NAEP study on special education

Project Description:

- Background to the client's research:

The data we have in this study was gathered from the National Assessment of Educational Progress (NAEP) which is the largest assessment of America's students' in different fields including but not limited to mathematics, reading, science, writing, the arts, civics, economics, geography, U.S. history, and in Technology and Engineering Literacy (TEL).

The data is survey-based and contains 166640 students including 1220 students with emotional disturbances and 13250 students with other disabilities. The questionnaire asked about different teaching strategies and professional development experience and how these could affect reading score of the 4^{th} grade students. This data recorded in 2009, 2011 and 2013 but students and teachers are not the same during these times and thus no longitudinal analysis is needed.

Besides teaching strategy and professional development, the researchers also collected information about teachers' educational background and students' disability status. These two factors could potentially play an important role on students' performance in reading. In the current version of the data set, the factors have been changed to binary values. The original data set was in the format of the categorical data. This was done either to simplify the data set or for confidentiality but it causes many problems and limits to our analysis.

- Research questions associated with this project
    - Which kinds of professional development experience and/or teaching strategies yield best student outcomes (4^{th} grader's reading score)?

- Whether teachers' educational background significantly moderates the abovementioned association?
- Whether students' disability status significantly moderates the abovementioned association?
- How the SEM model evolve over time?

We only need to do the first research question.

- Experimental Design/ Data Collection:

As I mentioned in the background the data was collected from a survey and has many samples based on observational study. Table 1 shows the structure of the data.

Student ID	Teacher ID	Disability level	Teacher's background	Teaching strategy and professional development (54 variables)	Reading score
1	1	2	6	1	265.8
2	3	3	7	1	171.95
3	3	1	9	0	145.43

Table 1. The structure of the data. This is an example of how data would look like and it's not copied from the data.

- Variable list with description:

There are 54 variables (questions/items) in teaching strategy and professional development. These variables are recorded as "yes" and "no" and thus they have two levels. Each of the variables are related to a specific question. For instance, in variable "T106003r" the question was:

Have you used a computer or other technological resources to practice spelling and grammar?

And the response to this was either yes or no (NA shows missing values). Similarly the rest of the 54 variables are asking about different methods or strategies that have been used in the class. Some of them related to teaching strategies and some related to professional development. For instance the following questions are related to teaching strategies:

- Whether language arts are integrated with other subjects?
- Whether you have provided instruction in fiction?
- Whether you have provided instruction in literary nonfiction?

And the following examples are related to professional development:

- Whether your school system provides you with the materials and other resources you need for reading/language arts instruction?
- Have you learned “how students learn reading” during the past two years?
- Have you learned “content standards in reading” during the past two years?

The details of these questions were provided by the client later. Later when I do the EFA I will use these details to analyze why a latent factor could integrate some of these questions in one module.

The teacher’s background has 9 different levels: "Multiple", "High-school diploma", "Assoc-deg/voc cert", “Bachelor's degree", "Master's degree", "Education specialist", "Doctorate", "Professional degree" and "Omitted".

The students’ disability has 3 levels: "Typical", "Emotional disturbance" and "Other disability (without ED)”.

The reading score is a continuous variable range from 37.22 to 354.70 with 1071 missing values. The histogram of the reading score is presented in Figure1 and it seems that it's normally distributed.

