# Case study: Evaluating multilevel latent profile models of course engagement with the latents R package

Mohammed Saqr, University of Eastern Finland
Sonsoles López-Pernas, University of Eastern Finland

26 November 2026

## Contents

Data and starting model . . . . . 2
Classification uncertainty . . . . . 4
Residual dependence . . . . . 6
Class counts and covariance assumptions . . . . . 7
Uncertainty in profile means . . . . . 11
Limits of the interpretation . . . . . 13
References . . . . . 13

Multilevel latent profile analysis (MLPA) estimates latent profiles at the observation level while simultaneously modeling how the probabilities of those profiles vary across higher-level latent classes. In this framework, the observation-level profiles are defined by their characteristic indicator means, whereas the higher-level classes are distinguished by the probabilities with which those profiles occur.

Model evaluation therefore involves several related questions. First, it is important to determine whether the estimated profiles provide meaningful separation among observations. Second, the measurement assumptions should be examined to assess whether the model adequately represents the relationships among the observed indicators. Third, alternative model specifications can be compared to determine whether a different number of observation-level profiles or higher-level classes provides a more appropriate representation of the data.

The three-profile, three-class model introduced in the introductory vignette serves as the starting point for these evaluation procedures. The following sections use this model as a reference while examining profile separation, measurement assumptions, and alternative model specifications.

```
library(latents)
```

## Data and starting model

The `engagement_momentum` dataset contains repeated course enrolments nested within students. The bundled data are synthetic, and their relationship to the longitudinal engagement studies of Saqr and López-Pernas (2021) and López-Pernas and Saqr (2024) is described in the introductory vignette. Engagement is measured using five course-level indicators: course-page views (`browse`), lecture-video views (`lectures`), forum posts read (`forum_read`), forum posts written (`forum_post`), and days active (`active_days`). Each indicator is simulated as a continuous value representing a log-activity scale and standardized within course. The vector `vars` holds their names.

```
vars <- c("browse", "lectures", "forum_read", "forum_post", "active_days")
```

The variable `student` identifies the higher-level unit, while `sequence` records the order of each student's course enrolments. In the present model, enrolments are treated as observations nested within students. Courses themselves are not modeled as an additional grouping level. The multilevel structure therefore distinguishes between variation across enrolments and variation across students.

The starting model contains three Gaussian enrolment-level profiles and three student-level latent classes. A central quantity in the model is the conditional probability

$$\pi_{k|m} = P(C_{it} = k \mid G_i = m),$$

where $C_{it}$ denotes the latent profile of student $i$ at occasion $t$, and $G_i$ denotes that student's higher-level latent class. Each probability lies between zero and one, and the profile probabilities sum to one within each student class. These probabilities describe how frequently the different engagement profiles are expected to occur among students belonging to each latent class. The profile-specific means and variances are shared across student classes, so the classes differ in their profile probabilities rather than in the definitions of the profiles themselves.

The model uses a diagonal covariance structure, which assumes that the engagement indicators are conditionally independent within each latent profile. It also assumes that enrolments are conditionally independent given the student's latent class and that students are independent of one another. In practice, these assumptions may be violated. For example, indicators may remain associated because several measures reflect related forms of online activity, engagement may persist across adjacent courses taken by the same student, or students may share unmodeled course-level influences.

Like other finite-mixture models, MLPA can also be sensitive to local likelihood maxima. Using multiple starting values reduces the risk that estimation will settle on a suboptimal local solution, although it does

not guarantee that the global maximum has been identified. Model evaluation should therefore consider both the substantive interpretation of the solution and evidence concerning the stability of the likelihood across different starting values.

The `multilpa()` function estimates the model by maximum likelihood. A stricter convergence tolerance (`tol`) is used here so that the optimization proceeds closer to a stationary solution before standard errors are calculated. This provides a more stringent convergence criterion for the starting model used in the subsequent evaluation steps.

```
fit <- multilpa(engagement_momentum, vars, id = "student", time = "sequence",
                n_profiles = 3, n_group_classes = 3, n_starts = 10,
                max_iter = 2000, tol = 1e-10, seed = 1)
fit
#> Two-level latent profile analysis: 3 profiles, 3 group classes
#> 1310 individuals in 106 groups; varying diagonal residual covariance (VVI)
#> Log likelihood: -7487.954642 | AIC: 15051.909 | BIC (groups): 15153.120
#> Converged: TRUE | iterations: 24 | best log likelihood reached by 10 of 10 starts (10 converged)
#> Smallest standardized Gaussian eigenvalue: 0.3005 (profiles, profile_1); Gaussian mass: 442.5; weight effec
#>
#> Class sizes   profile_1 33.8% (442)   profile_2 24.4% (320)   profile_3 41.8% (548)
#> Group classes group_class_1 19.7% (21)  group_class_2 44.6% (47)  group_class_3 35.7% (38)
#>
#> Means
#>    indicator profile_1 profile_2 profile_3
#>       browse    -0.928    -0.012     0.756
#>     lectures    -0.794    -0.142     0.724
#>   forum_read    -0.980    -0.079     0.837
#>   forum_post    -0.873    -0.055     0.737
#>  active_days    -0.961    -0.029     0.793
#>
#> Variances and standard errors: get_results(x, "profiles").
#> Every other table: get_results(x, what = ), or get_results(x, "all").
```

The fitted model summarizes 1,310 course enrolments contributed by 106 students. The three estimated profiles show a clear ordering across the engagement indicators. Profile 3 has the highest mean on every indicator, Profile 2 has means close to the course average, and Profile 1 has the lowest means; following the introductory vignette, they are the **active**, **average** and **disengaged** profiles. For example, the estimated means of days active are 0.793, -0.029, and -0.961 standard deviations relative to the corresponding course mean for Profiles 3, 2, and 1, respectively.

