Final Report, STAT 528

Yuan Wang, Libin Liang, Songqian Chen

March 14, 2017

Contents

1. Summary

It is identified that staff perception of school-wide positive behavior support (SWPBS) is a major barrier to the SWPBS implementation in many schools. To help teams better understand the staff perceptions, a survey tool called Staff Perceptions of Behavior & Discipline (SPBD) was developed. The survey consists of 23 core items, 19 supplementary items and 3 open-ended items. In your original paper [5] and two previous consulting appointments, factor analysis and hierarchical linear models (HLM) have been used to analyze the survey data. In this report, we used the new data from fall 2014 to fall 2016, including response of 5362 staff members from 147 schools.

A PCA-based factor analysis was performed to study the internal validity of the SPBD. 5 factors were identified, which supports the findings from the original paper. Insight and a general profile of staff perception of each factor were also provided according to the analysis results. On the other hand, a hierarchical linear model was used to investigate the effect of some selected school/staff characteristics on staff support level of SWPBS. It is suggested that school level (SL), self-rate of the communication of the school (COMMUN), level of understanding of SWPBS (LOU) and support or commitment to the SWPBS (SOC) are significantly associated with staff support level of SWPBS.

2. Scientific Background

The schoolwide positive behavior support (SWPBS) is an evidence-based approach for restructuring schoolwide discipline through a multi-tiered framework of social, emotional, and behavioral supports [4], to effectively and efficiently support student(and staff) behavior. It helps establish the social culture and behavioral supports needed for students in a school to achieve both social and academic success.

Many schools in the U.S. have been transitioning to SWPBS from a traditional model of school discipline. However, some schools have difficulty in the SWPBS implementation. It is identified that staff perception of SWPBS is a major barrier to the SWPBS implementation in many schools. Hence in your paper, a catalyst survey tool called Staff Perceptions of Behavior & Discipline (SPBD) was developed to help teams understand the staff perceptions in their school [5].

In the reports of previous two consulting appointments (December 2015 and March 2016), factor analysis and linear mixed effects model have been used to analyze the survey data. Since more data have been obtained after that, you are interested in continuing the analysis using the new data set. Your research goal and interested methods are very similar to last time. Your current interests focus more on factor analysis, to study the internal validity of the SPBD. Besides that, a hierarchical linear model can be considered to understand which aspects of the school/staff correspond with differences in staff overall support level of SWPBS.

3. Research Methods

We use the SPBD survey data obtained from fall 2014 to fall 2016, including response of 5362 staff members from 147 schools. The data we use is in the *Coded Data Core SPBD Items* spreadsheet (which is different from the raw data). In the SPBD survey, there are 23 core items, and each item is one description or opinion developed to measure staff perception of SWPBS. For 17 of the 23 core items, there are four response choices: *strongly disagree*, *disagree*, *agree* and *strongly agree*. For the other 6 core items, besides the four response choices mentioned above, there is one more response choice which is phrased like “I feel that I do not know my colleagues well enough to answer this question”, “I don’t know”, etc. These “I don’t know”-type answers are viewed as neutral attitudes in our analysis. Hence, to quantify answers of the core items, 4-point scale scores with a neutral option are used as below:

Response Options	Scores
strongly disagree	1
disagree	2
"I don't know"-type answers	2.5
agree	3
strongly agree	4

Table 1: Response options and the corresponding scores

Among the 23 core items, 9 items are phrased positively and 14 items are phrased negatively. To make all the answers consistent, the 14 negatively phrased items are re-scaled to have 1 represent unsupportive attitudes to SWPBS and 4 represent supportive attitudes to SWPBS. The data of 23 core items provides individual-level information and is what we will focus on in chapter **6.2**.

Besides 23 core items, 19 supplementary items are designed to collect more information, both individual-level and school-level. School-level supplementary items include state, school level, self-reported level of SWPBS implementation, years of implementation, etc.; individual-level supplementary items include respondents' role at school, years of experience, hours of professional development in SWPBS, level of understanding of the concepts and procedures of SWPBS, etc.

Based on our research interests, we will select some of the supplementary items to analyze in chapter **6.3**.

Also, three open-ended items are included at the end of the survey, which is supposed to collect information about staff needs, concerns and existing capacities for SWPBS. The three items are:

- When you think about schoolwide positive behavior supports, what concerns do you have? Please be frank and answer in complete sentences. (*assess concerns*)
- When it comes to behavior and discipline, what is working well in this school? (*assess existing capacities*)
- What is needed to make it better? (*assess needs*)

We will briefly discuss the open-ended items in our final comments, chapter **7.2.2**. In the main body of this report, only core items and supplementary items will be considered.

4. Statistical Questions

The statistical questions include:

- To study the validity and consistency of SPBD, how might the survey items be best grouped into subscales?
- What are the associations between school/staff characteristics and staff support level of SWPBS?
- How to deal with missing data and "I don't know"-type answers in the survey?

To answer the first question, we will conduct a factor analysis in chapter **6.2**. To answer the second question, we will perform a hierarchical linear modeling in chapter **6.3**. For the third question, as mentioned in chapter **3**, "I don't know"-type answers will be regarded as neutral answers and transformed into median of the 1-4 scale (2.5). Missing data will be discussed in chapter **6**.

5. The Statistical Methods

For the questions raised in chapter **4**, we utilized two methods following your original paper [5]: PCA-based factor analysis and hierarchical linear model. Brief introductions are given below:

PCA-based factor analysis is actually not a kind of factor analysis since it is totally based on principal component analysis (PCA), which is theoretically different from factor analysis. It is an extension of PCA by performing a rotation on the loadings from the PCA so as to enforce some sparsity into the components and thus to improve the interpretability. In our further discussion in chapter **6.2**, we would conduct the varimax rotation on the loadings, identify the factors after rotation, and provide some insights by digging deeper into the factors/components.

Hierarchical linear model (HLM) is an extension of the ordinary linear model. It is designed to deal with the situation where there are variables from different levels because HLM can take care of within and between group variation simultaneously and make our inference more reliable. In the following discussion in chapter **6.3**, we try to investigate the relationship between some key variables and the support level of SWPBS. The interested variables are from different levels, hence the HLM is needed in this case.

6. Results and Interpretations

6.1 Descriptive Analysis

In this part, we provide the descriptive analysis of 23 core items and selected supplementary items. Barplots of variables mentioned in chapter 6.1.1 and 6.1.2 can be found in **Appendix A**.

6.1.1 School-Level Supplementary Items

For data from the *Coded Data Core SPBD Items* spreadsheet, some data are missing for unknown reasons. For the supplementary items, the missing rates are high. This survey includes respondents from 147 schools. 78 of the 147 schools missed the state information. Among the rest 69 schools, 33 schools are in Washington, 13 in Oregon, 5 in California, 5 in Florida, 5 in North Carolina, 3 in Virginia, 2 in Rhode Island, 1 in Massachusetts, 1 in Minnesota and 1 in Pennsylvania.

76 of the 147 schools missed the school level information. For the rest 71 schools, there are 41 elementary schools, 18 middle schools, 6 high schools, 3 preschools, 1 treatment center, 1 K-8 school and 1 alternative K-12 school.

97 of the 147 schools missed the self-reported level of implementation and years of implementation. For the rest 50 schools, 20 reported partial implementation, 15 reported planning implementation, 14 reported full implementation, and 1 reported unknown. Still for these 50 schools, 22 reported 1 year of implementation, 12 reported 2 years, 3 reported 3 years, 4 reported 4 years, 5 reported 5 years, 1 reported 7 years, 2 reported 10 years and 1 reported unknown.

6.1.2 Individual-Level Supplementary Items

There are 5362 respondents in our data. For staff's role at school, 65.8% of the respondents are certificated teachers, 18.1% are classified staff (e.g., office staff, kitchen staff, security). 14.5% of the respondents are in roles like administrator, certificated support personnel, etc. People who work as a certificated teacher or classified staff but have additional roles, are also included in the 14.5% group. Only 1.6% of the respondents missed the information of their roles.

For years of experience, 13.6% have 0-1 years' experience, 16.1% have 2-3 years, 13.8% have 4-6 years, 14.0% have 7-10 years, 15.1% have 11-15 years, 11.2% have 16-20 years, 12.9% have more than 20 years and 3.2% missed the information of years of experience.

Last, for the understanding level of concepts and procedures of positive behavior supports, 56.3% have basic understanding, 23.2% have high understanding, 17.9% have limited understanding, 2.3% do not know what it is and 0.3% are missing.

For the analysis in chapter **6.3**, we will select some of the supplementary items and remove observations with missing values in those items, which is a large amount of data and is a limitation of the analysis.

6.1.3 Core Items

For core items, the missing rates are much lower. Item 14 has the highest missing rate (10.3%) for unknown reason. Besides that, all the other 22 core items have a missing rate lower than 0.9%. Hence, we remove observations with any missing values in the 23 core items, and there are 4589 observations left for analysis.