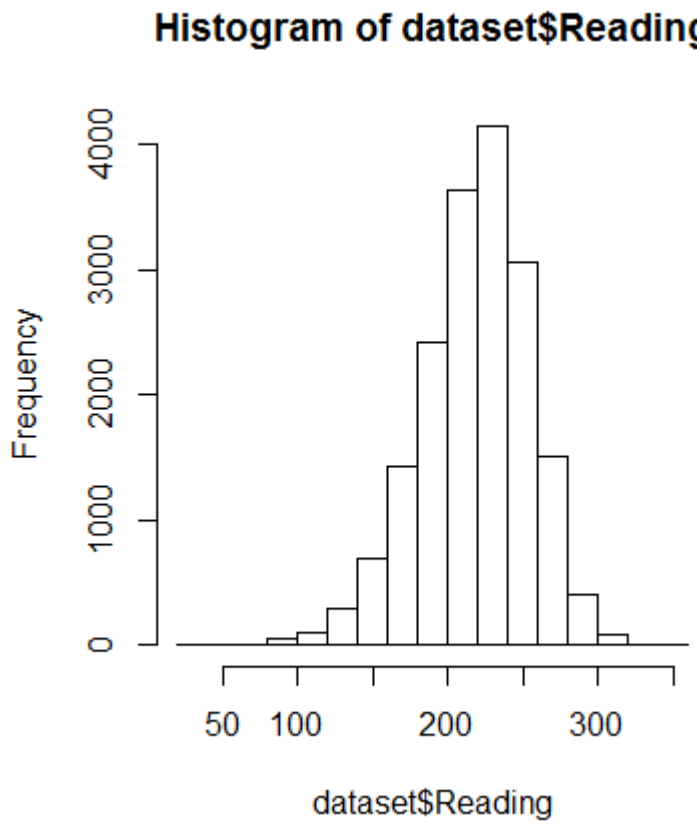

Figure 1. Histogram of the reading score

The teaching strategies, professional development, teachers' background and students' disability are independent variables and the reading score is the dependent (response) variable.

- Anticipated relationship between variables

The goal of this research was to investigate possible moderator on reading score while teaching strategies' (TS) and professional developments' (PD) factors are playing key roles for having a

greater scores in reading. Also they were interested to examine the effect of the educational background and disability status on these factors. In other words, whether there are any interactions between these factors or not that could potentially affect the reading scores. Figure 2 simplified these relationships.

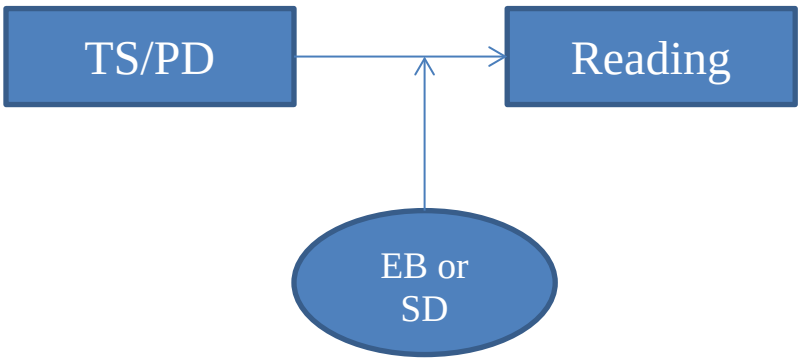

Figure 2. How different factors could impact reading score (both main effects and their interactions)

Methods:

I used exploratory factor analysis (EFA) to find the main predictors (from TS and PD variables) that could change the reading score. When we have so many factors that are correlated with each other or assessing different aspects of a particular predictor and they could each cause some effects on the response variable, we can summarize them using EFA which tries to map these factors in to a bigger (latent) “macro-factor”. In other words, this method groups the “micro-factors” that have similarities or higher correlations. Then these “macro-factors” could be used in any analysis e.g. linear regression. Another method to do this analysis is to use the simple regression model but with regularization. Since these “micro-factors” are correlated they would cause collinearity. To deal with this problem, one can use ridge, lasso or any other regularization form.

Figure 3 describes the EFA method. In that, the micF1 to micF66 are the micro-factors. In our dataset, I have 54 of them. Then I mapped them to macro-factors which I chose to have only 2 in this assignment and then each of them could be a predictor for reading score (linear regression).

