# Supplementary Material (SM)

## Contents

**SM 1. Detailed Analysis Steps**

As an approach to segmentation, latent class analysis has been used to explore variations in patients' use of complementary medicine (Strizich et al., 2015), how attitudes toward mental health are formed (Mannarini, Boffo, Rossi, & Balottin, 2018), and stakeholder expectations toward Corporate Responsibility (Hillenbrand & Money, 2009). These latent models can also include covariates, which allow the prior probabilities of latent class assignment to vary for each respondent (Linzer & Lewis, 2011). For example, Strizich and colleagues found higher use of complementary medicines to be associated with high levels of exercise and healthier eating habits (Strizich et al., 2015). Following the approach adopted by these aforementioned studies, the current case study applied latent class analysis in an exploratory approach to gauge and segment student expectations of learning analytics services, addressing RQ1 and RQ2. Covariates were also included in the latent class model in order to gain a greater understanding of what characteristics typically define the groups identified, which answered RQ3. For RQ4, a contingency table was created to explore whether student class assignment was stable or variable across the two expectation scale (ideal and predicted).

To address research questions one (RQ1) and two (RQ2), the raw data was analysed using the three-step approach to latent class analysis (Vermunt, 2010), which was carried out in Mplus 8.1 (Muthén & Muthén, 2017). The traditional one-step method was not used as various disadvantages of this approach have been outlined (Vermunt, 2010). An example of how the one step method is disadvantageous is in relation to the number of classes to extract, as the solution will change with the inclusion or exclusion of covariates (Vermunt, 2010). To overcome these issues, Vermunt (2010) presented the three-step method to latent class analysis. This is a step-wise approach in which the latent class model is first estimated with indicator variables alone, then a most likely class variable is generated, which is then regressed onto the predictor variables (Asparouhov & Muthén, 2014; Vermunt, 2010). Thus, the three-step method does not change the initial measurement model through the introduction of covariates, as is the case with the one-step approach (Vermunt, 2010).

For the analysis of the collected data, the ideal and predicted expectation scales were analysed separately. An assessment of the response distributions for each scale shows the data to contain ceiling effects (SM 4 and 5), particularly with regards to the ideal expectation scale. This is anticipated as the ideal expectation scale corresponds to a desired level of service so responses on this scale are likely to be high. Therefore, the data collected from the SELAQ was treated as categorical. As for the model covariates, the age variable was treated as continuous; whereas, the remaining variables were dummy coded. These dummy coded variables were gender (0 = male, 1= female), management, science, and technology (0 = culture and jurisprudence, 1 = management, science, and technology), psychology and education (0 = culture and jurisprudence, 1 = psychology and education), Postgraduate Student (0 = Undergraduate Student, 1 = Postgraduate Student), European Student (0 = Dutch Student, 1 = European Student), and Overseas Student (0 = Dutch Student, 1 = Overseas Student). These covariates allowed for the exploration of whether gender, age, faculty, level of study, or student type were associated with latent class assignment.

As for the latent class model building, the steps outlined by Masyn (2013) will be followed, which can be decomposed into assessments of absolute fit, relative fit, classification diagnostics, and class interpretation. When assessing absolute fit, the absolute values of standardised residuals will examined. According to Masyn (2013), values exceeding 3 are indicative of poor fitting response frequencies. Given the large number of response frequencies that are possible due to both the number of latent class indicators (n = 12 per expectation scale) and response options (n = 7), it is difficult to determine what constitutes a poor fitting model. A useful guideline was proposed by Masyn (2013), which states that large standardised residual values in “notable excess” of 5% would lead to a model being considered as poor fitting (p. 567).

With regards to the relative fit of each model, this examined using both an inferential and information-heuristic approach (Masyn, 2013). In terms of the inferential approach, there are two tests used which compare a *K* class model to a *K* – 1 class model (e.g., compare a 3 class model to a 2 class model), which are the adjusted Lo-Mendell-Rubin likelihood ratio test (LMR-LRT; (Lo, Mendell, & Rubin, 2001) and the bootstrap likelihood ratio test (BLRT; McLachlan & Peel, 2000). In the case of either test, if the likelihood ratio difference is found to be statistically significant then the model containing a greater number of classes is considered to fit better (Masyn, 2013). As for the information heuristic approach, the Bayesian Information Criterion (BIC; Schwarz, 1978) is most commonly used to determine the best fitting model (Nylund, Asparouhov, & Muthén, 2007). This decision is usually based on the number of classes where the BIC value is lowest (Nylund et al., 2007) or from “elbow” plots (Masyn, 2013). There are other indexes that can be used such as Akaike’s Information Criterion (AIC; Akaike, 1987); however, it has been shown that the BIC is the best information criterion (Nylund et al., 2007). Therefore, only the BIC of each model will be plotted and decisions regarding model selection will be based on the “elbow criterion” (Masyn, 2013). If, in conjunction with the findings of the inferential approach, there is no clear contender for a model (e.g., no *K* + 1 model is rejected) then a plot of log likelihood values will also be examined (Masyn, 2013). As with the BIC value plot, an “elbow” in the plot of log likelihood values can also be used to identify a candidate model (Masyn, 2013).

For assessing the classification precision, the relative entropy will be one of the diagnostic statistics used (Ramaswamy, Desarbo, Reibstein, & Robinson, 1993). It is intended to provide a summary of classification accuracy across each latent class, with values lying between 0 (classification no better than chance) and 1 (classification is perfect)

(Ramaswamy et al., 1993). As a means to selecting the number of classes to extract, the relative entropy should not be used as even with high values there is likely to be assignment error (Masyn, 2013). Therefore, three additional classification diagnostic statistics will be examined: the average posterior class probability (AvePP), the odds of correct classification ratio (OCC), and the modal class assignment proportion (Masyn, 2013). The AvePP provides a class-specific measure of assignment accuracy between 0 and 1, with values greater than .70 being suggestive of good accuracy (Nagin, 2005). The OCC is also used to assess both assignment accuracy and class separation, with values exceeding 5 being good (Nagin, 2005). Finally, the mcaP is the proportion of those individuals modally assigned to a specific class and this is compared to the model-estimated proportions of this class ($\hat{\pi}_k$) (Masyn, 2013). The size of the discrepancies between the mcaP and $\hat{\pi}_k$ provides an indication of whether there are errors in the class assignment, specifically when the discrepancy size is large (Masyn, 2013).