As we mentioned in chapter 3, the answers to the core items are 4-point scale scores. First, we give the boxplot of each core item as below:

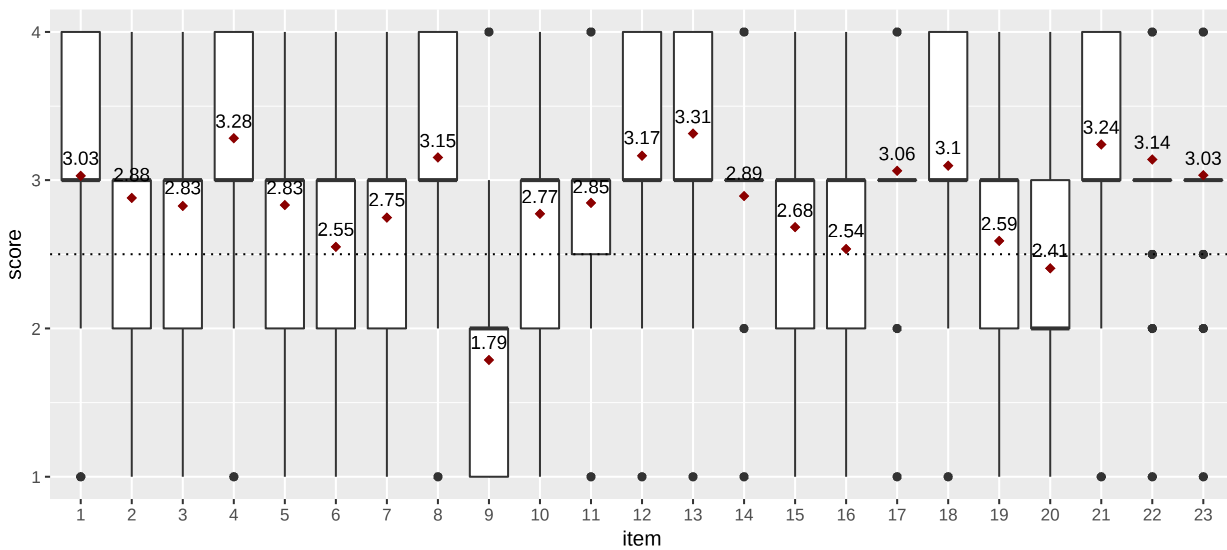

Figure 1: Boxplot of core items

In each box, the red point and the number in text indicates the mean value of answers to each item. Black points are outliers. The lower edge and upper edge of the box indicate the 25% quartile and 75% quartile of the item answers. For all the boxes, median is overlapped with at least one quartile. Besides item 9 and item 20, the median of all the other core items is 3.

Also, we give the summary of means as below:

Mean	Item
$[3, 3.5)$	1, 4, 8, 12, 13, 17, 18, 21, 22, 23
$[2.5, 3)$	2, 3, 5, 6, 7, 10, 11, 14, 15, 16, 19
$(1.5, 2.5)$	9, 20

Table 2: Mean of the answers to each core item

There are 10 items having a mean larger than 3, indicating that on average, people have a supportive attitude towards SWPBS in those items. Among them, the three largest means are from item 13(mean 3.31), item 4(mean 3.28) and item 21(mean 3.24).

There are 11 items have a mean smaller than 3 but larger than 2.5, indicating that on average, people have a neutral to slightly supportive attitude towards SWPBS in those items.

Only 2 items have a mean smaller than 2.5. They are item 9(mean 1.79) and item 20(mean 2.41). Hence in the two items, especially in item 9, people have an unsupportive attitude on average towards SWPBS.

Item 9 has the smallest mean among all 23 core items. It reflects that on average, people largely support the negatively phrased opinion "the students at this school need to be held more responsible for their own behavior".

Next, we give the heatmap of correlations between core items, along with the result of a hierarchical clustering:

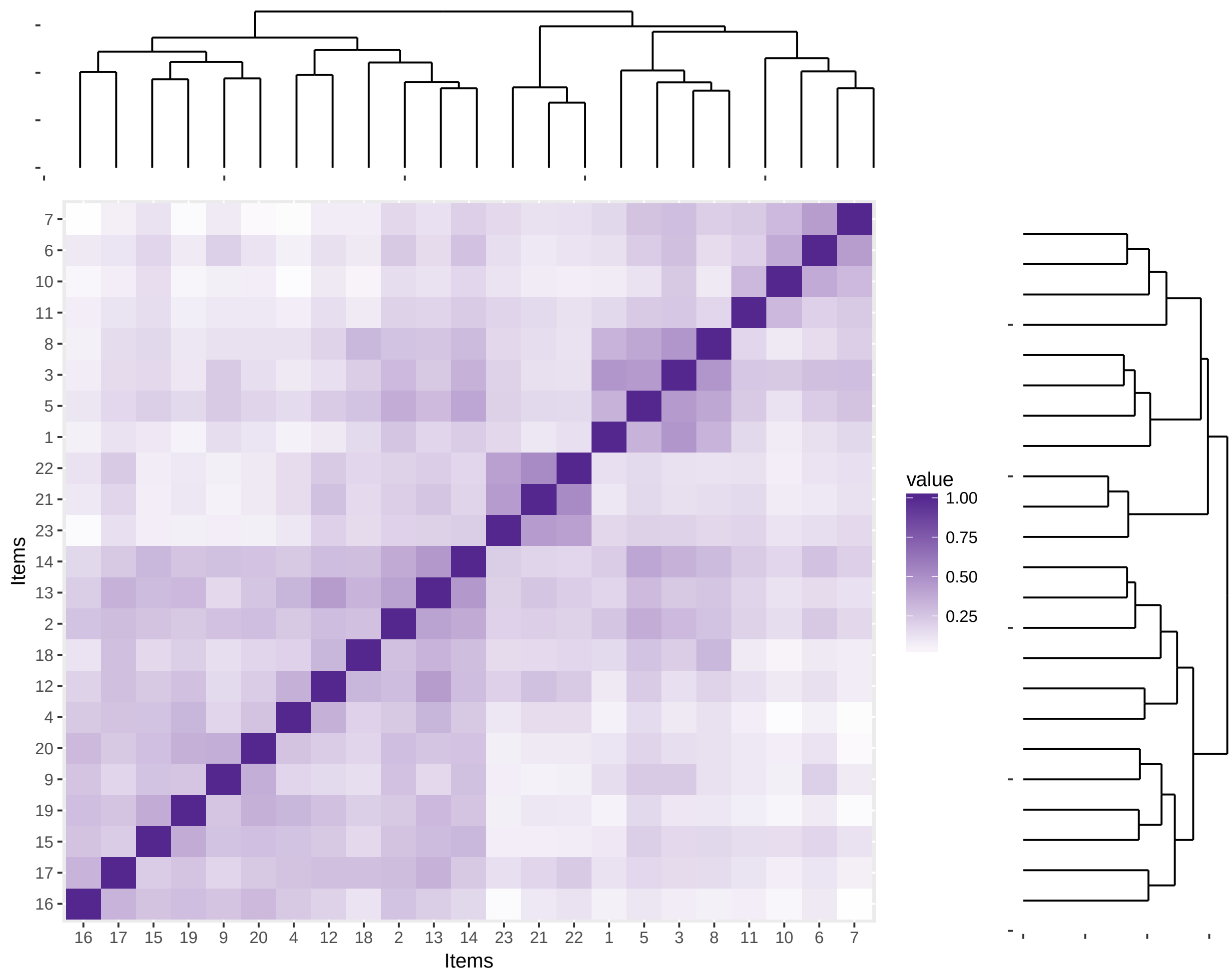

Figure 2: Heatmap of correlation between core items, with hierachical clustering

In the heatmap, a darker color means a larger correlation. From the plot, we can find some groups, within which the items have relatively larger correlations (darker colors). (11, 10, 6, 7), (1, 5, 3, 8), (23, 21, 22) and (16, 17, 15, 19, 9, 20, 4, 12, 18, 2, 13, 14) are four groups we can identify. Items within the same group are more correlated to each other.

The dendrogram of the hierarchical clustering is also shown in the plot. The hierarchical clustering is based on the correlation between items. Items that are more correlated are more likely to be clustered into the same group. The result of hierarchical clustering is consistent with the heatmap.

6.2 PCA-based Factor Analysis

6.2.1 Suitability of Factor Analysis

Following your original paper [5], we assess the suitability of the data for factor analysis via Cronbach's alpha coefficient and Kaiser-Meyer-Oklin (KMO) index.

Cronbach's alpha coefficient is a measurement of internal consistency with values scaling from 0 to 1. Generally, a set of items with Cronbach's alpha coefficient larger than .8 would be considered to be closely related to each other as a group. With your new SPBD data set, we have 0.84 as the overall Cronbach's alpha coefficient, indicating a good internal consistency within the data and that the factor analysis would be useful to find the latent correlations behind the items.

