013) (p = 0.432) | -0.011 (0.013) (p = 0.803) | -0.007 (0.013) (p = 0.718) | -0.028 (0.014) (p = 0.978) | 0.003 (0.013) (p = 0.421) | -0.013 (0.013) (p = 0.839) | 0.021 (0.022) (p = 0.167) | -0.011 (0.03) (p = 0.644) | 0.013 | 0.201 | 0.011 |
| Moderately low | 0.271 (0.084) (p < .001) | 0.024 (0.018) (p = 0.087) | 0.027 (0.017) (p = 0.061) | 0.002 (0.017) (p = 0.447) | 0 (0.018) (p = 0.499) | 0.033 (0.017) (p = 0.03) | -0.024 (0.021) (p = 0.869) | 0.02 (0.034) (p = 0.284) | -0.019 (0.049) (p = 0.649) | 0.014 | 0.288 | 0.036 |
| Challenged | 0.279 (0.075) (p < .001) | 0.018 (0.017) (p = 0.138) | -0.003 (0.017) (p = 0.568) | 0.029 (0.016) (p = 0.036) | -0.005 (0.017) (p = 0.606) | -0.006 (0.017) (p = 0.648) | -0.011 (0.019) (p = 0.725) | -0.042 (0.032) (p = 0.906) | -0.032 (0.045) (p = 0.762) | 0.007 | 0.253 | 0.022 |
| Highly challenged | 0.128 (0.035) (p < .001) | -0.004 (0.013) (p = 0.63) | -0.008 (0.013) (p = 0.722) | -0.016 (0.012) (p = 0.9) | 0.023 (0.014) (p = 0.048) | -0.022 (0.013) (p = 0.951) | -0.015 (0.009) (p = 0.959) | -0.009 (0.015) (p = 0.721) | 0.013 (0.021) (p = 0.263) | 0.025 | 0.108 | 0.000 |
| Engaged and competent but not challenged | 0.071 (0.071) (p = 0.161) | -0.018 (0.017) (p = 0.854) | 0.017 (0.017) (p = 0.164) | -0.027 (0.016) (p = 0.948) | -0.013 (0.018) (p = 0.759) | 0.017 (0.017) (p = 0.168) | 0.034 (0.018) (p = 0.034) | 0.046 (0.032) (p = 0.079) | -0.001 (0.045) (p = 0.509) | 0.013 | 0.279 | 0.000 |
| Full | 0.053 (0.085) (p = 0.266) | -0.015 (0.014) (p = 0.852) | -0.026 (0.014) (p = 0.967) | 0.022 (0.013) (p = 0.048) | 0.029 (0.015) (p = 0.026) | -0.022 (0.014) (p = 0.938) | 0.023 (0.021) (p = 0.148) | -0.043 (0.038) (p = 0.868) | 0.063 (0.055) (p = 0.127) | 0.021 | 0.534 | 0.000 |

Table 7.8: Results of mixed effects models with interest and other characteristics and the composite work with data

| model | intercept | dm_composite | overall_pre_interest | gender_female | urm | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|---|---|---|
| Universally low | 0.077 (0.041) (p = 0.031) | -0.004 (0.003) (p = 0.908) | -0.004 (0.01) (p = 0.632) | 0.017 (0.018) (p = 0.168) | -0.002 (0.025) (p = 0.533) | 0.007 | 0.189 | 0.004 |
| Competent but not engaged or challenged | 0.144 (0.05) (p = 0.002) | -0.008 (0.003) (p = 0.992) | -0.013 (0.013) (p = 0.841) | 0.022 (0.022) (p = 0.159) | -0.012 (0.03) (p = 0.65) | 0.012 | 0.201 | 0.012 |
| Moderately low | 0.273 (0.083) (p < .001) | 0.017 (0.004) (p < .001) | -0.024 (0.021) (p = 0.871) | 0.02 (0.034) (p = 0.278) | -0.019 (0.049) (p = 0.653) | 0.013 | 0.288 | 0.034 |
| Challenged | 0.28 (0.075) (p < .001) | 0.007 (0.004) (p = 0.044) | -0.011 (0.019) (p = 0.717) | -0.042 (0.032) (p = 0.905) | -0.032 (0.045) (p = 0.763) | 0.006 | 0.253 | 0.022 |
| Highly challenged | 0.125 (0.035) (p < .001) | -0.005 (0.003) (p = 0.951) | -0.016 (0.009) (p = 0.964) | -0.01 (0.015) (p = 0.733) | 0.014 (0.021) (p = 0.254) | 0.024 | 0.109 | 0.000 |
| Engaged and competent but not challenged | 0.072 (0.071) (p = 0.156) | -0.006 (0.004) (p = 0.911) | 0.033 (0.018) (p = 0.035) | 0.046 (0.032) (p = 0.078) | 0 (0.045) (p = 0.503) | 0.012 | 0.279 | 0.000 |
| Full | 0.05 (0.085) (p = 0.278) | -0.001 (0.004) (p = 0.657) | 0.023 (0.021) (p = 0.143) | -0.045 (0.038) (p = 0.878) | 0.063 (0.055) (p = 0.126) | 0.022 | 0.532 | 0.000 |

