Research on Factors Affecting Campus Bullying Based on Factor Analysis Model

Dexiang Ruan, Baokun Hu*
School of Hangzhou International Service Engineering College, Hangzhou Normal University, Hangzhou, China

*Corresponding author e-mail: hubaokun@gmail.com

Abstract. School bullying has frequently appeared in major primary and secondary schools and even universities. It has become a topic of great concern to all sectors of society. Campus safety issues have greatly affected the physical and mental health of young people, and even have an indelible shadow on their future. The rapid development of technology has provided a powerful help for studying the factors affecting contemporary students' school bullying. This paper uses the factor analysis model to compare the dominant factors affecting school bullying from the horizontal aspect, which can provide reference for parents and schools to judge potential campus bullying, and has positive significance for preventing and controlling campus bullying.

1. Introduction

All along, school bullying has been the headline event of major news websites, which has aroused widespread concern from parents, schools and the community. School bullying and violence are no longer unfamiliar words, and their seriousness and concealment are on the rise. A sample survey conducted by the China Youth Research Center on primary and middle school students in 10 provinces and cities shows that 32.5% of people have this tendency of people are occasionally bullied, and 6.1% are often bullied by their elders. Bullying and violence in school are common, shocking youth is the future of the country, is the hope of the nation, education campus should be the safest, the sunniest place. But, student bullying and violence in a few places not only endanger the physical and mental health of minors, but also easily form public hot spots and affect social stability. Regarding the phenomenon of bullying and violence on campus, many NPC deputies believe that prevent and control bullying and violence in schools can not be achieved by a single paper literature. It needs to be long term government, family, society and schools. Hard, meticulous work and joint efforts. It is recommended to further define the "campus bullying" improve relevant legislation at the national level, formulate anti-bullying special law, and increase the protection of minors[1]. International empirical research shows that school bullying whether it has a significant negative impact on adolescent health[2].

Bullying is a deliberate, repetitive, and damaging behavioral event that occurs between individuals with unbalanced power[3]. Campus bullying will seriously harm the physical and mental health, academic development and social adaptability of young people in the transition period, and have an irreversible impact on the long-term development of students[4]. Studies have shown that students who suffer from bullying on campus have lower academic performance than their peers who have not been bullied, and have greater risk of detecting psychological symptoms such as anxiety, depression, loneliness and suicidal tendencies. Bullying is a serious obstacle. The academic ability of students and
the development of sound personality [5, 6].

At present, there are many methods for analyzing the factors affecting school bullying. Huang Liang chose two levels of Bernoulli model to analyze the main factors of campus bullying [7]; Chen Chunjin et al. used Logistic model to analyze the main factors of campus bullying [8]; Wang Jian, Ru Fuxia, etc. Using multi-factor logistic regression to analyze the main factors of campus bullying [9, 10]; Li Ling et al. used structural equation modeling to analyze the main factors of campus bullying [11]; based on this, this paper uses the horizontal dimension factor analysis model to make practical reference and basis for establishing the prevention and control mechanism of campus bullying.

2. Data source

The data used in this paper is from the Organization for Economic Cooperation and Development (OECD). Planned “Programme for International Student Assessment” (PISA) for the global 2015. PISA This unobservable social phenomenon in school culture will be assessed by measuring student perceptions of school bullying [12]. First, PISA for junior high school students around 15 years old, Comprehensive testing of mathematics, reading and scientific literacy, it also collects information on all aspects of adolescent growth, especially the detailed description of the student’s family background and school information. Second, the PISA project collects data from multiple countries and regions, and the scientific and representative nature of sampling has good reliability and validity. Once again, the background of the PISA project has a rigorous theoretical concept. The preparation of the test questions and background questionnaires is scientific and rigorous. The test organization process specifications and the data have high reliability. PISA 2015 measured the bullying behavior of young people on campus for the first time, which provides rich and reliable data for studying bullying in schools across the country. A total of 9,841 students from 268 schools in four provinces (Beijing, Shanghai, Guangdong, and Jiangsu) participated in the evaluation. This provides favorable conditions for analyzing the bullying status and influencing factors of students in four provinces and cities in China.

3. Factor analysis model

Factor analysis is a data reduction technique. By analyzing the internal interdependence between variables, we explore the basic structure of data. The basic data structure is represented by several abstract variables. These abstract variables are called factors and can reflect the main variables of many variables. information. The original variable is the apparent variable that is observable, and the factor is the unobservable latent variable.

3.1 Mathematical model

The factor analysis type is divided into R type factor analysis and Q type factor analysis. The factor analysis of the R type is a factor analysis of the variables, and the Q type factor analysis is a factor analysis of the samples. This paper selects the R-factor analysis. The common factor in R-factor analysis refers to the common influencing factors that can not be observed directly but exist objectively.