The Kaiser-Meyer-Olkin (KMO) index was proposed by Kaiser [1] in 1970 and modified in 1974 [2], which can also tell how suitable the data is for factor analysis by measuring sampling adequacy of the data set. The value of KMO index is between 0 and 1. Generally, a KMO index larger than .8 would be a great indicator that the data is suitable for factor analysis.

6.2.2 Principal Component Analysis and Varimax Rotation

With good indications of the suitability for factor analysis, we conducted a principal component analysis (PCA), followed by the varimax rotation on the loadings we derived from the PCA. The varimax rotation is an orthogonal rotation on the loadings so as to increase the interpretability of the principal components, which we will name as factors in the further analysis. More specifically, when the original PCA have many non-zero loadings for factors, we can use a varimax rotation to rotate the axis. By doing that, some factors would have loadings close to 1, while others would have loadings close to 0, and interpretability can be increased in that way.

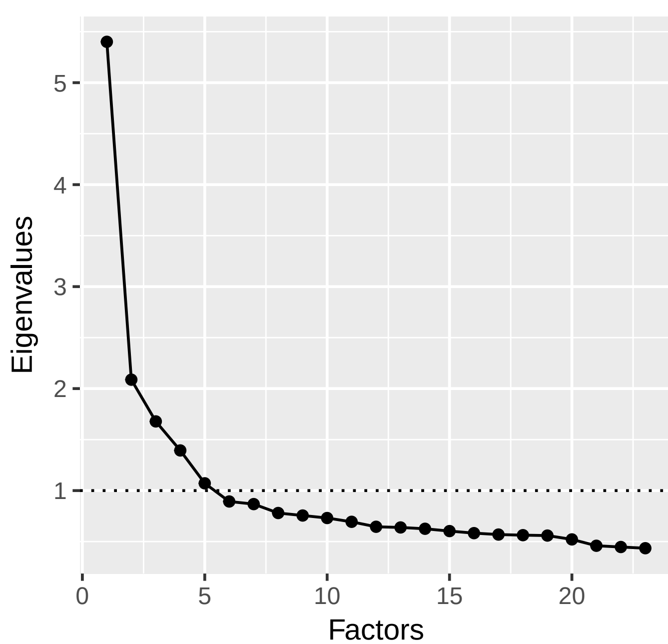

Figure 3: Values of eigenvalues with respect to factors in order

Before moving on to the further analysis, the number of factors remained in the model could be determined from the scree plot in Figure 3. It displays the eigenvalues versus the number of the factors. Two observations could be seen from Figure 3. First, there are 5 eigenvalues larger than 1, which means that they have eigenvalues larger than the average value, since the summation of all eigenvalues equals the total numbers of factors we have. Second, the slightly decreasing trend after the fifth eigenvalue indicates that most variability are explained by the first five factors. Therefore, we remain 5 factors for our further analysis.

The result of PCA followed by varimax rotation with 5 remained factors is shown in the Table 3, where the threshold for loadings to display is 0.4. The first five factors explain 11.5%, 11.1%, 10.1%, 9.2% and 8.6% of variability inside the data respectively, and 50.6% cumulatively. In Table 3, the ID for each item is the same as the original item order in the survey.

			Factor Loadings				
ID	Items	Means (SD)	1	2	3	4	5
13*	Schoolwide behaviour supports may work in other schools, but I doubt it will work in ours.	3.29 (0.61)	0.70				
12*	I don't have time to teach the schoolwide behavioural expectations.	3.14 (0.65)	0.68				
18*	I resent being asked to do one more thing.	3.10 (0.67)	0.60				
4*	We should not have to teach students how to behave at school.	3.27 (0.75)	0.54				
17*	I feel that rewarding students is the same as bribing them.	3.05 (0.64)	0.45				
14*	Behaviour plans do not work well in our school.	2.89 (0.75)	0.43				
19*	If students are not disciplined at home, they are not likely to accept any discipline at school.	2.57 (0.83)	0.42		0.53		
3	The climate at this school is positive.	2.83 (0.78)		0.74			
1	I have trust in my administrator's ability to lead us through change.	3.03 (0.90)		0.71			
5	I believe our school has (or will have) the necessary resources to support schoolwide positive behaviour support.	2.82 (0.76)		0.63			
8	Overall, I am satisfied with my job.	3.15 (0.70)		0.68			
9*	The students at this school need to be held more responsible for their own behaviour.	1.78 (0.69)			0.69		
20*	When problem behaviours occur, we need to get tougher.	2.39 (0.74)			0.68		
16*	I believe we should reserve rewards for students exceeding expectations, not simply for meeting them.	2.51 (0.84)			0.64		
15*	Parents in the community don't seem to care about how their children behave at school.	2.68 (0.73)			0.47		
6*	I suspect that my colleagues will not (or are not) consistently implementing the agreed upon schoolwide behaviour plan.	2.55 (0.76)				0.71	
10*	The staff at this school tends to resist change with concerns such as we don't do it that way here.	2.77 (0.71)				0.75	
7	My colleagues and I share a common philosophy for behaviour and discipline.	2.73 (0.65)				0.66	
11	This school has successfully implemented change efforts in the past.	2.82 (0.58)				0.52	
21	Currently, I teach the agreed upon schoolwide behaviour expectations to students.	3.21 (0.60)					0.80
22	Currently, I acknowledge/reward students for meeting the agreed upon schoolwide behaviour expectations.	3.11 (0.59)					0.80
23	Currently, I apply the agreed upon schoolwide disciplinary consequences.	3.01 (0.55)					0.72
2	Schoolwide behaviour support is likely to be yet another fad that comes and goes in this school.	2.86 (0.79)					

Table 3: Questions with their corresponding means, standard deviations and loadings for the five factors. Cells colored with green have the five largest mean values, while cells colored with red have the five smallest mean values.

We could find that items in above factors are similar to the item groups we identified in the correlation-based heatmap (Figure 2). In fact, the PCA-based factor analysis and the correlation-based hierarchical clustering are theoretically different. The PCA-based factor analysis makes a rotation so as to enforce sparsity on loadings, while the hierarchical clustering in Figure 2 is totally based on the correlation between items.

However, it's not surprising that the two results are similar, since PCA-based factor analysis finds latent structures in the data, and items with high correlations are more likely to be from the same structure.

The results from Table 3 are quite similar to your previous study [5], although the order of the factors might slightly change. Therefore we would stick with the interpretation of each factor that proposed in your study. However, this time we would further interpret the results by digging more into staff's attitude to each factor.

6.2.3 Interpretation

As we could see from Table 3, we colored the five largest mean values with green while the five smallest mean values with light red. After we colored these cells out, an interesting observation would be that most items in factor 1 are covered in green, meaning that staff tend to support SWPBS in their answers to those items. While most items in factor 3 are in red, indicating that staff seem to hold an unsupportive attitude to SWPBS in those items. Summary information in Table 4 and Figure 4 would give us a more comprehensive view of these observations.

	Factors	Means (SD)	Cronbach's α	Item ID
1	Teaching and acknowledging behavioral expectations	3.05 (0.74)	0.74	(4,12,13,14,17,18,19)
2	Sustainable change in school climate	2.96 (0.80)	0.72	(1,3,5,8)
3	Philosophical views of behavior and discipline	2.39 (0.83)	0.68	(9,15,16,19,20)
4	Cohesiveness and openness to change	2.72 (0.69)	0.64	(6,7,10,11)
5	Implementation integrity	3.11 (0.59)	0.73	(21,22,23)
-	Total	2.84 (0.79)	0.84	-

Table 4: Summaries for each factors

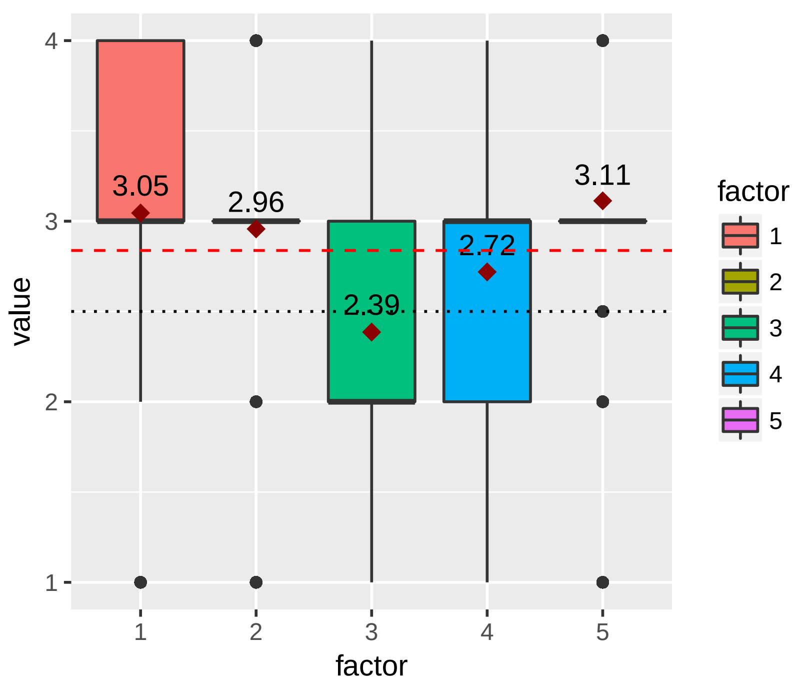

Figure 4: Boxplot for each factor. The red diamond points in the middle represent mean values for each factor. The red dash line shows the overall mean among all factors, and the black dash line shows 2.5.