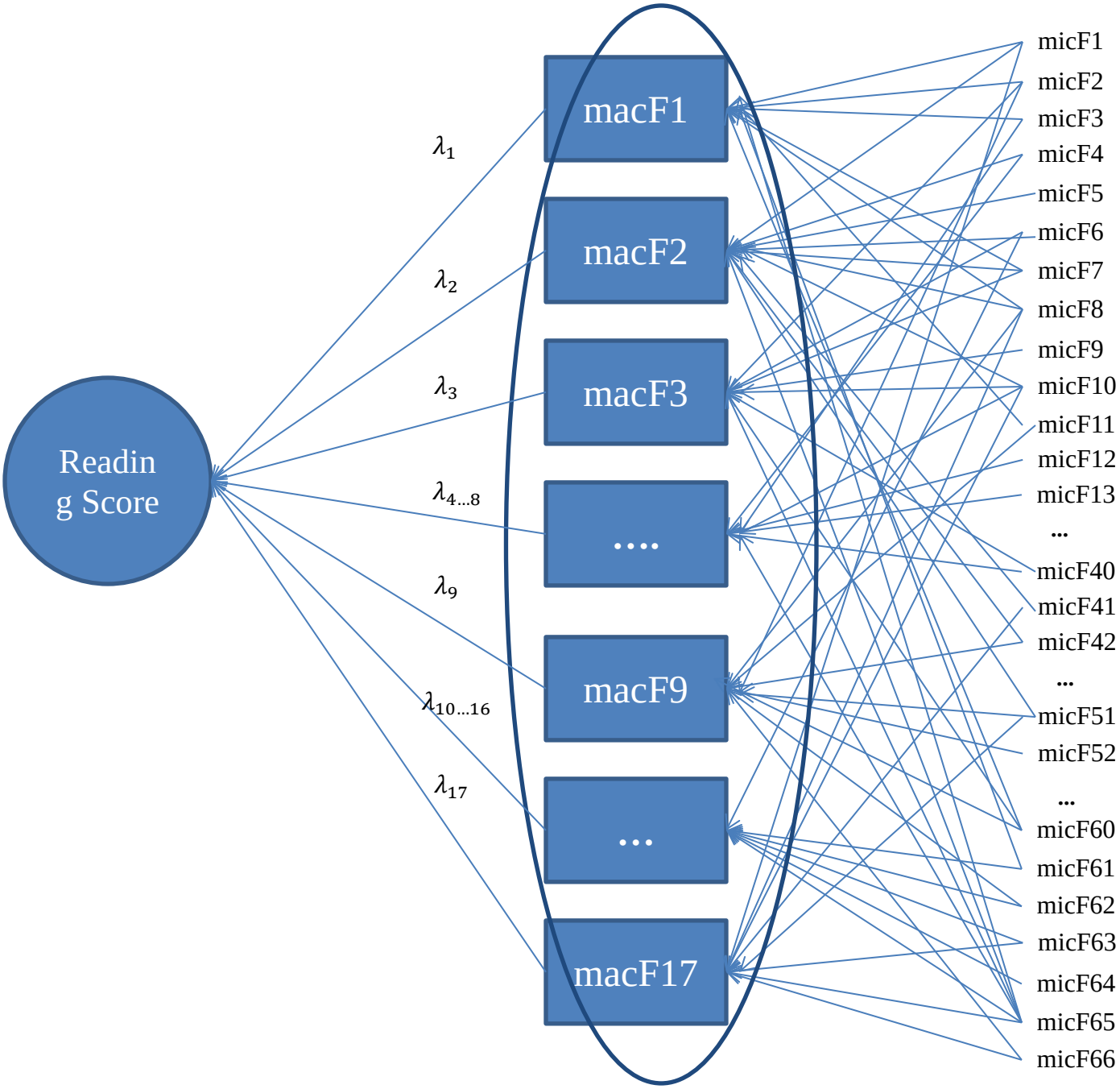

Figure 3. EFA integrates the micro-factors into macro-factors.

First I changed the format of the data to an appropriate format of zeros and ones. Then I used the thetrachoric correlation to find the correlation between these variables. The result is not really informative but a 54 by 54 correlation matrix along with tau (the tetrachoric cut points).

Some of the variables are highly correlated. Also in many of them there are only “yes” or “1” with missing values (no “0”). Thus I decided to exclude these variables. The remaining variables are 35 (reduced from 54). After removing the “NA” s from the data the number of subjects reduced to 251. Now dealing with the correlation matrix is easier. Feeding the correlation matrix into the “factanal” function to run the EFA, we have the chi-square of 231.9658 with 526 as the degrees of freedom which means that we can’t reject the null hypothesis of the perfect fit. So I don’t need to increase the number of factors.

