EDRS 821

Exploratory Factor Analysis

1. Using the “factor analysis” data set, develop a factor structure that best fits your data **(maximizes explained variance, parsimonious, theoretically plausible)**. There are 22 items about teacher beliefs: items t1-t22. All 22 items were evaluated on a 1-5 point Likert scale.
* Be sure that your data is cleaned before running your analysis.
* Be sure to evaluate the adequacy of your data for factor analysis including:

Sample size

Normality, Linearity **(determinant)** and Outliers

Multicollinearity

Factorability of matrix

 Prior to conducting the factor analysis, all data were screened for model assumption of normality and linearity, and correlation using histograms, and P-P plots and a correlation matrix with Bartlett’s test of sphericity. Model assumptions of normality and linearity were met. Other assumptions met included:

* Sample size (N = 163) was reasonably close to fair (100 poor, 200 fair)
* A reasonable ratio of **163/7 (N/k)** was present**;** and,
* \*At least some of the *r* correlations differ significantly from zero indicating some degree of collinearity among the variables but not an extreme degree or singularity among the variables, X2 (66)= 691.686, p < .001. The degree of common variance among the twenty-two variables is almost “mediocre,” KMO = .574.

******

 Using descriptive to identify outliers, one score was deleted from the variable “the best teachers are passionate about their work’ as it was beyond the permissible values of the 1-5 likert scale (8). Univariate and multivariate analyses were performed to examine if outliers were present. Ten scores were found to either be above the three standard deviation or beyond the mahalanobis threshold of 33.92 (for twenty two variables); however, these were retained as it was deemed important to keep all the data in order to allow the PC analysis to discover the dimensionalities rather than cleaning to a point that scores may have been forced into groups.

1. You may use either a PC or PAF method and also choose the rotation method that you feel is most appropriate for your data. Examine both a varimax and an oblimin rotation.
2. Drop any items that do not have adequate communalities (remember that in class we discussed guidelines—you must make decisions based on your data patterns.)

|  |
| --- |
| Table 1 Items for Principal Component Analysis |
| Item |
| 1. When a student does better than usual, many times it is because I exert a little extra effort.
 |
| 1. Teaching is a skill that can only be earned and developed through practice.
 |
| 1. The hours in my class have little influence on students compared to the influence of their home environment.
 |
| 1. It is easy to recognize quality teaching.
 |
| 1. The amount a student can learn is primarily related to family background.
 |
| 1. The best teachers are passionate about their work.
 |
| 1. If students aren't disciplined at home, they \aren't likely to accept any \discipline.
 |
| 1. It is important to understand the theory behind teaching techniques.
 |
| 1. Knowing how to use and implement teaching techniques is the hallmark of a good teacher.
 |
| 1. Knowledge about instructional practices is the most important knowledge a teacher can have.
 |
| 1. Expertise in teaching can be developed after only a few years of practice.
 |
| 1. Anyone can be a teacher.
 |
| 1. Academic dishonesty is a serious problem in this school.
 |
| 1. A teacher is very limited in what he/she can achieve because a student's home environment has a large influence on his/her achievement.
 |
| 1. Teaching is a talent. Some people have it, and some people don't.
 |
| 1. Teachers are not a very powerful influence on student achievement when all factors are considered.
 |
| 1. Good teachers get through most of their day on instinct.
 |
| 1. Expert subject matter knowledge is necessary for effective teaching.
 |
| 1. Knowledge about how to motivate students is essential for teaching.
 |
| 1. If parents would do more for their children, I could do more.
 |
| 1. As long as teachers know how to manage a classroom students will learn.
 |
| 1. The influences of a student's home experiences can be overcome by good teaching.
 |