Throughout these abovementioned steps, it is necessary that the interpretability of the solution needs to be considered (Lanza & Rhoades, 2013). For instance, there may be problems regarding the local fit of the model (e.g., proportion of standardised residuals greater than 5%), which can be addressed by increasing the number of classes that are extracted. However, this additional class may not be easily interpreted; thus, based on parsimony, the *K*-1 model would be more suitable. For Lanza and Rhoades (2013), they recommend that class interpretability should be guided by a clear separation between classes, classes being easily labelled, and patterns that are logical. To assist in decisions regarding the interpretability of a solution, we will follow the step taken by Oberski (2016) and use profile plots. These plots provide the estimated class means as opposed to the estimated distributions (Oberski, 2016). This is because there are seven possible categories (1 = Strongly Disagree, 7 = Strongly Agree), which makes plots of estimated distributions difficult to read (Oberski, 2016).

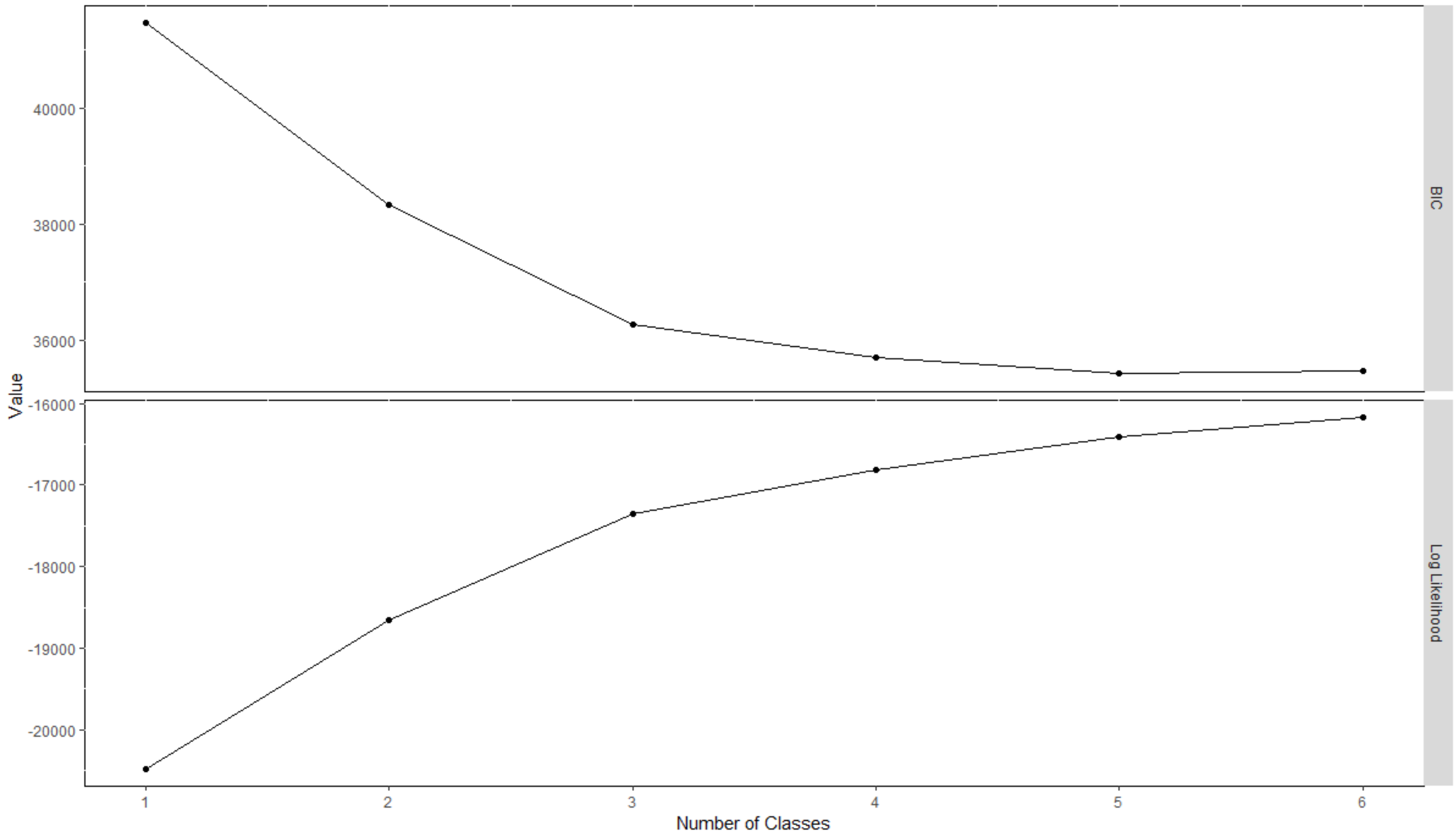
Thus, to provide an overview of the steps taken in this analysis, we increased the number of classes to extract until either the solution could not be identified or the number of classes would affect the interpretability of the solution. These models would then be compared on the basis of their relative fit using both the inferential and information-heuristic approaches. From this, a selection of possible models will be selected and then compared on the basis of their classification accuracy and local fit. Throughout each stage, decisions regarding the selection of a candidate model will also be determined by the class interpretability. Once a suitable candidate model has been identified, the latent class regression is then ran, which addresses research question three (RQ3). For the purpose of this paper, the alpha level is set at 5% for determining whether an effect is considered to be statistically significant.

**SM 2. Detailed Results for the Ideal Expectation Scale**

One to six latent class models were estimated from the data. Based on the BIC values obtained from these six models, the three class model appeared to meet the “elbow criterion” as the addition of more classes did not provide more information (SM Figure 1). It was also found that at the six class solution, the BIC value began to increase. Thus, on the BIC values alone the final model would be a three class solution.

In order to further test the suitability of this three class solution, the relative fit of this model over a two class solution was assessed using the adjusted LMR-LRT and BLRT. The results obtained from these relative fit tests did not provide clear evidence to support a three class solution over a two class solution as the adjusted LMR-LRT was not statistically significant (LMR-LRT = 2584.362, $p = .763$), but the BLRT was statistically significant (BLRT = 2589.332, $p < .001$). In contrast, both the LMR-LRT and BLRT were statistically significant (LMR-LRT = 3647.126, $p < .001$; BLRT = 3654.238, $p < .001$) for the comparison of a two class solution against a one class solution.