These differences in indicator means define the substantive interpretation of the fitted profiles. In this example, the profiles can therefore be understood as representing progressively lower levels of engagement across the observed activity measures. However, the presence of distinct estimated means does not by itself establish that the profiles are well supported. The diagnostic checks that follow examine the extent to which the data support this separation and whether the fitted solution is numerically stable.

The command `get_results(fit, "starts")` summarizes the outcome of each model initialization, including convergence information and the maximized log likelihood. When multiple converged starts arrive at the same or very similar log-likelihood value, this provides evidence that the selected solution is numerically stable. By contrast, substantially different likelihood values across converged starts may indicate the presence of competing local maxima. Similarly, variance estimates that reach active parameter bounds can signal estimation difficulties or an inadequately specified solution. Either pattern warrants further investigation before the fitted profiles are interpreted in detail.

```
get_results(fit, "starts")
#>    start log_likelihood converged iterations error boundary
#> 1      1       -7487.95      TRUE         24  <NA>    FALSE
#> 2      2       -7487.95      TRUE         24  <NA>    FALSE
#> 3      3       -7487.95      TRUE         24  <NA>    FALSE
#> 4      4       -7487.95      TRUE         24  <NA>    FALSE
#> 5      5       -7487.95      TRUE         24  <NA>    FALSE
#> 6      6       -7487.95      TRUE         24  <NA>    FALSE
#> 7      7       -7487.95      TRUE         24  <NA>    FALSE
#> 8      8       -7487.95      TRUE         24  <NA>    FALSE
#> 9      9       -7487.95      TRUE         24  <NA>    FALSE
#> 10    10       -7487.95      TRUE         24  <NA>    FALSE
```

All ten starting values converged to the same solution, with no reported boundary issues and a maximized log likelihood of approximately -7487.955. This replication across starts provides evidence that the fitted solution is numerically stable. The comparisons of alternative profile and class counts below therefore begin from a solution that was recovered consistently across multiple initializations.

## Classification uncertainty

An important aspect of model evaluation is the degree of uncertainty in assigning observations to latent profiles or classes. Relative entropy summarizes this uncertainty by measuring how concentrated the posterior membership probabilities are. For $N$ units, $K$ latent classes, and posterior probability $p_{uk}$ that unit $u$ belongs to class $k$, relative entropy is defined as:

$$E = 1 - \frac{-\sum_{u=1}^{N}\sum_{k=1}^{K} p_{uk}\log p_{uk}}{N\log K}.$$

Relative entropy ranges from zero to one. A value of zero occurs when each unit has equal posterior probability of belonging to every class, which is maximal classification uncertainty. A value of one occurs when every unit has posterior probability one for a single class and zero for all others. Values between these extremes reflect varying degrees of certainty in the posterior assignments.

In a multilevel latent profile model, relative entropy is calculated separately for the enrolment and student levels. Enrolment-level entropy summarizes the concentration of posterior probabilities for the observation-level profiles, whereas student-level entropy summarizes uncertainty in assignment to the higher-level latent classes. These quantities should therefore be interpreted separately because classification may be more distinct at one level than at the other.

Relative entropy is conditional on the fitted model and should not be interpreted as a general measure of model correctness. A misspecified model can still produce highly concentrated posterior probabilities. For example, inappropriate assumptions about indicator independence or the number of latent profiles may yield confident classifications even when the model provides an inadequate representation of the data. Entropy should therefore be considered alongside other diagnostic and model-comparison measures. Relative entropy is also undefined at any level containing only a single latent class because there is no classification uncertainty to evaluate.

The command `get_results(fit, "entropy")` reports both the unnormalized entropy sum and the corresponding relative entropy for the enrolment and student levels. These results provide a descriptive summary of how clearly the fitted model assigns units to the latent profiles and classes at each level.

```
get_results(fit, "entropy")
#>         level n_classes n_units entropy_sum relative_entropy
```

```
#> 1 individuals         3    1310    269.5407         0.812713
#> 2      groups         3     106     19.0746         0.836203
```

The fitted model has a relative entropy of 0.813 at the enrolment level and 0.836 at the student level. These values indicate that posterior membership probabilities are reasonably concentrated at both levels, but that classification remains uncertain for some enrolments and students. Consequently, reducing each posterior probability distribution to a single most likely profile or class results in a loss of information. Where possible, subsequent analyses should therefore retain the uncertainty represented by the full posterior probabilities.