1. A Principal Component (PC) factor analysis was performed on a set of **22 items** to assess the dimensionality that account for the patterns of collinearity among the variables. The default criterion to retain only factors with eigenvalues greater than 1, and the varimax and Oblimin rotations were requested. The items consisted of self-reporting ratings on teacher beliefs (see table 1 for list of items). Each item was rated on a 5-point scale that ranged from 1 (somewhat true) to 5 (mostly true). The component analysis indicated that the items formed into eight separate groups. One variable was dropped from the analysis as it did not meet the set threshold of greater than or equal to 5 communality value, h2 = .447 (T15).
2. Because the T15 was dropped, the PC analysis was rerun. The component matrix indicated that the **twenty-one items seemed to form into eight separate groups** with communalities of greater than 5. The component matrix (CM) indicated the 8 groups were not patterned and upon closer inspection it was decided that it did not make sense to attempt to interpret this eight-factor model as it contained 2 factors with only two items loaded on it, in addition the model showed cross-loading. A rotated component matrix (RCM) corrected the cross-loading but the RCM also contained 4 factors with item loadings of 2.
3. This made it difficult to theorize about the nature of the latent variable the factors measured. Therefore, the **model was reduced to 7 factors with 21 items**. The test of sufficient variance in the model was miserable, .514. In addition, the communality for item T20 was .425. The decision was made to **delete** it
4. The model was rerun with **20 items and an extraction of 7 factors,** which yielded communality values of above .5, but the test of sufficient variance in the model was miserable, .514. However the variance for the model was a whopping 69%, but the loading on the component matrix showed cross-loadings and it contained 4 factors with less than 3 items. The rotated component model fixed the cross-loading problem but the model retained 4 factors with loading less then 3 items.
5. The model extraction was reduced to **6 factors with 20 items** in order to explore if a smaller set of factors would improve variable loadings for each factor. With 6 factors the communality of T2 (.238) and T3 (.397) were below the threshold value of 50%. Therefore the decision was made to drop item T2.
6. The model with **6 factors and 19 items** showed a marked improvement in the communality, all variables were at or above .5; in addition the cumulative variance was 66 percent. The test of sufficient variance in the model was improved but still miserable, .532. The factor loadings were a problem too. Two factors had only one item loaded in the CM, and there were two cross-loadings. Using the RCM, the cross loading was remedied but 1 of the 6 factors had only 1 item loaded on it. The decision was made to reduce the factors.
7. A model of **19 items with five factors** for extraction was run; the communality of two variables was a problem, T3 (.391) and T11 (.362). It was decided to drop both T11and T3 (one at a time) and rerun the analysis.
8. The model with **17 items and 5 factors** resulted in a KMO of miserable, .544; however, the variance explained by the factors was 63%. The communality for all factors was at or above the .5 thresholds. However, the CM showed one cross-loading and 2 factors with loading of 2 items. The RCM fixed the cross-loading issue but not the loadings. It was decided to reduce the factors to four.
9. In the **17 items with 4 factor model** there were three items with low communality values, T1 (.383), T6 (.412) and T19 (1.68). Item T19 was deleted as it was the lowest value and it was theorized that dropping this item may improve the other item values.
10. The analysis was rerun for **16 items with 4 factors** extracted. Items T1 (.275) and T6 (.364) showed low communality values. Item T1 was deleted.
11. After rerunning the analysis for **15 items with 4 factors** item T6 (.374) was deleted.
12. The PC analysis was rerun with **14 items and 4 factors** extracted. The communality for extracted factors was at or above 50%. The variance explained was 63%, however, the test of sufficient variance in the model was still miserable, .544. In addition the component matrix showed 2 factors with only 2 items loading, the rotated component model also showed 2 factors with only 2 items loaded. The decision was made to further reduce the factors.
13. The PC analysis was rerun with **14 items and 3 factors**. Item T18 (.146) and T13 (.288) were below the threshold value of communality (.5). T18 was deleted.
14. A **13 item with 3 factor** model showed communality value problems with T13 (.267) continued, therefore it was deleted.
15. The final Principal Component model was a **12-item component with 3 factors.** The model had a Determinant of Correlation Matrix value of .013 (above the .00001) that showed the matrix could be inverted (linear). The test of sufficient variance in the model was .574, almost mediocre. All factors had a communality of .50 or greater, and the total variance explained by these factors was 59%.

 **Final Results**

 An exploratory factor analysis was performed to evaluate whether an eight-factors model made sense for the data. The SPSS factor analysis procedure was used, the method of extraction was PC, only factors with eigenvalues >1 were retained, and varimax and oblimin rotations were requested. Three factors were retained and rotated. Results from this analysis, including rotated factor loadings, communalities, and SSL for the retained factors, are summarized in Table 2.