Given the discrepancies between these two evaluations of relative fit for the three class solution, it is important to also consider a plot of log likelihood values (SM Figure 1). As with the plot of BIC values, there was a clear “elbow” for the three class solution. Thus, the evidence seemingly supported the three class solution as a candidate model. However, given the non-significant LMR-LRT it was important to compare the classification diagnostics between the two and three class solutions.




**SM Figure 1.** Index Values across Six Latent Class Models

To assess the classification accuracy of the two and three class solutions, the relative entropy of both models were initially compared. For the two class solution, the entropy value was .931, which was greater than the value of .919 for the three class solution. In both cases, the relative entropy values showed either solution (k = 2 and k = 3) to have good classification precision, but it should not be used to justify the selection of a candidate model. For the purpose of selecting a candidate model on the basis of classification diagnostics, the AvePP, OCC, and mcaP were used (SM Tables 1 and 2).

SM Table 1 shows that for the two class solution, the discrepancies between model estimated proportions for each class ($\hat{\pi}_k$) and modal class assignment proportions ($mcaP_k$) were not large (absolute difference of .004 for both class one and two). All AvePP values exceeded .70 (class one = .984; class two = .974) and both OCC values were larger than 5 (24.755 and 93.066 for class one and two, respectively).

**SM Table 1.** Two Class Classification Accuracy Diagnostics

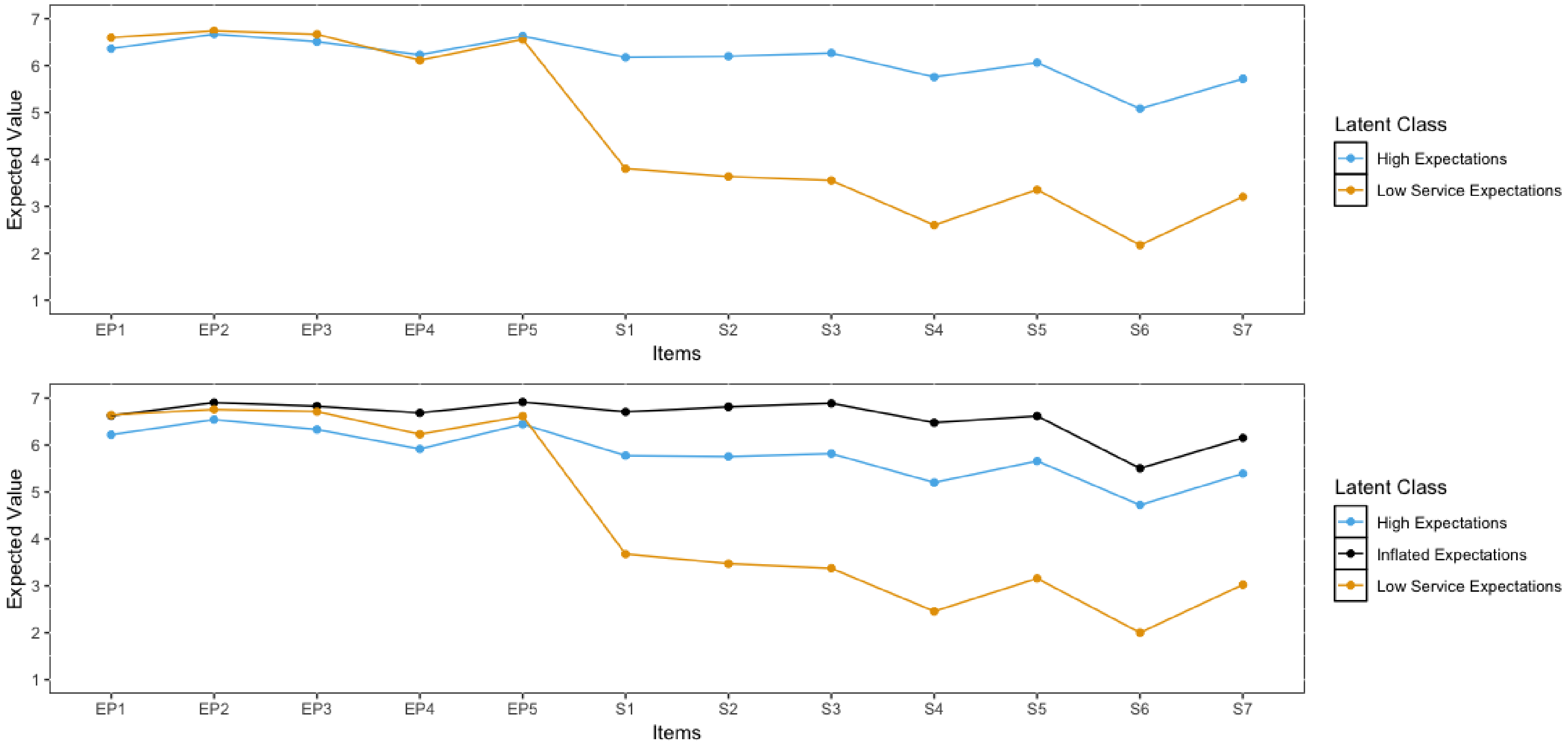
| Class *k* | $\hat{\pi}_k$ | $mcaP_k$ | $AvePP_k$ | $OCC_k$ |
|---|---|---|---|---|
| Class One | .713 | .717 | .984 | 24.755 |
| Class Two | .287 | .283 | .974 | 93.066 |

SM Table 2 presents the classification accuracy diagnostics for the three class model. Discrepancies between model estimated proportions for each class ($\hat{\pi}_k$) and modal class assignment proportions ($mcaP_k$) were small (absolute values of .004, .002, and .007 for classes one, two, and three, respectively). AvePP values were greater than .70 (class one = .972, class two = .969, and class three = .956), and all OCC values exceeded 5 (91.980, 94.276, and 23.823 for classes one, two, and three, respectively).

**SM Table 2.** Three Class Classification Accuracy Diagnostics

| Class *k* | $\hat{\pi}_k$ | $mcaP_k$ | $AvePP_k$ | $OCC_k$ |
|---|---|---|---|---|
| Class One | .274 | .269 | .972 | 91.980 |
| Class Two | .249 | .247 | .969 | 94.276 |
| Class Three | .477 | .484 | .956 | 23.823 |

From the classification accuracy diagnostics, it appeared that either the two or three class solutions had high classification accuracies. Therefore, it was necessary to explore the class separation of each model. To do this, the approach adopted by Oberski (2016) was used, which is to present the means of each latent class in what is known as a profile plot (SM Figure 2).