A complementary measure of classification certainty is the average posterior probability within each assigned class. Let $A_k$ denote the set of units whose most probable assignment is to class $k$. The average posterior probability for that class is

$$\bar{p}_k = \frac{1}{|A_k|} \sum_{u \in A_k} p_{uk},$$

where $p_{uk}$ is the posterior probability that unit $u$ belongs to class $k$. This quantity summarizes, among units assigned to a particular class, how strongly the model supports those assignments. Values closer to one indicate that units assigned to the class generally have high posterior probabilities for that class, whereas lower values indicate greater ambiguity in the assignments.

Average posterior probabilities provide class-specific information that is not visible from a single overall entropy measure. For example, a model may show relatively high overall entropy while still containing one profile or class whose members are considerably more difficult to classify than those in the others. Examining average posterior probabilities by class can therefore help identify where classification uncertainty is concentrated.

Like relative entropy, however, average posterior probability is conditional on the fitted model and its estimated classes. It describes the certainty of the resulting modal assignments but does not establish that the chosen number of classes, measurement assumptions, or substantive interpretation is correct. Classification diagnostics should therefore be considered alongside other evidence about model fit and specification.

The command `get_results(fit, "classification")` reports these class-specific classification summaries together with the number of units assigned to each profile or class. The class counts are also useful for identifying very small groups, for which classification summaries and subsequent substantive comparisons may be less stable.

```
get_results(fit, "classification")
#>       level class n_modal proportion_modal estimated_n estimated_proportion average_posterior
#> 1 individuals     1     441        0.336641    442.4812            0.337772         0.952706
#> 2 individuals     2     320        0.244275    319.5231            0.243911         0.826204
#> 3 individuals     3     549        0.419084    547.9957            0.418317         0.937542
#> 4      groups     1      21        0.198113     20.8314            0.196523         0.950178
#> 5      groups     2      47        0.443396     47.2963            0.446191         0.919908
#> 6      groups     3      38        0.358491     37.8723            0.357286         0.920675
#>   odds_correct_classification
#> 1                     39.4945
#> 2                     14.7363
#> 3                     20.8730
#> 4                     77.9736
#> 5                     14.2558
#> 6                     20.8785
```

The class-specific results show that classification certainty varies across profiles and student classes. At the enrolment level, 320 enrolments are assigned to the average Profile 2, with an average posterior probability

of 0.826. By comparison, the 441 enrolments assigned to the disengaged Profile 1 and the 549 assigned to the active Profile 3 have average posterior probabilities of 0.953 and 0.938. The average profile lies between the other two, so enrolments classified into it are, on average, more ambiguous.

A similar pattern appears at the student level. Students assigned to the fluctuating Class 2 and the engaged Class 3 have average posterior probabilities of 0.920 and 0.921, whereas those assigned to the mostly disengaged Class 1 have an average posterior probability of 0.950. The student-level classifications are therefore not equally certain across classes, even though the overall relative entropy is similar across the two levels.

These differences matter for subsequent analyses because modal class assignments treat the most likely class as if it were observed without error. When average posterior probabilities are lower, that simplification discards more of the uncertainty contained in the posterior distribution. Later analyses should therefore retain or adjust for classification uncertainty.

The covariate vignette describes approaches for incorporating this uncertainty when relating latent profiles and classes to external variables, including methods that adjust explicitly for classification error.

## Residual dependence

An important assumption of the diagonal covariance model is that the indicators are conditionally independent within each latent profile. Bivariate residual correlations can be used to examine whether meaningful associations remain between pairs of indicators after the fitted profile structure has been taken into account. For indicators $a$ and $b$ within profile $k$, the residual correlation is defined as

$$r_{ab,k}^{\text{observed}} - r_{ab,k}^{\text{expected}}.$$

The observed correlation is calculated using posterior probabilities of profile membership as weights. Under the diagonal covariance model, the expected within-profile correlation is zero. A residual close to zero therefore indicates that the conditional-independence assumption provides a reasonable approximation for that particular indicator pair. Larger residuals indicate associations that remain unexplained by the fitted measurement model, while the sign of the residual indicates whether the remaining association is positive or negative.

These residual correlations are diagnostic quantities. Their values depend on the estimated posterior profile memberships and therefore on the fitted model itself. In addition, the reported p-values are descriptive approximations. They do not account fully for uncertainty in the estimated measurement parameters or posterior memberships, nor do they adjust for the dependence created by repeated enrolments from the same student. The magnitude and pattern of the residuals are therefore more useful for identifying measurement assumptions that warrant further examination than the p-values are for making formal inferential decisions.

The command `get_results(fit, "residuals")` returns the indicator pairs ordered by the absolute magnitude of their residual correlations. Examining the largest residuals provides a convenient way to identify the pairs for which the diagonal covariance assumption is least consistent with the observed associations.

```
get_results(fit, "residuals") |> head(4)
#>    profile indicator_1 indicator_2    kind  observed expected  residual effective_n statistic
#> 1 profile_1   lectures active_days gaussian  0.1150778        0  0.1150778    442.481  2.42320
#> 2 profile_2  forum_read active_days gaussian -0.0948800        0 -0.0948800    319.523 -1.69311
#> 3 profile_3  forum_post active_days gaussian -0.0887335        0 -0.0887335    547.996 -2.07696
#> 4 profile_2   lectures  forum_read gaussian  0.0794574        0  0.0794574    319.523  1.41662
#>   df   p_value p_adjusted
#> 1 NA 0.0153843   0.461530
#> 2 NA 0.0904341   0.678256
#> 3 NA 0.0378054   0.567081
#> 4 NA 0.1565941   0.786502
```

The largest residual correlation is 0.115, between lecture views and days active in the disengaged Profile 1, and none of the residuals remains significant after the Benjamini-Hochberg adjustment (the smallest adjusted p-value is 0.46). Within each profile, the indicators are therefore approximately uncorrelated, as the diagonal covariance structure assumes. This is expected for these synthetic data, which were generated with indicators that are conditionally independent within each engagement state; in applied data, a concentration of large residuals on particular indicator pairs would identify the part of the measurement structure where the assumption is too restrictive.