|  |
| --- |
| **Table 2 Rotated Factor Loading from Final Analysis: Principal Component with Oblimin and Varimax** |
| Nine self-Rated Teacher Belief Items |
|  Oblimin Rotation Varimax Rottion |
|  | Factor1  | Factor2 | Factor3 | Factor1  | Factor2 | Factor3 | Communality h2 |
| Knowledge about instructional practices is the most important knowledge a teacher can have. | **.757** | -.289 | -.223 | **.755** | -.305 | -.266 | .733 |
| Knowing how to use and implement teaching techniques is the hallmark of a good teacher. | **.737** | -.054 | .125 | **.739** | -.059 | .084 | .556 |
| Teaching is a skill that can only be learned and developed through practice.  | **.681** | -.035 | -.087 | **.678** | -.046 | -.125 | .478 |
| It is important to understand the theory behind teaching techniques. | **.665** | .119 | -.108 | **.661** | .107 | -.143 | .469 |
| It is easy to recognize quality teaching. | **.622** | .022 | .269 | **.625** | .023 | .235 | .447 |
| A teacher is very limited in what he/she can achieve because a student's home environment has a large influence on his/her achievement. | -.077 | **.840** | -.106 | -.087 | **.837** | -.094 | .717 |
| The amount a student can learn is primarily related to family background | -.013 | **.782** | .074 | -.020 | **.785** | .082 | .623 |
| If students aren’t disciplined at home, they aren’t likely to accept any discipline. | -.138 | **.671** | -.324 | -.150 | **.663** | -.310 | .558 |
| Teachers are not a very powerful influence on student achievement when all factors are considered. | .086 | **.645** | .238 | .084 | **.651** | .239 | .488 |
| Good teachers get through most of their day on instinct. | .379 | .315 | **.742** | .387 | .333 | **.724** | .785 |
| The influences of a student’s home experiences can be overcome by good teaching. | .213 | .230 | **-.690** | .199 | .206 | **-.700** | .572 |
| As long as teachers know how to manage a classroom students will learn. | -.533 | .033 | **.575** | -.524 | .057 | **.605** | .643 |
| Sum of Shared Loading | 2.905 | 2.447 | 1.717 | 2.935 | 2.454 | 1.718 |  |
| % Explained Variance | 24.21% | 20.40% | 14.31% | 25.45% | 19.71% | 13.72% |  |

 After varimax and oblimin rotations, the factors accounted for 58.9% of the variance using Varimax and 58.9% variance using oblimin; the variance contribution by all three factors showed an acceptable contribution pattern (see table 2). Communalities for variables were generally reasonably high, ranging from a low of .447 to a high of .785. Rotated factor loading were examined to assess the nature of these three retained oblimin and varimax-loadings. An arbitrary criterion was used to decide which factor loadings were large; a loading was interpreted as large if it exceeded .40 in absolute magnitude. There were five high positive correlations among the ratings on ‘Knowledge about instructional practices is the most important knowledge a teacher can have’ Knowing how to use and implement teaching techniques is the hallmark of a good teacher,’ ‘Teaching is a skill that can only be learned and developed through practice,’ ‘It is important to understand the theory behind teaching techniques,’ and ‘It is easy to recognize quality teaching;’ these items were identified as “teacher knowledge” **(factor 1)** as the variables all seem to indicate experience and expertise of the teacher. There were four high positive correlations among the items ‘A teacher is very limited in what he/she can achieve because a student's home environment has a large influence on his/her,’ ‘The amount a student can learn is primarily related to family background,’ ‘If students aren't disciplined at home, they aren't likely to accept any discipline,’ and, ‘Teachers are not a very powerful influence on student achievement when all factors are considered; this factor was difficult to name as it seems to relate to teacher beliefs about socio-economic influences and the quality of child rearing/supervision, it was decided to identify these items as “home supports” **(factor 2)** as the variables relate to levels of support *from home/family/neighborhood*.

 Three items with moderate positive and negative correlations were ‘Good teachers get through most of their day on instinct,’ ‘The influences of a student's home experiences an be overcome by good teaching,’ and, ‘As long as teachers know how to manage a classroom students will learn;’ these items were all identified as “teacher qualities” **(factor 3)**, as they seem to relate to teacher abilities and talents as influencers of student outcomes.

 \*On the final results write-up, I will use the varimax rotation to report my results. Although both produced similar results, the varimax uncorrelated values seem be closer to each other, easier pattern to examine/interpret groupings. Another thing that I will keep in mind for the future is to reference the reproduced correlation matrix earlier in the analysis when attempting to determine how items are co-varying together.

In your APA style write-up, include a brief summary of your sample and then summarize the results of your solution, talk about the similarities and differences between methods you examined, and give a rationale for which model you would suggest accepting.

1. Cluster Analysis
2. Using the data set Miller\_cluster.sav, classify students on the variables of Desired moral Approbation Self, Desired Moral Approbation Others-Praise, and Desired Moral Approbation Others-Blame.

The sample consists of 777 high school students. Assume adequacy of sample and theoretical justification for the clustering variables.

1. Your analysis should include the following steps:
* Clean data
* Choose cluster method
* Justification of number of clusters
* Interpretation of clusters
* Validate your cluster solution (there are other variables in the data set which may be used for this portion of the assignment).
1. In your APA style write-up, include a brief summary of your sample and summarize the results of your analysis.