**SM Figure 2.** Profile Plot: Estimated Means for Ideal Expectation Items for Two and Three Class Solutions

For the two class solution (top plot in SM Figure 2), both classes were found to have high scores on the Ethical and Privacy Expectation items (EP1, EP2, EP3, EP4, and EP5). Where the two classes separated, however, were on the Service Expectations items (S1, S2, S3, S4, S5, S6, and S7). More specifically, individuals in class one had high scores across all Service Expectation items, whilst those in class two had low scores on these seven Service Expectation variables. The additional third class (bottom plot in SM Figure 2) was found to have high responses for all Ethical and Privacy Expectation items. As for the Service Expectation items, class three showed a similar response pattern to class one in that responses tended to be high. However, class one seemingly showed inflated expectations across each item, whilst the expectations of those in class three appeared to be more moderate.

A final step taken in choosing between the two and three class solutions was to assess the local fit of each model by examining the standardised residuals. For the two class solution, there were 434 of the 3234 (13.42%) absolute standardised residuals that exceeded 3; 196 (6.06%) of these were greater than 5. Improved local fit was found with the three class solution, with only 211 (6.52%) residuals exceeding 3 and 88 (2.72%) of these were greater than 5. An improved local fit would continue to be achieved if more classes were extracted (e.g., four or five classes). However, this would come at cost as the interpretability of the solution would have become increasingly difficult. Thus, on the basis of the relative fit, classification accuracy, class interpretability, and local fit the three class solution was selected as the candidate model. As noted, 6.52% of the absolute standardised residuals for this model did exceed 3, this is not excessive as in the case of the two class model (13.42% of residuals exceeding 3), but interpretation of the results was still taken with caution. For the three class solution, the following labels were given: the *Inflated Ideal Expectation* group (Class One; n = 334, 26.94%), the *Low Ideal Service Expectation* group (Class Two; n = 306, 24.68%), and the *High Ideal Expectation* group (Class Three; n = 600, 48.39%).

The logistic regression results from the three class model are presented in SM Table 3, which used class three as the baseline group. For class one, the covariates of gender, management, science, and technology, psychology and education, Postgraduate Student, European Student, or Overseas Student were not statistically significant at the 5% level. As for those variables that were statistically significant, the results found that those in class one are more likely to be older students ($p = .004$). As for class two, the covariates of gender, management, science, and technology, psychology and education, Postgraduate Student, European Student, and Overseas Student were not statistically significant at the 5% level. Only age was found to be statistically significant ($p = .032$) in that there was more chance of being in class two with increased age.

**SM Table 3.** Logistic Regressions using the Three Step Method with the Three Class Solution

| | Class One | | | Class Two | | |
|---|---|---|---|---|---|---|
| Covariate | Estimate | Standard Error | P-Value | Estimate | Standard Error | P-Value |
| Gender | .028 | .157 | .860 | .249 | .165 | .133 |
| Age | .018 | .006 | .004 | .014 | .006 | .032 |
| Management, Science, and Technology | .356 | .196 | .069 | -.113 | .211 | .592 |
| Psychology and Education | .251 | .190 | .187 | -.037 | .188 | .844 |
| Postgraduate | .073 | .154 | .637 | -.304 | .174 | .082 |
| European Student | .332 | .251 | .186 | -.033 | .285 | .907 |
| Overseas Student | .059 | .674 | .930 | .235 | .636 | .712 |

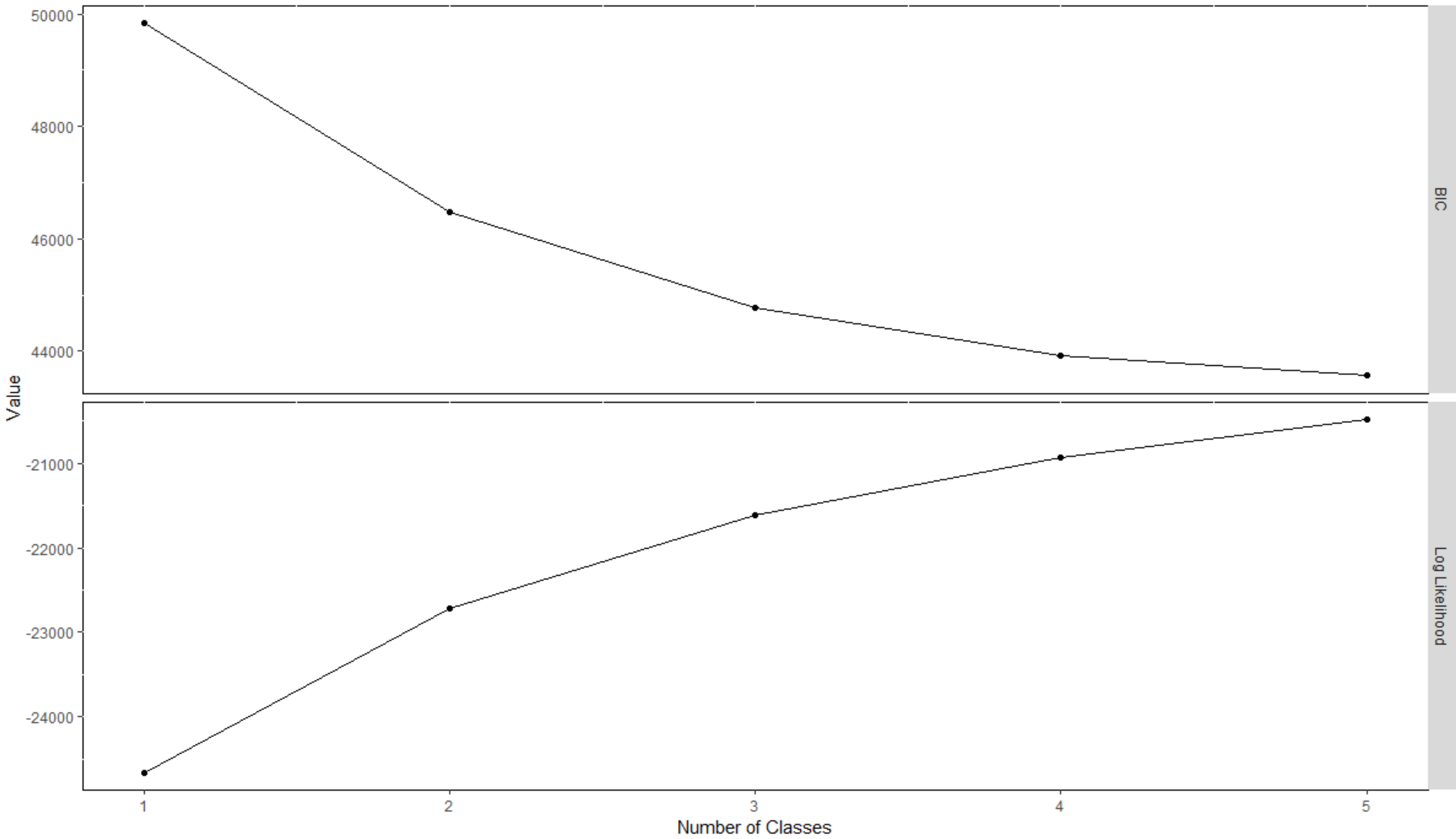
**SM 3. Detailed Results for the Predicted Expectation Scale**