## Class counts and covariance assumptions

Class enumeration evaluates alternative models with different numbers of enrolment-level profiles and student-level latent classes. One common criterion for comparing these models is the Bayesian information criterion (BIC),

$$\mathrm{BIC} = -2\ell + q\log(N),$$

where $\ell$ is the maximized log likelihood, $q$ is the number of freely estimated parameters, and $N$ is the sample size used in the penalty term. Lower BIC values indicate a more favorable balance between model fit and complexity within the set of candidate models being compared. In the present analysis, `bic_groups` uses the number of students as $N$ for every candidate, so the BIC values are calculated on a consistent basis across models.

The comparison also examines whether assumptions about the covariance structure affect the preferred numbers of profiles and student classes. With `model = "VVI"`, each profile has its own indicator variances, but residual covariances between indicators are fixed to zero. This specification corresponds to the conditional-independence assumption examined in the preceding section. By contrast, `model = "VVV"` estimates a separate unrestricted covariance matrix for each profile, so the indicators can remain correlated within profiles.

Both covariance structures retain Gaussian enrolment-level profiles and use common measurement parameters across student classes. The main difference is therefore the amount of within-profile dependence that the measurement model is permitted to capture. Because unrestricted covariance matrices require substantially more parameters, full-covariance models can be more difficult to estimate. Numerical instability is particularly likely when some profiles contain relatively few observations or when several likelihood maxima provide similar fits.

The `enumerate_classes()` function fits every requested combination of profile count, student-class count, and covariance structure to the same enrolments and indicators. The resulting candidate models can then be summarized with `get_results()`, which reports BIC together with numerical diagnostics such as convergence status, active parameter bounds, and replication of the best likelihood across starting values.