From Table 4, we would see that cronbach'a alpha for all factors are larger than 0.6, which reveals great internal consistency within each factor. In Figure 4, we add two lines to the box plot. The first line (dash black line) represents the value of 2.5, which is the median of the scale 1-4. The second line shows the overall mean if we get all factors together, which is 2.84 for our data. We can find that factor 1,2 and 5 have an average response larger than both 2.84 and 2.5, meaning that staff tend to have positive attitudes to SWPBS in these factors. On the opposite, factor 3 has a mean value smaller than both 2.84 and 2.5, showing that staff tend to have a negative attitude to SWPBS in factor 3. Finally, the mean value of factor 4 is between the two lines, hence staff tend to have a neutral attitude to SWPBS in factor 4.

Based on the above response analysis for each factor, we could further interpret these factors by combining staff attitudes with the items these factors have:

- *Teaching and acknowledging behavioral expectations:* Most responses of items in this factor would tend to be "supportive" for SWPBS, hence this factor actually reflects a supportive attitude to teach and acknowledge the schoolwide behavioral expectations.
- *Sustainable change in school climate:* This factor reflects that staff are satisfied with the schools' climate, ability and resources.
- *Philosophical views of behavior and discipline:* This factor has a mean smaller than both 2.5 and the overall mean, indicating a strong "unsupportive" attitude to SWPBS in answers of items in this factor. With a deeper look into the items, they are basically asking for staff's opinion on the philosophy behind the behaviors support system. Therefore, we could see that to some degree, staff disagree with the philosophy behind the positive behavior support principal, although they would still accept the schoolwide behavioral support in their work based on the analysis of factor 1.
- *Cohesiveness and openness to change:* This factor contains most items related to colleagues. According to the result, staff tend to have a neutral response for these items. This is reasonable since it is always hard to rate others' perception to a certain concept, hence people are likely to give neutral responses for this kind of questions.
- *Implementation integrity:* The items in the last factor are straightforward and based on the "agree"-likely response, most of the staff have good integrity with implementation. However, we should be more careful when interpreting this factor since these items are subjective and is a form of self-rating, where people are more likely to give a positive response for themselves. It would be suggested to take into consideration the information we have from other factors when interpreting this factor.

To sum up, the survey gives us a good profile of the general staff perceptions of the schoolwide positive behavior support. On one hand, staff members accept the schoolwide implementation of positive behavior support principal, and regard themselves having good integrity when implementing this discipline. While on the other hand, they cannot fully agree with the philosophy underlying the positive behavior support, and also have much uncertainty about how consistent their colleagues would implement the support.

6.2.4 Diagnosis

In this part, we use the raw data instead of the data from the *Coded Data Core SPBD Items* spreadsheet. As you have mentioned before, it would be suggested to check whether the demographic information have influence on the factor analysis. Since for the raw data, only the "states" and "school level" are complete (that is, there is generally no missing data for these two columns), we mainly conduct the diagnosis on these two demographic information. More specifically, based on our previous analysis, we would project the data onto the first two factors and try to see whether the data points for different states or different school levels will have different clusters.

To begin with, for different states, we have 2449 observations for WA, 1037 for CA, 667 for OR, 365 for SC, 249 for FL, 223 for NC, 91 for VA, 69 for MA, 51 for MN, 47 for RI, 43 for PA, 14 for BC, 12 for WY and 45 are missing. For simplicity, we remain WA, OR and CA that occupy a large proportion of observations while grouping observations from other states together. By this mean, we would have the scatter plot of the data onto the first factors with different states shown in Figure 5. As we could see from the Figure 5, data points from different states seem to have similar mean and variance. There is no obvious evidence that different states tend to have different clusters of data points.

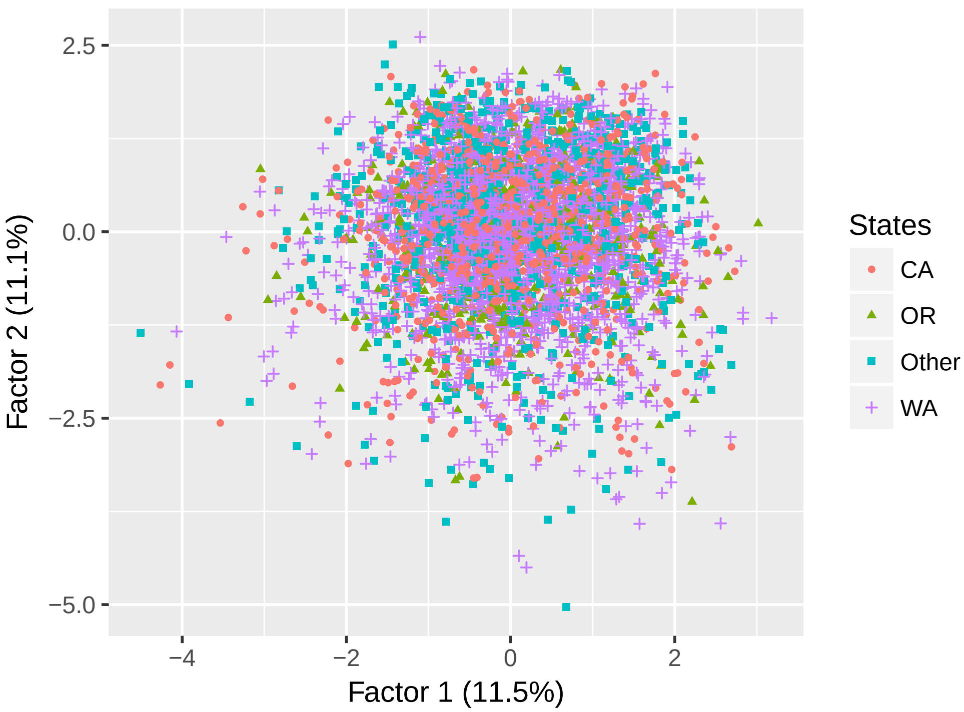

Figure 5: Raw data onto the first two factors by different states

Similar diagnosis was conducted on the school level. For school levels, we have 2692 observations for elementary school, 1419 observations for middle school, 856 for high school and 359 for others. Scatter plot of the projected data was drawn in Figure 6. It could be observed that there is no obvious difference between four clusters of points for these four categories of schools.

In conclusion, based on the diagnosis above, the two demographic information “states” and “school level” do not have significant influence on the PCA analysis we conducted.

6.3 Hierarchical Linear Model Analysis

In this part, we try to apply the hierarchical linear model (HLM) to investigate the relationship between the support level of SWPBS and some interesting variables.

6.3.1 Variables in the Analysis

The responses in the analysis is the total score of the 23 core items in the SPBS survey, which can reflect the overall support level of SWPBS.

5 interested variables will be involved in the following analysis. Some of them are school-level variables, such as school level (SL) and self-reported level of implementation (IL) while some of them are staff-level variables such as self-rate of the communication of the school (COMMUN), level of understanding of SWPBS (LOU) and support or commitment to the SWPBS (SOC).