One to six latent class models were estimated; however, the six class solution was not identified. Therefore, only the results of the one to five class solutions will be presented. With regards to the BIC values (SM Figure 3), either a two or three class solution would be supported on the basis of the “elbow criterion”.

To determine which of these two solutions (k =2 or k = 3) should be selected as a candidate model, the relative fit was assessed using the adjusted LMR-LRT and BLRT. For the two class solution, both tests showed this model to be a significant improvement over a one class solution (LMR-LRT = 3877.154, $p < .001$; BLRT = 3884.714, $p < .001$). Likewise, the fit of the three class solution was found to be a significant improvement over the two class solution (LMR-LRT = 2207.610, $p < .001$; BLRT = 2211.855, $p < .001$). At four classes, the adjusted LMR-LRT showed this solution to not provide a significantly improved fit over the three class solution (LMR-LRT = 1394.582, $p = .762$), but the BLRT output did support the four class model (BLRT = 1397.264, $p < .001$).

Taking the aforementioned evidence into consideration, it was clear that either the two or three class solution could still be selected as candidate models. The BLRT did support the four class solution, but there is a risk of this test never reaching a non-significant *p*-value. Thus, it was advisable to inspect a plot of log likelihood values for each solution and as with the BIC values, assess whether there is an “elbow”. From an examination of the plot of log likelihood values in SM Figure 3, a pronounced “elbow” was found at the two class solution.

From the evaluations of relative fit, it appeared that either the two or three class solutions were permissible solutions. Extraction of further classes (e.g., a four class solution) was not supported on the basis of the BIC and log likelihood plots (SM Figure 3) or the adjusted LMR-LRT. In light of these findings, it was decided that both the two and three class solutions would be compared in regards to classification accuracy, interpretability, and local fit.




**SM Figure 3.** Index Values across Five Latent Class Models

The relative entropy of the two and three class solutions were found to be .887 and .901, respectively. Thus, either model was considered to have good overall classification precision. To reiterate, however, the relative entropy values are not intended to be used in decisions of model selection. Rather, such decisions should be informed by an examination of the following classification diagnostics: AvePP, OCC, and mcaP (SM Tables 4 and 5).

SM Table 4 presents the classification accuracy measures for the two class model. It can be seen that the average posterior class probability (AvePP) for class one and two all exceeded .70, which shows the classes to be well separated. As for the odds of correction classification ratio (OCC), both values were greater than five, which is indicative of good assignment accuracy. As for the absolute differences between modal class assignment and model estimated proportions for each class, they were small (.004 and .005 for class one and two, respectively).

**SM Table 4.** Two Class Classification Accuracy Diagnostics

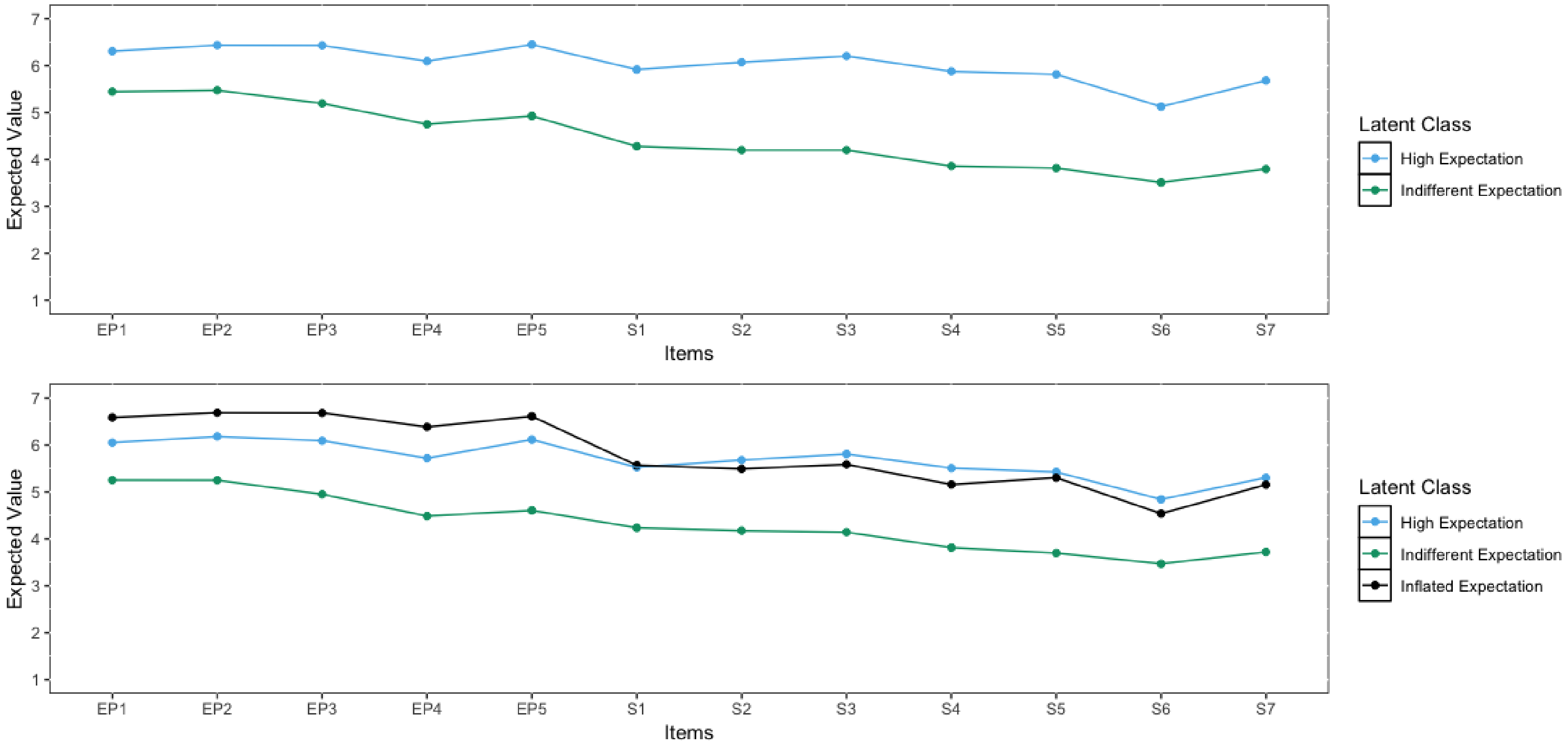
| Class *k* | $\hat{\pi}_k$ | $mcaP_k$ | $AvePP_k$ | $OCC_k$ |
|---|---|---|---|---|
| Class One | .472 | .468 | .971 | 37.455 |
| Class Two | .527 | .532 | .966 | 25.501 |