These diagnostics should be considered before interpreting differences in BIC. Comparisons are most informative when the candidate models have converged, no relevant parameter bounds are active, and the best likelihood is reproduced across multiple starts. BIC then provides evidence about relative fit within the specified candidate set, but it does not by itself establish that the Gaussian measurement assumptions are correct or that a unique number of latent profiles or classes exists.

```
candidates <- enumerate_classes(engagement_momentum, vars, id = "student",
                                n_profiles = 2:4, n_group_classes = 2:4,
                                model = c("VVI", "VVV"), n_starts = 6,
                                max_iter = 2000, tol = 1e-10, seed = 1)
#> 3 of 18 candidates reached their best log likelihood from one start only (profiles4_groups4_VVI, profiles4_g
get_results(candidates, "candidates")
#>              candidate n_profiles n_group_classes model log_likelihood n_parameters    aic
#> 1  profiles2_groups2_VVI         2              2   VVI      -7665.30          23 15376.6
```

```
#> 2  profiles3_groups2_VVI         3              2   VVI      -7526.39          35 15122.8
#> 3  profiles4_groups2_VVI         4              2   VVI      -7510.81          47 15115.6
#> 4  profiles2_groups3_VVI         2              3   VVI      -7636.81          25 15323.6
#> 5  profiles3_groups3_VVI         3              3   VVI      -7487.95          38 15051.9
#> 6  profiles4_groups3_VVI         4              3   VVI      -7470.41          51 15042.8
#> 7  profiles2_groups4_VVI         2              4   VVI      -7636.15          27 15326.3
#> 8  profiles3_groups4_VVI         3              4   VVI      -7486.77          41 15055.5
#> 9  profiles4_groups4_VVI         4              4   VVI      -7469.29          55 15048.6
#> 10 profiles2_groups2_VVV         2              2   VVV      -7551.30          43 15188.6
#> 11 profiles3_groups2_VVV         3              2   VVV      -7504.91          65 15139.8
#> 12 profiles4_groups2_VVV         4              2   VVV      -7476.88          87 15127.8
#> 13 profiles2_groups3_VVV         2              3   VVV      -7526.13          45 15142.3
#> 14 profiles3_groups3_VVV         3              3   VVV      -7465.45          68 15066.9
#> 15 profiles4_groups3_VVV         4              3   VVV      -7430.05          91 15042.1
#> 16 profiles2_groups4_VVV         2              4   VVV      -7525.72          47 15145.4
#> 17 profiles3_groups4_VVV         3              4   VVV      -7463.97          71 15069.9
#> 18 profiles4_groups4_VVV         4              4   VVV      -7427.89          95 15045.8
#>      kic bic_groups bic_individual sabic_groups sabic_individual caic_groups caic_individual
#> 1  15402.6    15437.9       15495.7     15365.2         15422.6     15460.9         15518.7
#> 2  15160.8    15216.0       15304.0     15105.4         15192.8     15251.0         15339.0
#> 3  15165.6    15240.8       15359.0     15092.3         15209.7     15287.8         15406.0
#> 4  15351.6    15390.2       15453.1     15311.2         15373.6     15415.2         15478.1
#> 5  15092.9    15153.1       15248.7     15033.1         15128.0     15191.1         15286.7
#> 6  15096.8    15178.6       15306.9     15017.5         15144.9     15229.6         15357.9
#> 7  15356.3    15398.2       15466.1     15312.9         15380.3     15425.2         15493.1
#> 8  15099.5    15164.7       15267.8     15035.2         15137.6     15205.7         15308.8
#> 9  15106.6    15195.1       15333.4     15021.3         15158.7     15250.1         15388.4
#> 10 15234.6    15303.1       15411.2     15167.3         15274.7     15346.1         15454.2
#> 11 15207.8    15312.9       15476.4     15107.6         15269.9     15377.9         15541.4
#> 12 15217.8    15359.5       15578.2     15084.6         15301.9     15446.5         15665.2
#> 13 15190.3    15262.1       15375.3     15119.9         15232.3     15307.1         15420.3
#> 14 15137.9    15248.0       15419.0     15033.2         15203.0     15316.0         15487.0
#> 15 15136.1    15284.5       15513.3     14997.0         15224.2     15375.5         15604.3
#> 16 15195.4    15270.6       15388.8     15122.1         15239.5     15317.6         15435.8
#> 17 15143.9    15259.1       15437.6     15034.7         15212.0     15330.1         15508.6
#> 18 15143.8    15298.8       15537.7     14998.7         15235.9     15393.8         15632.7
#>   awe_groups awe_individual icl_groups icl_individual clc_groups clc_individual profile_entropy
#> 1    15775.2       15890.9   15598.9       15656.8   15491.7       15491.7       0.913912
#> 2    16035.7       16211.7   15767.5       15855.5   15604.3       15604.3       0.809795
#> 3    16158.2       16394.5   15798.0       15916.2   15578.8       15578.8       0.847727
#> 4    15774.6       15900.3   15583.0       15645.8   15466.4       15466.4       0.913354
#> 5    16011.0       16202.1   15719.8       15815.4   15542.6       15542.6       0.812713
#> 6    16125.8       16382.3   15735.0       15863.2   15497.2       15497.2       0.853074
#> 7    15849.5       15985.3   15642.6       15710.5   15516.7       15516.7       0.913778
#> 8    16094.2       16300.3   15780.0       15883.1   15588.8       15588.8       0.813111
#> 9    16230.2       16506.8   15808.8       15947.0   15552.3       15552.3       0.852620
#> 10   15871.8       16088.0   15542.3       15650.4   15341.8       15341.8       0.868829
#> 11   16257.2       16584.0   15759.0       15922.5   15455.9       15455.9       0.846473
#> 12   16501.8       16939.3   15835.1       16053.9   15429.4       15429.4       0.870262
#> 13   15882.4       16108.7   15537.6       15650.7   15327.7       15327.7       0.865778
#> 14   16248.9       16590.8   15727.7       15898.7   15410.6       15410.6       0.843319
#> 15   16506.5       16964.1   15809.1       16037.9   15384.7       15384.7       0.863312
#> 16   15934.3       16170.6   15574.1       15692.3   15354.9       15354.9       0.865181
```

```
#> 17   16331.5       16688.5   15787.4       15965.9   15456.3       15456.3       0.843804
#> 18   16550.9       17028.6   15822.8       16061.7   15379.8       15379.8       0.865895
#>    group_entropy avepp_min avepp_max avepp_mean avepp_below group_avepp_min group_avepp_max
#> 1       0.929549  0.974914  0.977550   0.976232           0        0.980823        0.983280
#> 2       0.953063  0.832406  0.947038   0.904811           0        0.988011        0.992342
#> 3       0.953455  0.713916  0.954650   0.857902           1        0.982183        0.993283
#> 4       0.811681  0.971376  0.978428   0.974902           0        0.895011        0.954745
#> 5       0.836203  0.826204  0.952706   0.905484           0        0.919908        0.950178
#> 6       0.851568  0.768376  0.957413   0.870250           1        0.917883        0.965488
#> 7       0.671184  0.972491  0.978134   0.975313           0        0.634443        0.946763
#> 8       0.692482  0.827418  0.952919   0.905669           0        0.668375        0.978782
#> 9       0.687555  0.777115  0.956299   0.872535           1        0.653907        0.981753
#> 10      0.948991  0.951652  0.972085   0.961868           0        0.987030        0.990114
#> 11      0.953519  0.818670  0.963116   0.912100           0        0.986899        0.994537
#> 12      0.949996  0.785228  0.962043   0.878374           1        0.978128        0.990688
#> 13      0.800218  0.956026  0.965483   0.960754           0        0.869398        0.952762
#> 14      0.836164  0.815401  0.962345   0.910886           0        0.901696        0.970395
#> 15      0.844613  0.734570  0.963552   0.865876           1        0.910252        0.949602
#> 16      0.739331  0.953928  0.966512   0.960220           0        0.532182        0.943460
#> 17      0.695484  0.817464  0.962743   0.911954           0        0.684357        0.973265
#> 18      0.837372  0.744678  0.962562   0.869227           1        0.671909        0.929821
#>    group_avepp_mean group_avepp_below converged boundary min_standardized_eigenvalue
#> 1          0.982052                0      TRUE    FALSE                   0.37577508
#> 2          0.990177                0      TRUE    FALSE                   0.29937168
#> 3          0.987733                0      TRUE    FALSE                   0.03232118
#> 4          0.918358                0      TRUE    FALSE                   0.38455785
#> 5          0.930254                0      TRUE    FALSE                   0.30047371
#> 6          0.938557                0      TRUE    FALSE                   0.03470778
#> 7          0.814465                1      TRUE    FALSE                   0.38512770
#> 8          0.817482                1      TRUE    FALSE                   0.30019881
#> 9          0.819033                2      TRUE    FALSE                   0.03445162
#> 10         0.988572                0      TRUE    FALSE                   0.31074473
#> 11         0.990718                0      TRUE    FALSE                   0.22781144
#> 12         0.984408                0      TRUE    FALSE                   0.00241103
#> 13         0.913903                0      TRUE    FALSE                   0.32559225
#> 14         0.928998                0      TRUE    FALSE                   0.24422535
#> 15         0.931381                0      TRUE    FALSE                   0.02947099
#> 16         0.774297                2      TRUE    FALSE                   0.32629434
#> 17         0.827306                1      TRUE    FALSE                   0.24685553
#> 18         0.858839                1      TRUE    FALSE                   0.02483378
#>    min_gaussian_mass n_best_replicated warnings error
#> 1           561.8366                 6          <NA>
#> 2           320.7660                 6          <NA>
#> 3            28.9257                 5          <NA>
#> 4           567.7480                 6          <NA>
#> 5           319.5231                 6          <NA>
#> 6            23.7104                 2          <NA>
#> 7           568.5995                 5          <NA>
#> 8           318.7681                 5          <NA>
#> 9            25.7986                 1          <NA>
#> 10          518.7588                 6          <NA>
#> 11          253.2844                 6          <NA>
#> 12           24.5006                 1          <NA>
```

```
#> 13          541.9110                6          <NA>
#> 14          258.8424                6          <NA>
#> 15           68.9344                4          <NA>
#> 16          542.7478                3          <NA>
#> 17          261.1041                5          <NA>
#> 18           62.1293                1          <NA>
```