For school level (SL), there are 5 levels and they are kindergarten (0), preschool (1), elementary school (2), middle school (3) and high school (4). For self-reported level of implementation (IL), there are 3 levels and

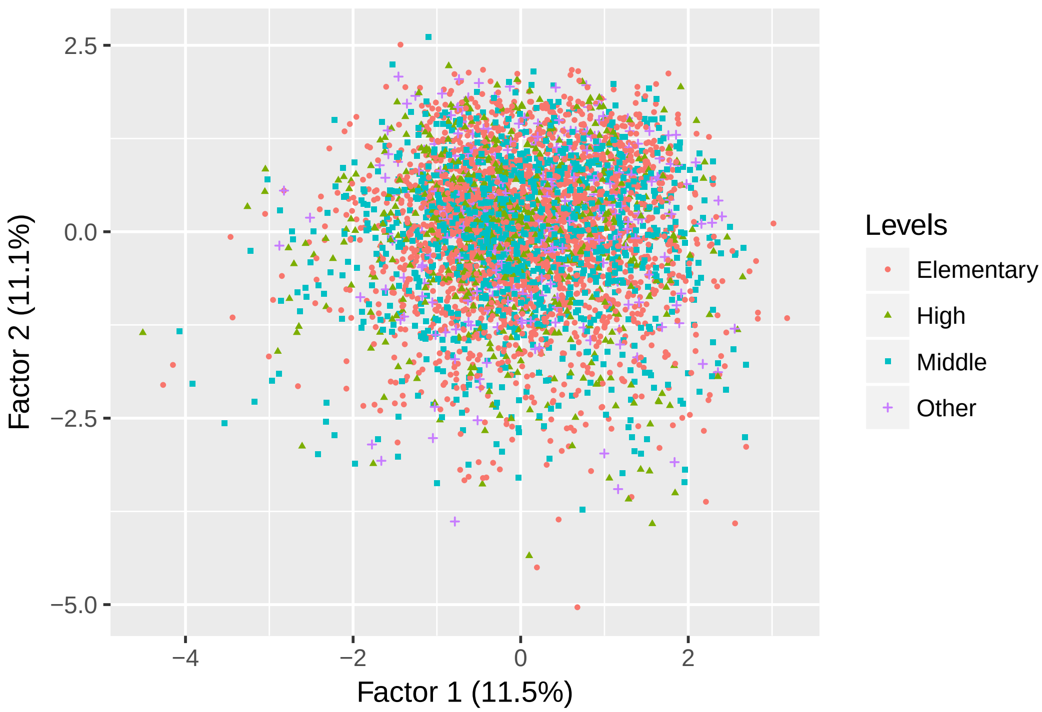

Figure 6: Data onto the first two factors for different states

they are planning (0), partial (1) and full (2). For self-rate of the communication of the school (COMMUN), there are 4 levels and they are Poor (0), Need improvement (1), Adequate (2) and Good (3). For level of understanding of SWPBS (LOU), there are 4 levels and they are Unfamiliar (0), Limited (1), Basic (2) and High (3). For support or commitment to the SWPBS (SOC), there are 5 levels and they are strongly disagree (0), disagree (1), unfamiliar (2), agree (3) and strongly agree (4).

6.3.2 Model Setting

As we mentioned above, the interested variables belong to different levels. If we fit the data with an ordinary linear model, then we will ignore the presence of group differences and variation. Considering that, we set up the hierarchical linear model below to investigate the relationship between support level of SWPBS and the covariates:

$$SPBD_{ij} = \beta_{0j} + \beta_{1j}COMMUN_{ij} + \beta_{2j}LOU_{ij} + \beta_{3j}SOC_{ij} + \epsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}SL_j + \gamma_{02}IL_j + U_{0j}$$

$$\beta_{kj} = \gamma_{k0} + U_{kj} \qquad k = 1, 2, 3$$

Where j corresponds to school-level while i corresponds to staff-level. $\epsilon_{ij} \sim N(0, \sigma^2)$, $U_{kj} \sim N(0, \tau_{kk}^2)$ for $k = 0, 1, 2, 3$.

Combining the models above, we obtain a mixed effects model as below:

$$\begin{aligned} SPBD_{ij} &= \gamma_{00} + \gamma_{01}SL_j + \gamma_{02}IL_j + \gamma_{10}COMMUN_{ij} + \gamma_{20}LOU_{ij} + \gamma_{30}SOC_{ij} \\ &+ U_{0j} + U_{1j}COMMUN_{ij} + U_{2j}LOU_{ij} + U_{3j}SOC_{ij} + \epsilon_{ij} \end{aligned}$$

With above mixed effects model, we can fit the HLM model with **lme** function in R.

6.3.3 Diagnosis

As the data of interested variables are not complete for all surveyed staff, we extracted the data points with complete data of the interested variables. There are 1353 data points with complete data of the interested variables, and they come from 49 different schools. We fit the model in **6.3.2** with the 1353 data points. For diagnostics, we drew the scatter plot of residuals versus fitted values and QQ-plots of random effects in different levels as below:

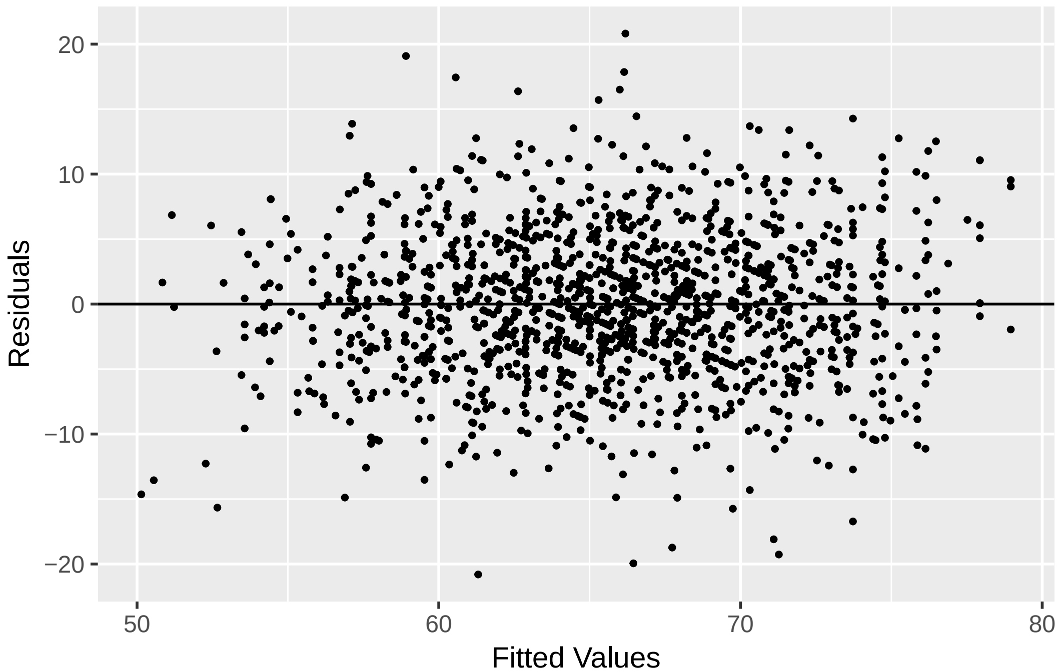

Figure 7: Residuals versus fitted value of the HLM

From the scatter plots above, we can see there is no obvious mean trend of residuals and the magnitude of the residuals do not have a significant change as the fitted value increases. Then we plot QQ-plot of the residuals and also of the random effects in the second level.

Based on the QQ-plots in Figure 8 and 9, there is no strong evidence supporting the violation of normality of residuals and second level random effects. Based on that, the HLM model here is plausible.

6.3.4 Results and Analysis

We have the fitting results of HLM as below:

From the results above, we can find that the school level (SL), self-rate of the communication of the school (COMMUN), level of understanding of SWPBS (LOU) and support or commitment to the SWPBS(SOC) are statistically significant in the hierarchical linear model. This suggests that school level (SL), self-rate of the communication of the school (COMMUN), level of understanding of SWPBS (LOU) and support or commitment to the SWPBS (SOC) have an effect on the overall support level of SWPBS.

Moreover, from the signs of the coefficients of the variables, we know that school level have a negative effect on the support level of SWPBS, that is, the higher the school level is, the lower the support level of SWPBS will be. This reflects that the staff in higher level schools tend to pay less attention on students' behaviors and disciplines and have negative perceptions of behaviors and disciplines. Considering that, the situation of SWPBS is worse in the higher level schools. Also, the communication of the school (COMMUN),