The classification accuracy results for the three class model are presented in SM Table 5. As with the two class solution, all AvePP values exceeded .70. With regards to the OCC values, these were all greater than 5. As for the discrepancies between the mcaP and model estimated proportions for each class, these absolute values were small (.001, .002, and .001 for class one, two, and three, respectively).

**SM Table 5.** Three Class Classification Accuracy Diagnostics

| Class *k* | $\hat{\pi}_k$ | $mcaP_k$ | $AvePP_k$ | $OCC_k$ |
|---|---|---|---|---|
| Class One | .436 | .435 | .954 | 26.828 |
| Class Two | .374 | .376 | .950 | 31.802 |
| Class Three | .190 | .189 | .966 | 121.124 |

Based on the classification accuracy diagnostics, either the two or three class models were found to be acceptable. Thus, the next step is to assess the interpretability and local fit of each latent class solution. The top plot in SM Figure 4 shows the two class solution, which shows class one to have high scores across all items. Class two, on the other hand, had high scores for the Ethical and Privacy Expectation items (EP1, EP2, EP3, EP4, and EP5), but for Service Expectation items (S1, S2, S3, S4, S5, S6, and S7) the scores are generally in the middle. As for the additional third class (bottom plot in SM Figure 4), this was not well differentiated from class one as it had high scores for both Ethical and Privacy Expectations and Service Expectations.




**SM Figure 4.** Profile Plot: Estimated Means for Ideal Expectation Items for Two and Three Class Solutions

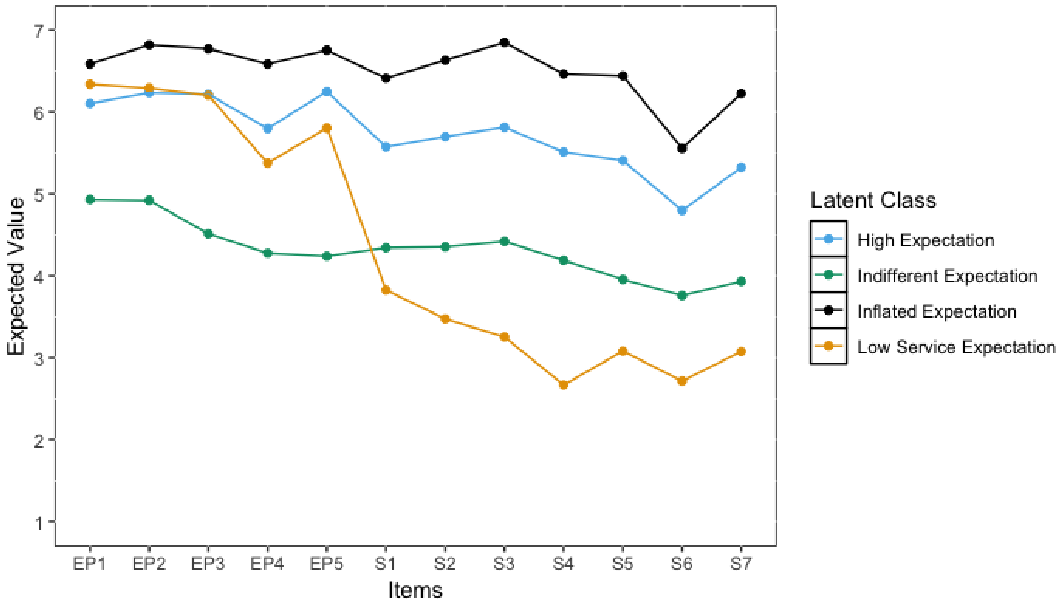
An examination of local fit for both models (k = 2 and k = 3), however, pointed to problems on account of the large proportion of high standardised residuals. For the two class model, 17.41% (n = 563) of the absolute standardised residual values exceeded 3 and 6.65% (n = 215) were greater than 5. With the three class solution, there was an improved local fit, but 10.45% (n = 338) of absolute standardised residual values exceeded 3, with 3.74% (n = 121) of values exceeding 5. Thus, it is clear that for both models the percentage of absolute standardised residual values that were greater than 3 was in excess of 5%. Given these local fit problems with both the two and three class solutions, it was necessary to assess whether the addition of a fourth class reduces the number of high standardised residuals and whether it provides an interpretable solution.

The classification accuracy diagnostics of the four class solution are presented in SM Table 6. It was found that the four class solution had good latent class assignment accuracy, as AvePP values exceeded .70, all OCC values exceeded 5, and the discrepancies between $\hat{\pi}$ and mcaP were small (absolute values = .001, .001, .001, .003 for class one, two, three, and four, respectively).

**SM Table 6.** Four Class Classification Accuracy Diagnostics

| Class *k* | $\hat{\pi}_k$ | $mcaP_k$ | $AvePP_k$ | $OCC_k$ |
|---|---|---|---|---|
| Class One | .402 | .403 | .954 | 30.851 |
| Class Two | .303 | .304 | .948 | 41.937 |
| Class Three | .138 | .139 | .967 | 183.038 |
| Class Four | .157 | .154 | .957 | 119.501 |

As can be seen from SM Figure 5, the addition of a fourth class did improve the interpretability of the model. Class four is shown to have high scores for the Ethical and Privacy Expectation items (EP1, EP2, EP3, EP4, and EP5), but low scores for the Service Expectation items (S1, S2, S3, S4, S5, S6, and S7). In terms of classes one and three, they were not well differentiated in the three class model; however, the differences became clearer with the use of a four class solution. More specifically, class three is characterised by inflated scores across all items; whereas, class one are at a lower level of expectation.