Under the diagonal covariance specification, the lowest BIC is obtained for the model with three profiles and three student classes, with a BIC of 15153.1, an interior minimum: removing a profile or a class (15390.2 for two profiles, 15216.0 for two classes) or adding one (15178.6 for four profiles, 15164.7 for four classes) raises the BIC. No full-covariance model reaches the BIC of the best diagonal model, although with two profiles the full-covariance models have the lower BIC of the two structures; the best of them, also with three profiles and three classes, has a BIC of 15248.0, 94.9 points above its diagonal counterpart, because its 30 additional covariance parameters do not improve the likelihood enough to justify them.

Most candidate models also show stable likelihood replication across starts. The exceptions are the larger candidates: the four-profile, four-class models with either covariance structure and the four-profile, two-class full-covariance model reach their best solution from only one of the six initializations. This weaker replication suggests greater numerical uncertainty for those candidates and illustrates why class enumeration should not be based on information criteria alone.

Taken together, these results support the three-profile, three-class specification with a diagonal covariance structure. Because the residual dependence examined above is negligible, freeing the within-profile covariances does not change the preferred numbers of profiles and classes; it only adds parameters. In applied data with substantial residual dependence, the same comparison can favour a different profile count once the covariances are estimated, so the two analyses should always be read together.

The `get_candidate()` function can be used to retrieve a fitted candidate with a particular combination of profile count, student-class count, and covariance structure. Examining its residual correlations provides a direct check of whether the estimated covariance parameters reproduce the within-profile associations observed in the data.

```
full <- get_candidate(candidates, n_profiles = 3, n_group_classes = 3,
                      model = "VVV")
get_results(full, "covariances") |>
  subset(indicator != indicator_2) |>
  head(6)
#>   profile   indicator indicator_2  covariance
#> 2       1    lectures      browse  0.02906203
#> 3       1  forum_read      browse -0.00191435
#> 4       1  forum_post      browse  0.02099442
#> 5       1 active_days      browse  0.03165217
#> 6       1      browse    lectures  0.02906203
#> 8       1  forum_read    lectures  0.00334423
```

In the three-profile, three-class full-covariance model, the covariances printed for the first profile are close to zero (-0.002 to 0.032). The full-covariance model therefore estimates the associations that the diagonal model fixes at zero and finds them negligible, which is why its additional parameters are penalized without a compensating gain in fit.

This agreement between the two specifications supports the diagonal measurement model for these data. It should not, however, be interpreted as independent evidence that the three profiles or the three student classes are substantively correct. Those questions require consideration of the broader pattern of model fit, numerical stability, classification quality, and substantive interpretability.