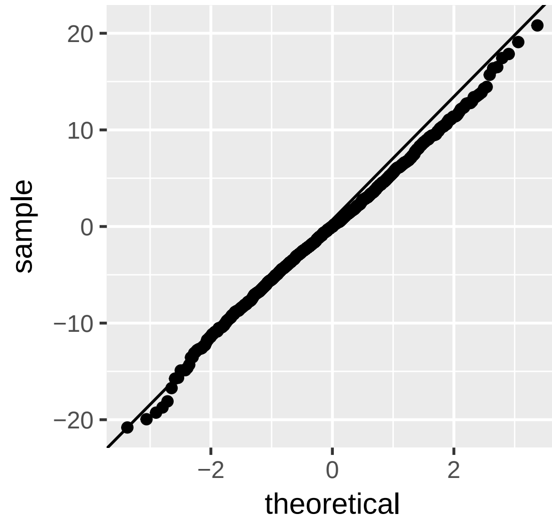

Figure 8: Norm QQ-plot of residuals

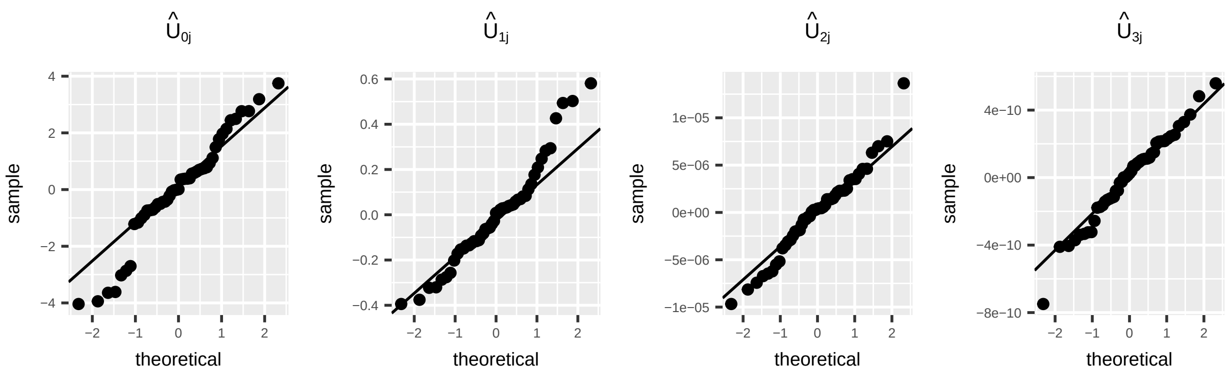

Figure 9: Norm QQ-plot of each second level random effect

Fixed Effect	Estimated Value	Standard Deviation	P value
Intercept(γ_{00})	55.91	1.54	<0.001
SL(γ_{01})	-2.12	0.48	<0.001
IL(γ_{02})	0.29	0.5	0.564
COMMUN(γ_{10})	3.42	0.21	<0.001
LOU(γ_{20})	1.43	0.24	<0.001
SOC(γ_{30})	1.77	0.19	<0.001
Random Effect	**Variance**		
Intercept(τ_{00}^2)	4.98		
COMMUN(τ_{11}^2)	0.25		
LOU(τ_{22}^2)	<0.01		
SOC(τ_{33}^2)	<0.01		
Residuals(σ^2)	33.33		

Table 5: Results for fixed effects and random effects for the hierarchical linear model

level of understanding of SWPBS (LOU) and the support or commitment to the SWPBS (SOC) all have a positive effect on the perceptions of behaviors and disciplines. This suggests that, to enhance the support level of SWPBS, we can improve the communication in the school, let the staff feel the improvement of the communication in their school, let the staff understand SWPBS better and talk with them about the benefits of SWPBS to make them give more support and commitment to SWPBS.

7. Final Comments

7.1 Further Discussions

Finally, there are some further discussions on the possible issues:

- Two lines were drawn in Figure 4 in chapter **6.2.3**, where one of them is the overall mean value of all factors. Our initial motivation for this line is to obtain the relative supportiveness of each factor compared to the overall rating level. However, as we mentioned in chapter **3**, the 14 negative items have been re-scaled so that their response values are actually in the opposite direction, which might let the overall mean cannot fully represent the average rating level. This is not a big issue because now the overall mean actually represents the average support level of SWPBS. Thus the overall mean value could still be utilized to compare the supportiveness of each factor to the average value.
- We also notice that there are problems of missing data in both the core items and the supplementary items. For the supplementary items, the problem is severe and only about one fourth of respondents have the complete responses in those items. To let our inference be more reliable, you should also ensure the completeness of the supplementary items. For the core items, the problem is much less serious and there are 4589 out of 5362 respondents have the complete responses in those items. In the analysis above, we just ignored the respondents with missing data while we may lose some information by doing that. Another promising solution is to impute the missing data by the EM algorithm.

7.2 Future Directions

7.2.1 HLM Analysis on the Factors from PCA

Based on the results of PCA-based factor analysis, the SPBD survey can be decomposed into five internal factors. And it can be imagined that the key variables we have mentioned in chapter **6.3** should have different effects on the five internal factors. Considering this, we can apply the HLM analysis to investigate the relationship between the key variables and each factor.

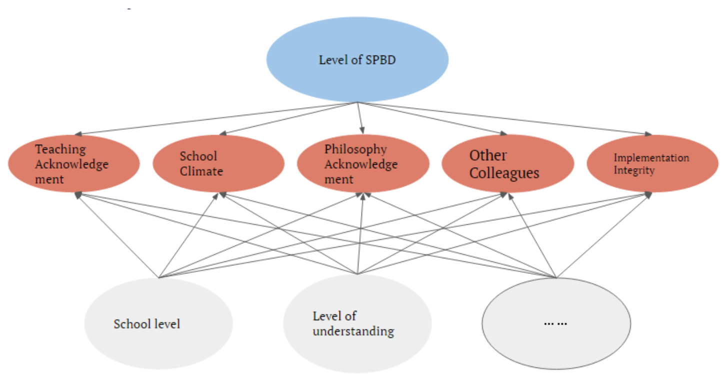

Figure 10: Structure of the relationship between the key variables and each factor

7.2.2 Concerns Extraction via Latent Directlet Allocation

As we dig deeper into the raw data set you provided later, we find that the three open-ended items are very informative. For example, in the item “When you think about schoolwide positive behavior supports, what concerns do you have”, many staff sincerely wrote down their concerns about the SWPBS, like:

- "I am concerned that all staff will not buy into the school wide supports, therefore **they will not be consistently implemented**, and students will receive mixed messages."
- "We only acknowledge the students that cause trouble. **What about the good students? The disruptive students should not be getting monetary awards for doing what they should be doing.** That is not teaching them to be responsible for their actions. Two weeks of appropriate behavior should not be rewarded."
- "I would like to have opportunities to receive behavior support **training**."

By reading many of the concerns, it seems that there exist some general topics that most of the staff are concerning about, like the consistency and some problems inside the rewarding system. To find out the latent topics behind their concerns, we tried to use a text-mining technique, latent dirichlet allocation (LDA), on the concerns data with 5 topics and 10 words each topic, then obtain the results as below.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
behavior	time	staff	student	student
student	need	concern	behavior	the
positive	staff	school	need	kid
support	PBI	consistency	teacher	get
reward	implement	implement	consequence	reward
need	support	support	school	don’t
think	train	teacher	parent	time
concern	program	need	support	class
will	think	will	classroom	teacher
work	manipulation	expectation	issue	behavior

Table 6: Topics result from latent dirichlet allocation with 5 topics

From the results of exploratory processing with LDA in Table 6, we could see that some keywords were found out by this text-mining technique (as we highlight in light green in the table). However, it is hard to interpret each topic because the parameters for the algorithm, like the number of topics, have not been optimized yet. For tuning the number of topics, we suggest to view the problem as a K-means clustering problem and use the technique to select the number of topics for this text-mining algorithm. There are many methods could be used to select the number of clusters in K-means, and we would like to recommend the one with GAP statistics raised by Tibshirani in 2001 [3]. We think this would be a promising direction for further studying open-ended items in the raw data set.

References

[1] Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401-415.

[2] Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.

[3] Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B* (Statistical Methodology), 63(2), 411-423.

[4] Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, 11(3), 133-144.

[5] Feuerborn, L. L., Tyre, A. D., & King, J. P. (2015). The Staff Perceptions of Behavior and Discipline Survey: A tool to help achieve systemic change through schoolwide positive behavior support. *Journal of Positive Behavior Interventions*, 17(2), 116-126.