**SM Figure 5.** Profile Plot: Estimated Means for Ideal Expectation Items for Four Class Solutions

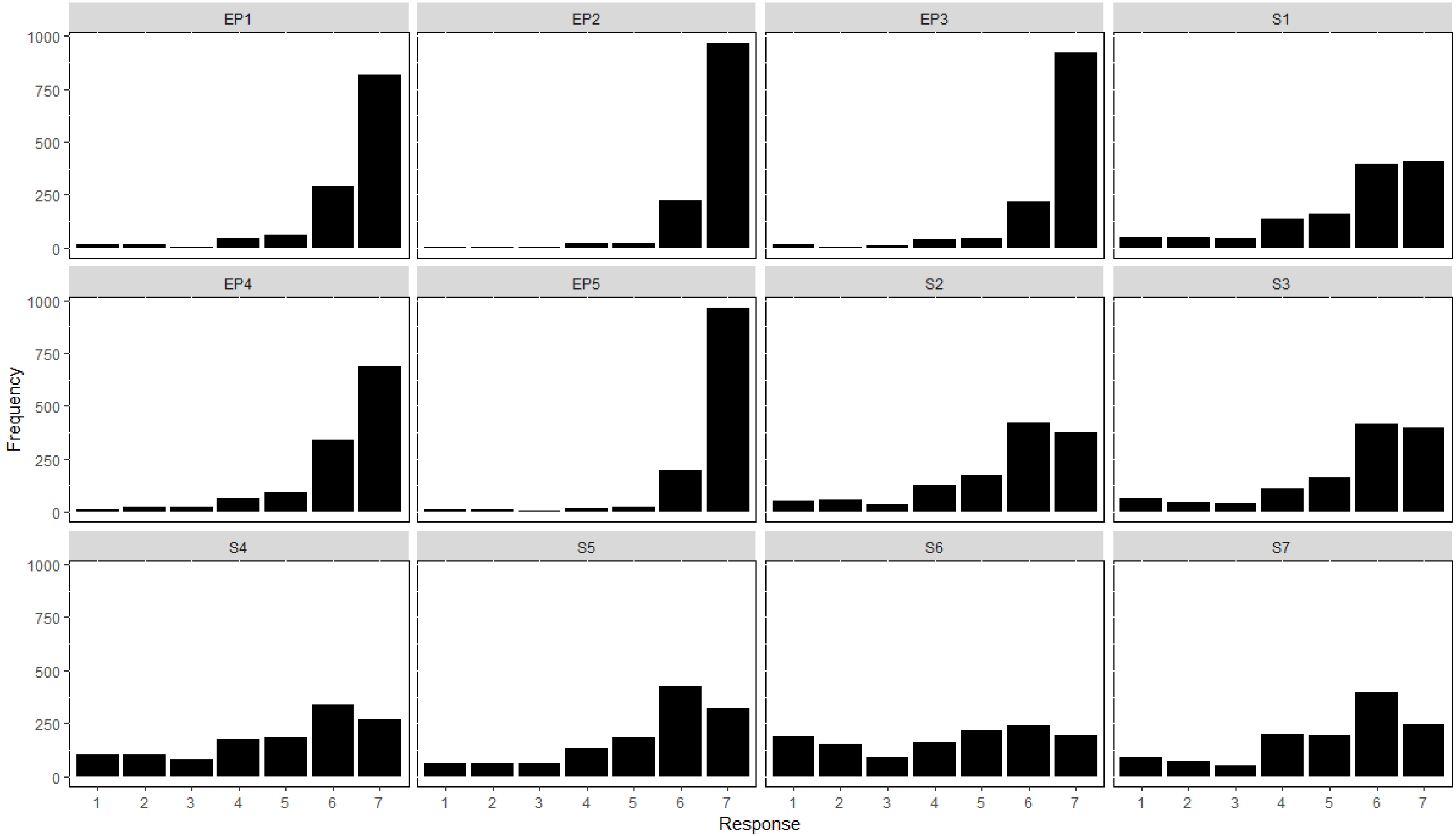
Along with the improved interpretability of the four class solution, the local fit was better than either the two or three class models. An examination of absolute standardised residual values shows 7.36% (n = 238) to exceed 3 and 2.54% (n = 82) to exceed 5. This showed that the addition of a fourth class did lead to a model with a better local fit. Even though the proportion of standardised residuals exceeding 3 remained greater than 5%, this is not as excessive as the proportions found for the two and three class solutions. Despite the information criteria (e.g., the BIC values) and adjusted LMR-LRT supporting either a two or three class solution, this also needs to be weighed up against the interpretability and local fit of each model. On the basis of the latter criteria, the four class model appeared more suitable and was supported by the BLRT; therefore, this was selected as the candidate model for the latent class regression. For this four class solution, the following labels were chosen: the *High Predicted Expectation* group (Class One; n = 500, 40.32%), the *Indifferent Predicted Expectation* group (Class Two; n = 377, 30.40%), the *Inflated Predicted Expectation* group (Class Three; n = 172, 13.87%), and the *Low Predicted Service Expectation* group (Class Four; n = 191, 15.40%).

For the latent class regression results (SM Table 7), class four was chosen as the baseline group. Starting with class one, older students are less likely to be assigned to this class ($p$ = .045). No other variable was found to be statistically significant at the 5% level for class one. As for class two, older students ($p$ = .003) and students who are European ($p$ = .015) are less likely to be assigned to this class. All remaining variables were found to not be statistically significant at the 5% level. Finally, with regards to class three, no variable was found to be statistically significant.