The first factor is heavily loaded by V13, V21, V22, V24 (the loads are more than 0.85). Looking back to data, these variables are related to

- Whether you have provided instruction in literary nonfiction?
- Have you used a different set of methods in teaching reading in some students?
- Have you supplemented the regular course curriculum with additional material for some students in teaching reading?
- Have you had some students engaged in different classroom activities?

These are all related to “Different teaching strategies for different group of students” and “Teacher’s teaching strategies/Practices”. I could set a lower threshold (e.g. loads greater than 0.6) to include other aspects. But before that I checked factor 2 with loads greater than 0.85. It seems that factor 2 is heavily loaded by V5, V8 and V9 which are related to

- Have you learned “how students learn reading” during the past two years?

- Have you learned “content standards in reading” during the past two years?
- Have you learned “instructional methods for teaching reading” during the past two years?

These are related to “Teachers’ Professional Development on Reading Instruction (PD)”. So with this conservative threshold, I found that among teaching strategies’ aspects, the “Different teaching strategies for different group of students” and “Teacher’s teaching strategies/Practices” and among professional development’s aspects the “Teachers’ Professional Development on Reading Instruction (PD)” are the main variables.

Note that I did similar analysis with the original data (without removing the “NA”) and I found 9 interesting factors, summarized in table 2:

Factors	Loaded by	PD/TS
Factor 1	V5, V6, V7, V8, V9, V10, V11	Teachers’ Professional Development on Reading Instruction (PD)
Factor 2	V21, V22, V23, V24, V25	Different teaching strategies for different group of students
Factor 3	V12, V13, V14, V15, V16	Teachers’ Teaching Strategies/Practices
Factor 4	V48, V49, V50	Technology Based Teaching Strategies on Reading
Factor 5	V1, V2, V3, V4	Teachers’ Educational Background (ED)
Factor 6	V18, V19, V20	Teachers’ Teaching Strategies/Practices
Factor 7	V39, V40, V41, V42	Teachers’ Teaching Strategies/Practices
Factor 8	V44, V46, V47	Teachers’ Teaching Strategies/Practices
Factor 9	V1, V18, V16, V45 (coefficients are smaller than other factors)	Teachers’ Educational Background (ED)/ Teachers’ Teaching Strategies/Practices

Table 2. Result of EFA on the original data without deleting the NA along with the loads

These results confirmed my first analysis but with more details. Having more factors with more data creates another predictor, Teachers' Educational Background (ED), to my factors. For simplicity I will use the results that I got by removing the "NA"s. There is a debate on whether "factanal" function could deal with Na variables correctly. Having only factors would explain 0.44 of the variances in the data (but with 9 factors it was about 0.63).

The scores from these factors can be calculated by

$$score = ZR^{-1}A$$

where A are the loads, R^{-1} is the correlation matrix and Z is the standardized data. It could be also calculated by setting scores="regression". These scores could be used for further analysis e.g. regression later. For instance in research question 2 and 3 we were asked to investigate the effect of teachers' educational background and students' disability status in the format of interaction. Simple averaging could also be useful after we figure out which variables are important using EFA. There are many methods to calculate these scores, [1].

I also analyzed the professional development and teaching strategies separately and the results are summarized below:

EFA on PD:

- 3 factors have been found.
- The loads are from "Teachers' Professional Development on Reading Instruction (PD)" limited to these items:

  Have you learned "how students learn reading" during the past two years?

Have you learned “content standards in reading” during the past two years?

Have you learned “curricular materials available in reading” during the past two years?

Have you learned “instructional methods for teaching reading” during the past two years?

Have you learned “methods for assessing students in reading” during the past two years?

Have you learned “preparation of students for district and state assessments” during the past two years?

Have you learned “strategies for teaching reading to students from diverse backgrounds” during the past two years?