Appendix

Appendix A: Barplot of Selected Supplementary Items

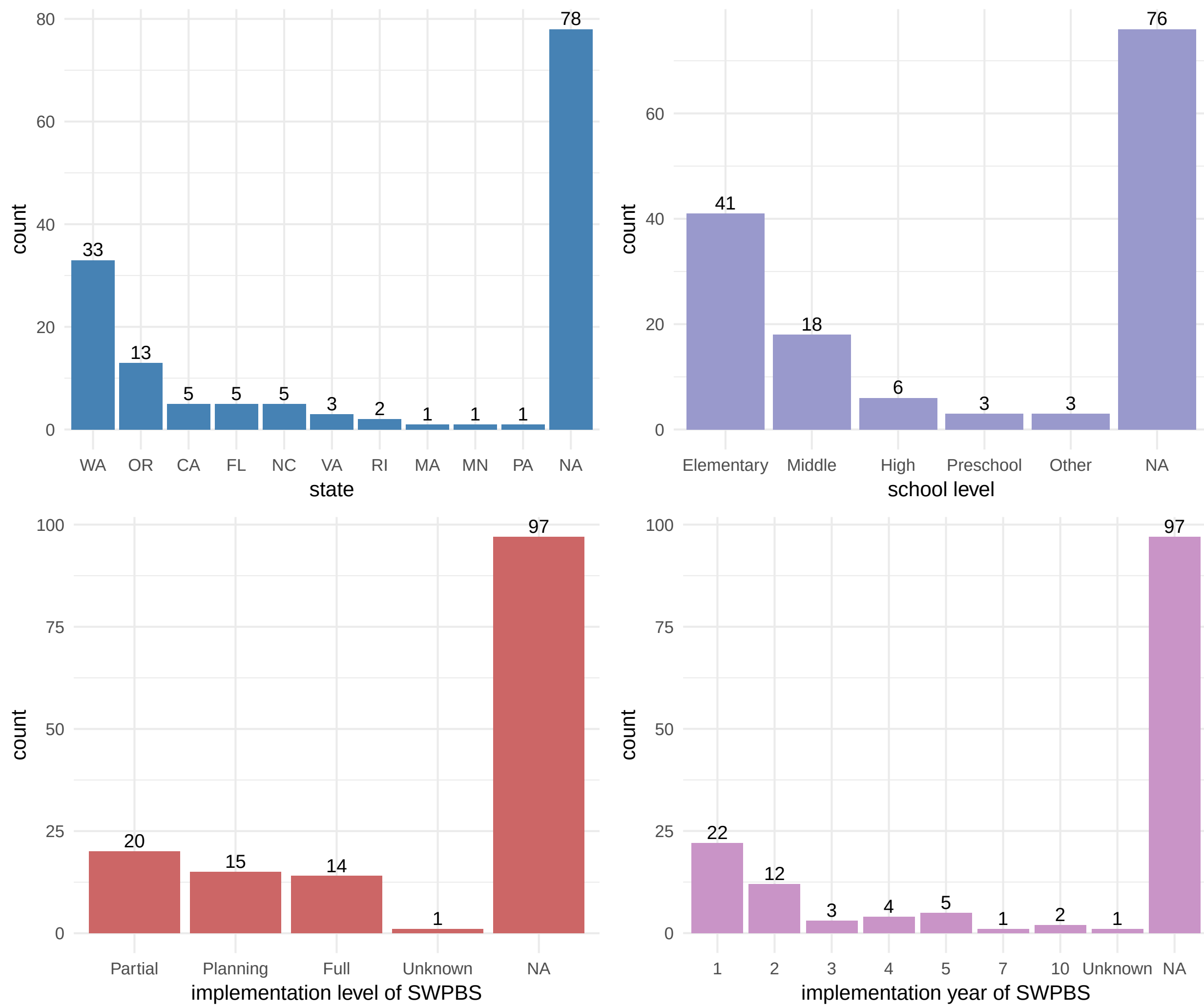

Figure 11: Barplot of selected school-level variables

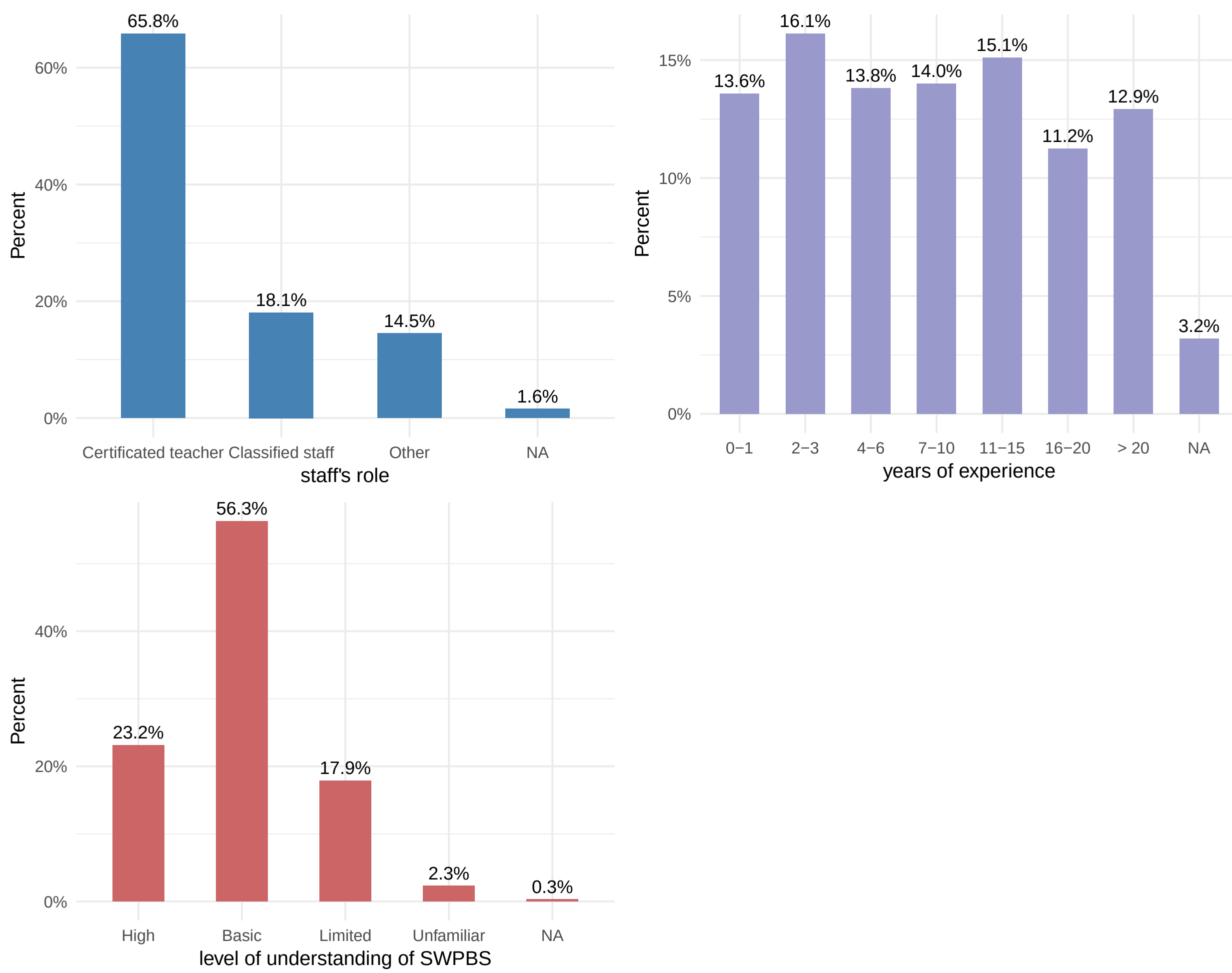

Figure 12: Barplot of selected individual-level variables

Appendix B: Code for Descriptive Analysis

```r
# Require data and files:
# @file: complete.csv

# not including code for plots
complete <- read.csv("complete.csv")
suplementary <- complete[, c(1:9, 12:15,36:37, 42:44)]
supp1 <- suplementary[, 1:8]
supp2 <- suplementary[, -(1:8)]

# school-level data
sup1 <- unique(supp1)
length(sup1$School.Name)
table(sup1$School.Name)

rownames(sup1) <- 1:nrow(sup1)
attach(sup1)
num41 <- sup1[School.Name == 41,]
num57 <- sup1[School.Name == 57,]

final.sup <- sup1[-c(36, 51),]
table(final.sup$School.Name)
length(final.sup$School.Name) # 147 schools

# School-level information
# 1. State
table(factor(final.sup$State))

# 2. School Level
table(factor(final.sup$School.Level))

# 3. Self-reported Level of Implementation
table(factor(final.sup$Self.reported.Level.of.
             Implementation..Planning..Partial..Full.))

# 4. Years of Implementation
table(factor(final.sup$Years.of.Implementation))

# Individual-level information
# 1. Role at school
table(supp2$Please.indicate.your.role.s..at.this.school.)
table(supp2$Please.indicate.your.role.s..at.this.school.)/5362

# 2. Years of experience
table(supp2$How.many.years.of.experience.do.you.have.in.your.current.role.)
table(supp2$How.many.years.of.experience.do.you.have.
      in.your.current.role.)/5362

# 3. Level of understanding of SWPBS
table(supp2$When.it.comes.to.the.concepts.and.procedures.
      of.positive.behavior.supports..my.level.of.understanding.is.)
table(supp2$When.it.comes.to.the.concepts.and.procedures.
      of.positive.behavior.supports..my.level.of.understanding.is.)/5362
```

Appendix C: Code for PCA-based Factor Analysis

Data Processing (R)

```
# Require data and files:
# @file: SPBD.csv

library(dplyr)
library(psych)
library(GPArotation)

SPBD <- read.csv("SPBD.csv") %>%
  dplyr::select(-starts_with("x")) %>%
  mutate_each(funs(as.character(.)), -(School.Name:Finish.Date))
questions <- names(SPBD)
questions <- questions[-(1:3)]
names(SPBD)[4:ncol(SPBD)] <- paste("col_", letters[1:(ncol(SPBD)-3)], sep="")


# normal case
SPBD[SPBD == "I feel that I do not know my colleagues well enough
     to answer this question."] <- 2.5
SPBD[SPBD == "I don't know."] <- 2.5
SPBD[SPBD == "Currently, my school does not have a common set of
     student expectations."] <- 2.5
SPBD[SPBD == "Currently, my school does not have a common set of
     consequences."] <- 2.5
SPBD[SPBD == "999"] <- NA
SPBD <- SPBD %>%
  mutate_each(funs(as.numeric(.)), -c(School.Name, ID, Finish.Date))

X <- SPBD %>% dplyr::select(-(School.Name:Finish.Date))
```