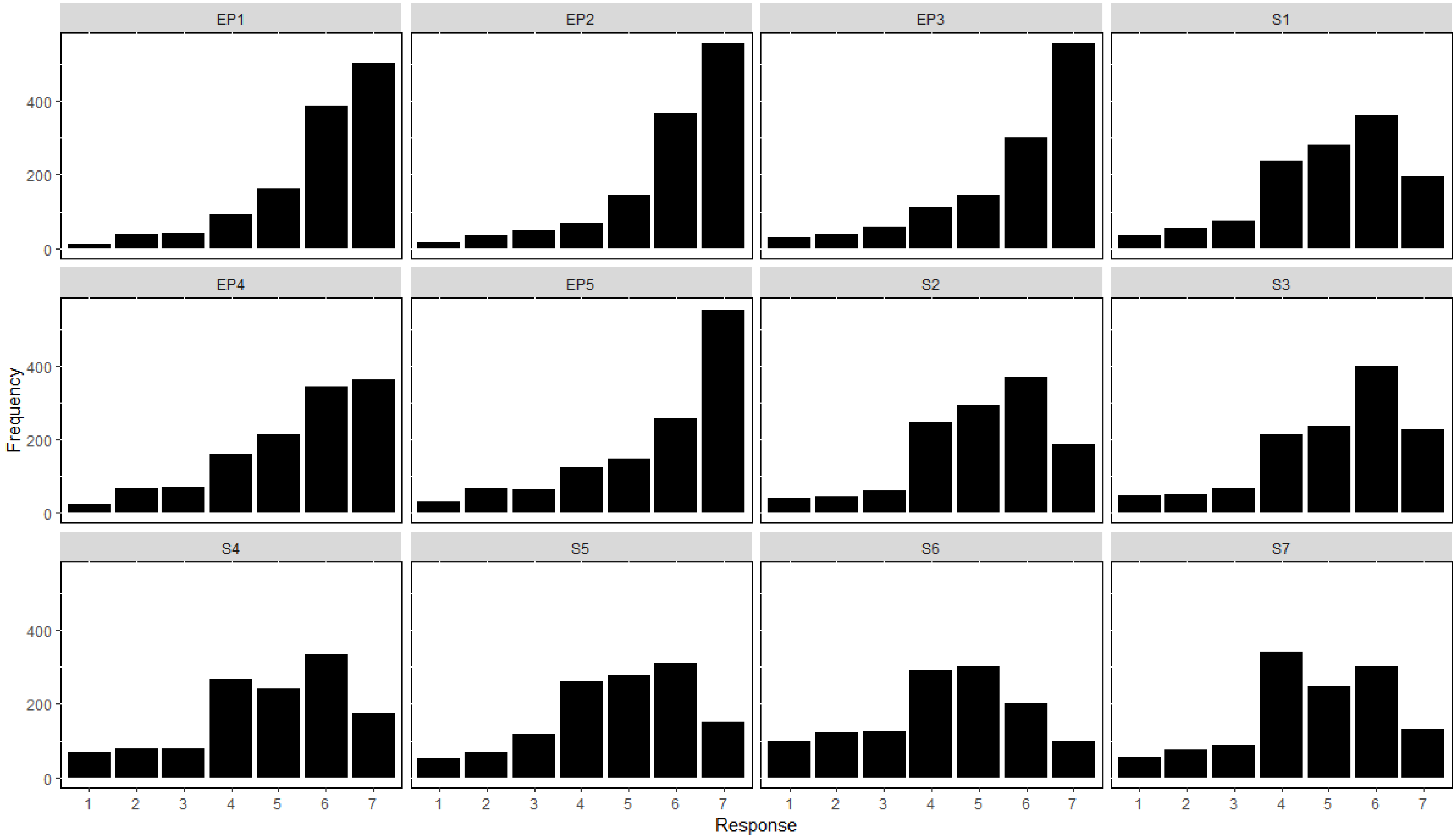
**SM Table 7.** Logistic Regressions using the Three Step Method with the Four Class Solution

| | Class One | | | Class Two | | | Class Three | | |
|---|---|---|---|---|---|---|---|---|---|
| Covariate | Estimate | Standard Error | P-Value | Estimate | Standard Error | P-Value | Estimate | Standard Error | P-Value |
| Gender | -.180 | .199 | .367 | -.359 | .211 | .089 | -.287 | .241 | .233 |
| Age | -.015 | .008 | .045 | -.024 | .008 | .003 | .010 | .009 | .272 |
| Management, Science, and Technology | .130 | .252 | .607 | -.058 | .267 | .828 | .250 | .297 | .401 |
| Psychology and Education | .281 | .232 | .226 | -.064 | .243 | .791 | .220 | .285 | .440 |
| Postgraduate | .236 | .207 | .256 | .075 | .222 | .737 | .083 | .244 | .733 |
| European Student | -.194 | .305 | .524 | -.927 | .382 | .015 | .476 | .337 | .158 |
| Overseas Student | .755 | 1.128 | .503 | -.189 | 1.307 | .885 | 2.066 | 1.154 | .073 |

**SM 4. Distribution Plots for Ideal Expectation Scale**



## SM 5. Distribution Plots for Predicted Expectation Scale



## References


Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*(3), 317–332. https://doi.org/10.1007/BF02294359

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341.

Hillenbrand, C., & Money, K. (2009). Segmenting stakeholders in terms of Corporate Responsibility: Implications for Reputation Management. *Australasian Marketing Journal (AMJ)*, *17*(2), 99–105. https://doi.org/10.1016/j.ausmj.2009.05.004

Lanza, S. T., & Rhoades, B. L. (2013). Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. *Prevention Science : The Official Journal of the Society for Prevention Research*, *14*(2), 157–168. https://doi.org/10.1007/s11121-011-0201-1

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, *42*(10), 1–29.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

Mannarini, S., Boffo, M., Rossi, A., & Balottin, L. (2018). Etiological Beliefs, Treatments, Stigmatizing Attitudes toward Schizophrenia. What Do Italians and Israelis Think? *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02289

Masyn, K. E. (2013). Latent Class Analysis and Finite Mixture Modeling. *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*. https://doi.org/10.1093/oxfordhb/9780199934898.013.0025

McLachlan, G., & Peel, D. (2000). *Finite Mixture Models* (1 edition). New York: Wiley-Blackwell.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User’s Guide* (Eighth Edition). Los Angeles, CA: Muthén & Muthén.

Nagin, D. (2005). *Group-Based Modeling of Development.* Cambridge: Harvard University Press.

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*(4), 535–569.

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In *Modern statistical methods for HCI* (pp. 275–287). Springer.

Ramaswamy, V., Desarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with Pims Data. *Marketing Science*, *12*(1), 103.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Strizich, G., Gammon, M. D., Jacobson, J. S., Wall, M., Abrahamson, P., Bradshaw, P. T., … Greenlee, H. (2015). Latent class analysis suggests four distinct classes of complementary medicine users among women with breast cancer. *BMC Complementary and Alternative Medicine*, *15*(1). https://doi.org/10.1186/s12906-015-0937-4

Vermunt, J. K. (2010). Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis*, *18*(04), 450–469. https://doi.org/10.1093/pan/mpq025