## Uncertainty in profile means

Parameter inference quantifies the sampling uncertainty associated with an estimated model parameter, $\hat{\theta}$. By default, the standard error is obtained from the observed information matrix. Specifically, the estimated variance of a parameter is taken from the corresponding diagonal element of the inverse observed information matrix, and the standard error is the square root of that value.

A conventional 95% Wald confidence interval is calculated as

$$\hat{\theta} \pm 1.96 \ \mathrm{SE}(\hat{\theta}).$$

These intervals provide a local approximation to parameter uncertainty within the fitted model. Their use requires a satisfactorily converged solution, inactive parameter bounds, and an invertible information matrix. If these conditions are not met, the resulting standard errors and confidence intervals may be unstable or unavailable.

The intervals are also conditional on the specified latent structure. In particular, they treat the selected numbers of enrolment-level profiles and student-level classes, together with the assumed covariance structure, as fixed. A parameter estimate can therefore have a narrow confidence interval even when an alternative model specification supports a different number of profiles. Wald intervals describe uncertainty about parameters within a chosen model; they do not incorporate uncertainty about the choice of model itself.

The `parameter_inference()` function calculates standard errors and confidence intervals for every parameter of the three-profile, three-class model. The discussion below uses the rows for the means of days active.

```
parameter_inference(fit)
#>          level      outcome          term   parameter   estimate standard_error  statistic
#> 1  measurement    profile_1        browse        mean -0.9275396      0.0322954 -28.720497
#> 2  measurement    profile_1      lectures        mean -0.7940260      0.0331091 -23.982135
#> 3  measurement    profile_1    forum_read        mean -0.9802813      0.0289440 -33.868156
#> 4  measurement    profile_1    forum_post        mean -0.8725968      0.0327350 -26.656393
#> 5  measurement    profile_1   active_days        mean -0.9607501      0.0332744 -28.873537
#> 6  measurement    profile_2        browse        mean -0.0121106      0.0538474  -0.224906
#> 7  measurement    profile_2      lectures        mean -0.1422332      0.0555643  -2.559796
#> 8  measurement    profile_2    forum_read        mean -0.0789085      0.0566160  -1.393749
#> 9  measurement    profile_2    forum_post        mean -0.0549890      0.0520043  -1.057394
#> 10 measurement    profile_2   active_days        mean -0.0294016      0.0505661  -0.581449
#> 11 measurement    profile_3        browse        mean  0.7558060      0.0346716  21.798971
#> 12 measurement    profile_3      lectures        mean  0.7239442      0.0384845  18.811299
#> 13 measurement    profile_3    forum_read        mean  0.8373042      0.0302118  27.714454
#> 14 measurement    profile_3    forum_post        mean  0.7367903      0.0377386  19.523533
#> 15 measurement    profile_3   active_days        mean  0.7926856      0.0336281  23.572118
#> 16 measurement    profile_1        browse    variance  0.3657645      0.0268615         NA
#> 17 measurement    profile_1      lectures    variance  0.4130822      0.0292085         NA
#> 18 measurement    profile_1    forum_read    variance  0.2931187      0.0223617         NA
#> 19 measurement    profile_1    forum_post    variance  0.3825361      0.0281862         NA
#> 20 measurement    profile_1   active_days    variance  0.3411264      0.0287758         NA
#> 21 measurement    profile_2        browse    variance  0.4624725      0.0472108         NA
#> 22 measurement    profile_2      lectures    variance  0.5500268      0.0525076         NA
#> 23 measurement    profile_2    forum_read    variance  0.4350730      0.0449940         NA
#> 24 measurement    profile_2    forum_post    variance  0.4388228      0.0437801         NA
#> 25 measurement    profile_2   active_days    variance  0.4056312      0.0446779         NA
#> 26 measurement    profile_3        browse    variance  0.5007539      0.0331284         NA
#> 27 measurement    profile_3      lectures    variance  0.6332103      0.0418390         NA
#> 28 measurement    profile_3    forum_read    variance  0.3610212      0.0244030         NA
```

```
#> 29 measurement     profile_3    forum_post    variance  0.6075119      0.0396958         NA
#> 30 measurement     profile_3   active_days    variance  0.4459738      0.0308463         NA
#> 31     profile     profile_1 group_class_1 probability  0.8596283      0.0357035         NA
#> 32     profile     profile_2 group_class_1 probability  0.1221521      0.0326782         NA
#> 33     profile     profile_3 group_class_1 probability  0.0182196      0.0127390         NA
#> 34     profile     profile_1 group_class_2 probability  0.3232298      0.0326865         NA
#> 35     profile     profile_2 group_class_2 probability  0.3495716      0.0277377         NA
#> 36     profile     profile_3 group_class_2 probability  0.3271986      0.0355504         NA
#> 37     profile     profile_1 group_class_3 probability  0.0660150      0.0165022         NA
#> 38     profile     profile_2 group_class_3 probability  0.1826631      0.0304012         NA
#> 39     profile     profile_3 group_class_3 probability  0.7513220      0.0362442         NA
#> 40       group group_class_1         <NA> probability  0.1965230      0.0425463         NA
#> 41       group group_class_2         <NA> probability  0.4461911      0.0617197         NA
#> 42       group group_class_3         <NA> probability  0.3572859      0.0616873         NA
#>         p_value   p_adjusted    conf_low  conf_high
#> 1  2.11660e-181 2.11660e-181 -0.99083743 -0.8642418
#> 2  4.27192e-127 4.27192e-127 -0.85891855 -0.7291334
#> 3  1.96155e-251 1.96155e-251 -1.03701060 -0.9235520
#> 4  1.50893e-156 1.50893e-156 -0.93675623 -0.8084374
#> 5  2.56659e-183 2.56659e-183 -1.02596674 -0.8955334
#> 6   8.22052e-01  8.22052e-01 -0.11764957  0.0934284
#> 7   1.04734e-02  1.04734e-02 -0.25113725 -0.0333292
#> 8   1.63393e-01  1.63393e-01 -0.18987393  0.0320569
#> 9   2.90332e-01  2.90332e-01 -0.15691555  0.0469375
#> 10  5.60938e-01  5.60938e-01 -0.12850931  0.0697061
#> 11 2.37291e-105 2.37291e-105  0.68785081  0.8237611
#> 12  6.10239e-79  6.10239e-79  0.64851592  0.7993725
#> 13 4.67554e-169 4.67554e-169  0.77809011  0.8965183
#> 14  6.92782e-85  6.92782e-85  0.66282403  0.8107565
#> 15 7.44804e-123 7.44804e-123  0.72677572  0.8585955
#> 16           NA           NA  0.31673058  0.4223895
#> 17           NA           NA  0.35962449  0.4744864
#> 18           NA           NA  0.25240987  0.3403931
#> 19           NA           NA  0.33109585  0.4419683
#> 20           NA           NA  0.28914261  0.4024562
#> 21           NA           NA  0.37861016  0.5649103
#> 22           NA           NA  0.45616820  0.6631973
#> 23           NA           NA  0.35524928  0.5328329
#> 24           NA           NA  0.36088363  0.5335944
#> 25           NA           NA  0.32687110  0.5033686
#> 26           NA           NA  0.43985686  0.5700820
#> 27           NA           NA  0.55629530  0.7207599
#> 28           NA           NA  0.31622513  0.4121631
#> 29           NA           NA  0.53448553  0.6905157
#> 30           NA           NA  0.38943501  0.5107209
#> 31           NA           NA  0.77421829  0.9162249
#> 32           NA           NA  0.07112754  0.2018268
#> 33           NA           NA  0.00457445  0.0697165
#> 34           NA           NA  0.26272920  0.3902871
#> 35           NA           NA  0.29733428  0.4056878
#> 36           NA           NA  0.26165243  0.4002633
#> 37           NA           NA  0.04015004  0.1066902
#> 38           NA           NA  0.13039029  0.2498704
```

```
#> 39           NA           NA  0.67380948  0.8154598
#> 40           NA           NA  0.12605742  0.2931651
#> 41           NA           NA  0.33056845  0.5679447
#> 42           NA           NA  0.24718582  0.4848423
```