- These three factors could explain 0.83 of the variations in the data, summarized in table 3.

```
Loadings:
    Factor1 Factor2 Factor3
V5  0.78    0.28    0.40
V7  0.63    0.48    0.36
V8  0.84    0.41    0.30
V9  0.66    0.57    0.30
V11 0.58    0.44    0.23
V10 0.35    0.75    0.32
V6  0.48    0.50    0.71

                Factor1 Factor2 Factor3
SS loadings       2.84    1.81    1.13
Proportion Var    0.41    0.26    0.16
Cumulative Var    0.41    0.66    0.83
```

Table 3. The result of EFA on PD

EFA on TS:

- 2 factors have been found.
- These factors could explain 0.39 of the variations in the data, summarized in table 4.
- The loads are from “Technology Based Teaching Strategies on Reading”, “Teachers’ Teaching Strategies/Practices” with the following items:

Technology:

Have you used a computer or other technological resources to build and practice vocabulary?

Have you used a computer or other technological resources to increase reading fluency and comprehension?

Have you used a computer or other technological resources to practice spelling and grammar?

Strategies/Practices:

Whether you’ve asked students to interpret the meaning of the passage when reading a story?

Whether you’ve asked students to question the motives or feelings of the characters when reading a story?

Whether you’ve asked students to identify the main themes of the passage when reading a story?

Whether you have provided instruction in fiction?

Whether you have provided instruction in literary nonfiction?

Loadings:

	Factor1	Factor2
V12	0.75	
V13	0.68	
V15	0.50	
V18	0.62	0.29
V19	0.69	
V20	0.65	0.23
V21	0.61	
V22	0.60	0.20
V23	0.56	
V39	0.67	
V40	0.77	
V41	0.73	
V42	0.68	
V44	0.53	
V47	0.61	
V48		0.94
V49		0.87
V50		0.86
V14	0.46	0.22
V16	0.46	0.26
V17	0.48	0.21
V24		0.50
V25		0.37
V38		0.33
V43		
V45	0.40	0.23
V46	0.45	0.24
V51	0.28	0.49
V52	0.32	0.37
V53		0.39

	Factor1	Factor2
SS loadings	8.04	3.61
Proportion Var	0.27	0.12
Cumulative Var	0.27	0.39

Table 4. The result of EFA on TS

Results from different analyses are almost the same except technology-based strategies that have been picked by EFA only after considering the teaching strategies separately. Thus I can conclude that among different teaching strategies and professional developments' factors, technology-based strategies, practices and reading instructions could be the most important factors.

Discussion:

There were 54 variables (2-level factors) in our data with more than 166000 samples. There are two methods to handle large data sets with many predictors that could be correlated. One is to use the linear regression with regularization. Since we want to select the predictors (feature selection) we should use LASSO. LASSO can make the coefficients of unnecessary predictors zero and thus eliminate them. Another method is to use exploratory factor analysis (EFA) which analyzes the data similar to principle component analysis. In EFA, we look for latent (macro) factors that could explain the variations in the data as much as possible. The number of factors is defined arbitrary but by increasing the number of factors we need to make sure that more variations (of the data) could be explained and also the p-value for the chi-square should not be significant.

Dealing with "NA" in "factanal" function is problematic so I deleted them and the analyzed the data. Based on that, I found two factors that are related to "Teachers' Professional Development on Reading Instruction (PD)" and "Different teaching strategies for different group of students" and "Teacher's teaching strategies/Practices". This means that these factors could potentially yield best student outcomes. Also I did similar analysis on professional development and teaching strategies separately and I found that besides the previous factors, "Technology Based Teaching Strategies on Reading" is also important.

Limitation:

1- Converting categorical data to binary data would affect the correlation matrix. In fact with this conversion, we lose information. For instance, one variable contains only "yes" or (1) for all students in the binary format while it could be different in the original data.

2- While EFA is helpful in identifying the important factors in our predictors, the accuracy of the results depends on the researchers decisions in so many levels (e.g. cut off rates, number of factors …).
3- EFA is very helpful in developing a theory rather than providing meanings to data set or making sense of factors. It's up to the researchers to provide meaning to these new structures from EFA.

Ref:

[1] DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. Practical Assessment, Research & Evaluation, 14(20), 1-11.