Table 7.9: Results of mixed effects models with the interactions between interest and other characactistics and the composite for work with data

| model | intercept | dm_composite | overall_pre_interest | gender_female | urm | overall_pre_interest:dm_composite | dm_composite:gender_female | dm_composite:urm | beep_ID_ICC | participant_ID_ICC | program_ID_ICC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Universally low | 0.113 (0.046) (p = 0.008) | -0.022 (0.012) (p = 0.973) | -0.014 (0.012) (p = 0.878) | 0.025 (0.021) (p = 0.113) | -0.012 (0.029) (p = 0.663) | 0.005 (0.003) (p = 0.038) | -0.004 (0.005) (p = 0.761) | 0.005 (0.007) (p = 0.225) | 0.007 | 0.189 | 0.004 |
| Competent but not engaged or challenged | 0.133 (0.056) (p = 0.009) | -0.002 (0.013) (p = 0.553) | -0.01 (0.015) (p = 0.761) | 0.031 (0.025) (p = 0.106) | -0.013 (0.035) (p = 0.643) | -0.001 (0.003) (p = 0.656) | -0.005 (0.006) (p = 0.784) | 0.001 (0.008) (p = 0.466) | 0.012 | 0.201 | 0.011 |
| Moderately low | 0.256 (0.089) (p = 0.002) | 0.026 (0.017) (p = 0.071) | -0.017 (0.023) (p = 0.774) | 0.015 (0.038) (p = 0.345) | -0.02 (0.053) (p = 0.643) | -0.003 (0.005) (p = 0.767) | 0.003 (0.008) (p = 0.376) | 0 (0.011) (p = 0.505) | 0.013 | 0.287 | 0.034 |
| Challenged | 0.262 (0.082) (p < .001) | 0.016 (0.017) (p = 0.178) | -0.002 (0.021) (p = 0.542) | -0.063 (0.035) (p = 0.961) | -0.03 (0.05) (p = 0.724) | -0.004 (0.005) (p = 0.83) | 0.011 (0.008) (p = 0.094) | -0.002 (0.011) (p = 0.573) | 0.006 | 0.253 | 0.022 |
| Highly challenged | 0.12 (0.042) (p = 0.002) | -0.003 (0.013) (p = 0.608) | -0.018 (0.011) (p = 0.951) | -0.015 (0.019) (p = 0.778) | 0.03 (0.026) (p = 0.126) | 0.001 (0.003) (p = 0.368) | 0.003 (0.006) (p = 0.317) | -0.008 (0.008) (p = 0.851) | 0.025 | 0.108 | 0.000 |
| Engaged and competent but not challenged | 0.07 (0.078) (p = 0.185) | -0.004 (0.017) (p = 0.593) | 0.029 (0.02) (p = 0.079) | 0.065 (0.036) (p = 0.034) | 0.007 (0.05) (p = 0.441) | 0.002 (0.005) (p = 0.319) | -0.01 (0.008) (p = 0.888) | -0.003 (0.011) (p = 0.624) | 0.012 | 0.277 | 0.000 |
| Full | 0.073 (0.088) (p = 0.205) | -0.014 (0.013) (p = 0.848) | 0.021 (0.022) (p = 0.172) | -0.057 (0.04) (p = 0.921) | 0.049 (0.057) (p = 0.196) | 0.001 (0.003) (p = 0.38) | 0.006 (0.006) (p = 0.148) | 0.007 (0.008) (p = 0.202) | 0.022 | 0.532 | 0.000 |