For the active Profile 3, the estimated mean of days active is 0.793 standard deviations above the course mean, with a standard error of 0.034 and a 95% Wald interval from 0.727 to 0.859. The average Profile 2 has an interval from −0.129 to 0.070, which includes the course mean, while the disengaged Profile 1 has an interval from −1.026 to −0.896.

Within the three-profile, three-class specification, these intervals show clear separation among the means of days active. This provides evidence that the fitted profiles differ substantially in their expected levels of activity under that model. However, the precision of these estimates should not be interpreted as resolving model-specification uncertainty. The confidence intervals quantify uncertainty in the profile means conditional on the three-by-three model; they do not address whether a different covariance structure or profile count provides a preferable representation of the data.

## Limits of the interpretation

The model comparisons presented here should be interpreted within the limits of the candidate set that was examined. Only a restricted range of profile counts, student-class counts, and covariance structures was considered, and the resulting BIC values also depend on the convention used to define the sample size in the BIC penalty. The comparison therefore identifies relative support among the fitted candidates only.

Additional caution is warranted for candidate models whose best likelihood was not reproduced across multiple starting values. In particular, full-covariance solutions with weak replication should be refitted with additional starts before they are used for substantive interpretation. More generally, all reported classification summaries, residual diagnostics, standard errors, and confidence intervals are conditional on the measurement specification from which they were obtained.

The multilevel structure accounts for the nesting of repeated enrolments within students by allowing students to differ in the composition of their engagement profiles. However, the model does not directly estimate serial dependence between adjacent enrolments. The argument `time = "sequence"` preserves the ordering of each student's enrolments and makes that ordering available for sequence-based summaries, but it does not by itself introduce a transition or autoregressive structure into the fitted model.

When the scientific question concerns how students move from one engagement profile to another over time, a transition model is more appropriate. The transition vignette extends the analysis by estimating probabilities of movement between profiles across successive enrolments.

## References

López-Pernas, S., & Saqr, M. (2024). How the dynamics of engagement explain the momentum of achievement and the inertia of disengagement: A complex systems theory approach. *Computers in Human Behavior, 153*, 108126. doi:10.1016/j.chb.2023.108126.

Saqr, M., & López-Pernas, S. (2021). The longitudinal trajectories of online engagement over a full program. *Computers & Education, 175*, 104325. doi:10.1016/j.compedu.2021.104325.