Model Construction (R)

```
# Require data and files:
# @data: X (from data processing part in Appendix C)

fit.varimax <- principal(X, nfactors = 5, rotate = "varimax")
fit.oblimin <- principal(X, nfactors = 5, rotate = "oblimin")
fit.promax <- principal(X, nfactors = 5, rotate = "promax")
fit.cluster <- principal(X, nfactors = 5, rotate = "cluster")

loading.2.matrix <- function(loading, nfactors) {
  return(matrix(loading, ncol=nfactors))
}

varimax.loading <- fit.varimax$loadings


## Cronbach's alpha coefficient and KMO index
cronbach <- psych::alpha(X)
c.alpha <- cronbach$total[1]

kmo <- KMO(X)$MSA
```

Appendix D: Code for Hierarchical Linear Model

Data Processing (Python)

```
# Require data and files:
# @file: SPBD.xlsx

import os
import csv
from openpyxl import load_workbook

wb=load_workbook(filename="SPBD.xlsx",read_only=True)
ws=wb['Coded Data All Schools']


dat=[]
for line in ws.rows:
    temp=[]
    for cell in line:
        temp.append(cell.value)
    dat.append(temp)

school_level=[]
implementation_level=[]
implementation_year=[]
level_of_understanding=[]
support_or_commitment=[]
development=[]
exp=[]
Commun=[]

for i in range(1,len(dat)):
    if dat[i][2] not in school_level:
        school_level.append(dat[i][2])
    if dat[i][6] not in implementation_level:
        implementation_level.append(dat[i][6])
    if dat[i][7] not in implementation_year:
        implementation_year.append(dat[i][7])
    if dat[i][36] not in level_of_understanding:
        level_of_understanding.append(dat[i][36])
    if dat[i][37] not in support_or_commitment:
        support_or_commitment.append(dat[i][37])
    if dat[i][42] not in development:
        development.append(dat[i][42])
    if dat[i][41] not in exp:
        exp.append(dat[i][41])
    if dat[i][43] not in Commun:
        Commun.append(dat[i][43])

w_school_level=[
    [school_level[6],school_level[9],school_level[4]],
    [school_level[3]],
    [school_level[0],school_level[5]],
    [school_level[2],school_level[7]],
```

```python
    [school_level[1],school_level[8]],
    [school_level[10]]
]
w_implementation_level=[
    [implementation_level[0],implementation_level[4],implementation_level[7]],
    [implementation_level[2],implementation_level[5]],
    [implementation_level[1],implementation_level[6]],
    [implementation_level[3],implementation_level[8]]
]
w_implementation_year=[implementation_year[3],implementation_year[8]]
w_level_of_understanding=[
    [level_of_understanding[3]],
    [level_of_understanding[1]],
    [level_of_understanding[2]],
    [level_of_understanding[0]],
    [level_of_understanding[4]]
]
w_support_or_commitment=[
    [support_or_commitment[3]], [support_or_commitment[4]], [support_or_commitment[1]],
    [support_or_commitment[2]], [support_or_commitment[0]], [support_or_commitment[5]]
]
w_development=[
    [development[0]],
    [development[1],development[2],development[5],development[6],development[7]],
    [development[3],development[4]]]
w_exp=[[exp[1]],[exp[3]],[exp[0]],[exp[6]],[exp[4]],[exp[8]],[exp[5]],[exp[2]],[exp[7]]]
w_Commun=[Commun[3]],[Commun[0]],[Commun[1]],[Commun[2]],[Commun[4]]

dat[i][9]

count=0

dat_organ=[]

for i in range(1,len(dat)):
    temp=[0]*9
    if dat[i][2] not in w_school_level[len(w_school_level)-1] and
       dat[i][6] not in w_implementation_level[len(w_implementation_level)-1] and
       dat[i][7] not in w_implementation_year and
       dat[i][36] not in w_level_of_understanding[len(w_level_of_understanding)-1] and
       dat[i][37] not in w_support_or_commitment[len(w_support_or_commitment)-1] and
       dat[i][42] not in w_development[len(w_development)-1] and
       dat[i][41] not in w_exp[len(w_exp)-1] and
       dat[i][43] not in w_Commun[len(w_Commun)-1]:
        count+=1
        temp[0]=dat[i][9]
        for j in range(len(w_school_level)-1):
            if dat[i][2] in w_school_level[j]:
                temp[1]=j
        for j in range(len(w_implementation_level)-1):
            if dat[i][6] in w_implementation_level[j]:
                temp[2]=j
        temp[3]=dat[i][7]
```

```
        for j in range(len(w_level_of_understanding)-1):
            if dat[i][36] in w_level_of_understanding[j]:
                temp[4]=j
        for j in range(len(w_support_or_commitment)-1):
            if dat[i][37] in w_support_or_commitment[j]:
                temp[5]=j
        for j in range(len(w_development)-1):
            if dat[i][42] in w_development[j]:
                temp[6]=j
        for j in range(len(w_exp)-1):
            if dat[i][41] in w_exp[j]:
                temp[7]=j
        for j in range(len(w_Commun)-1):
            if dat[i][43] in w_Commun[j]:
                temp[8]=j
        dat_organ.append(temp)

variables=['ID','school_level','implementation_level','implementation_year',
           'level_of_understanding','support_or_commitment','development',
           'exp','Commun']
csvf=file('dat_organ.csv','wb')
writer=csv.writer(csvf)
writer.writerow(variables)
writer.writerows(dat_organ)
csvf.close()

Commun=[]

for i in range(1,len(dat)):
    if dat[i][43] not in Commun:
        Commun.append(dat[i][43])
```

Model Construction (R)

```
# Require data and files:
# @files: dat_organ.csv (from data processing part in Appendix D)
# @data: SPBD (from data processing part in Appendix C)

library(dplyr)
library(nlme)
library(lmerTest)

X.hier <- SPBD %>% dplyr::select(-c(Finish.Date))


count <- 0
for(i in 1:dim(X)[1]) {
  if (sum(is.na(X[i,])) > 0){
    count <- count + 1
  }
}

dat_organ <- read.csv("dat_organ.csv")
tab <- merge(dat_organ, X.hier, by="ID")
```

```r
tab_temp <- NULL
for (i in 1:dim(tab)[1]) {
  if (sum(is.na(tab[i,])) == 0) {
    tab_temp <- rbind(tab_temp, tab[i,])
  }
}

colnames(tab_temp)[2:7] <- c('SL','IL','IY','LOU','SOC','DEV')
SPS <- apply(tab_temp[11:33], 1, sum)
attach(tab_temp)

l.mod <- lme(fixed = SPS ~ 1 + SL + IL + Commun + LOU + SOC, method = "REML",
             random = reStruct(~1 + Commun + LOU + SOC|School.Name,
                               pdClass = "pdDiag"))
```

Appendix E: Code for Latent Directlet Allocation

Model Construction (Python)

```python
# Require data and files:
# @files: concerns.xlsx

from nltk.tokenize import RegexpTokenizer
from stop_words import get_stop_words
from nltk.stem.porter import PorterStemmer
from gensim import corpora, models
import gensim
import xlrd

tokenizer = RegexpTokenizer(r'\w+')

# create English stop word list
en_stop = get_stop_words('en')
en_stop.extend([u'school', u't'])

# create p_stemmer of class PorterStemmer
p_stemmer = PorterStemmer()

# list of tokenized documents in loop
texts = []

print("Reading concerns data")

book = xlrd.open_workbook("concerns.xlsx")
sheet = book.sheet_by_index(0)
for row_index in xrange(1, sheet.nrows):
    concern = sheet.row_values(row_index, end_colx=1)
    concern = concern[0].lower()
    tokens = tokenizer.tokenize(concern)

    # remove stop words from tokens
    stopped_tokens = [i for i in tokens if not i in en_stop]
```

```python
    # stem toekns
    stemmed_tokens = [p_stemmer.stem(i) for i in stopped_tokens]

    # add tokens to list
    texts.append(stemmed_tokens)

print("Reading complete")


# turn our tokenized documents into a id <-> term dictionary
dictionary = corpora.Dictionary(texts)


# convert tokenized documents into a document-term matrix
corpus = [dictionary.doc2bow(text) for text in texts]

print("Ready for building LDA model")

# generate LDA model
lda_model = gensim.models.ldamodel.LdaModel(corpus, num_topics=5, id2word = dictionary,
                                            passes = 50)


print(lda_model.print_topics(num_topics = 5, num_words=10))
```