Every variable can be represented as the sum of a linear function and a special factor of a common factor,

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m + \varepsilon_i, (i = 1, 2, \ldots, p)$$  \hspace{1cm} (1)

In the above formula, $F_1$, $F_2$, $\ldots$, $F_m$ is called a common factor, $\varepsilon_i$ is called a special factor of $X_i$, and $X_i$ is a measurable variable.

The matrix form of the model is as follows:

$$X = AF + \varepsilon$$  \hspace{1cm} (2)

among them,
\[ A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pm} \end{bmatrix} = (A_1, A_2, \ldots, A_i) \quad (3) \]

\[ X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, \quad F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_p \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix} \quad (4) \]

This mathematical model needs to meet the following four aspects:

1. \( m < p \), the number of common factors extracted is less than the number of original variables.

2. \( \text{Cov}(F, \varepsilon) = 0 \), the common factor and the special factor are not related.

3. \( D(F) = \text{Im} \), the common variance of each common factor is 1.

4. \( \text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \), \( D(\varepsilon_i) = \sigma_j \), each particular factor is uncorrelated and has a different variance.

3.2 The nature of the factor analysis model

1. Decomposition of the covariance matrix of the original variable \( X \) by \( X = AF + \varepsilon \), \( \text{Cov}(X) = ACov(F)A^T + \text{Cov}(\varepsilon) \).

\[ \text{Cov}(X) = AA^T + \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2) \quad (5) \]

The smaller the value of \( \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2 \), the more components shared by the common factor.

2. The load matrix is not unique

Let \( T \) be an \( m \times m \) matrix, let \( A^* = AT, F^* = T^TF \), the model can be expressed as:

\[ X = A^*F^* + \varepsilon \quad (6) \]

3.3 Factor load matrix

3.3.1 Statistical significance of factor loading matrix

1. Statistical significance of factor load \( a_{ij} \)

For the factor model,

\[ X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m + \varepsilon_i, \quad (i = 1, 2, \ldots, p) \quad (7) \]

The covariance of \( X_i \) and \( F_j \) is:

\[ \text{Cov}(X_i, F_j) = \text{Cov}\left( \sum_{k=1}^m a_{ik}F_k + \varepsilon, F_j \right) \]

\[ = \text{Cov}\left( \sum_{k=1}^m a_{ik}F_k, F_j \right) + \text{Cov}(\varepsilon_i, F_j) \]

\[ = a_{ij} \quad (8) \]

If \( X_i \) is standardized, the standard deviation of \( X_i \) is 1, and the standard deviation of \( F_j \) is 1, there is

\[ \gamma_{X_iF_j} = \frac{\text{Cov}(X_i, F_j)}{\sqrt{D(X_i)D(F_j)}} = \text{Cov}(X_i, F_j) = a_{ij} \quad (9) \]

Then, for the standardized \( X_i, a_{ij} \) is the correlation coefficient of \( X_i \) and \( F_j \), indicating that \( X_i \) depends on the proportion of \( F_j \). The psychologist calls it the payload, which represents the load of the \( i \)-th variable on the \( j \)-th common factor, and reflects the relative importance of the \( i \)-th variable on the \( j \)-th common factor.

2. Statistical significance of the commonality of variables

Known by the factor model,
\[ D(X_i) = a_{i1}^2 D(F_1) + a_{i2}^2 D(F_2) + \cdots + a_{im}^2 D(F_m) + D(\varepsilon_i) \]
\[ = a_{i1}^2 + a_{i2}^2 + \cdots + a_{im}^2 + D(\varepsilon_i) \]
\[ = h_i^2 + \sigma_i^2 \] (10)

The commonality of the variable \( X_i \) is: \( h_i^2 = \sum_{j=1}^{m} a_{ij}^2 \quad i = 1, 2, \ldots, p. \)

If \( X_i \) is standardized, there is \( 1 = h_i^2 + \sigma_i^2. \)

(3) The statistical significance of the variance contribution rate \( g_j^2 \) of the common factor \( F_i \)

Let the factor load matrix be \( A \), the sum of the squares of the elements of the \( j \)th column,
\[ g_j^2 = \sum_{i=1}^{p} a_{ij}^2 \quad j = 1, 2, \ldots, m \] (11)

The contribution of the common factor \( X_i \) to \( F_j \), \( g_j^2 \), represents the sum of the variance contributions provided by the same common factor \( F_j \) for each variable, and the total effect of the factor on all variables is the relative importance of each common factor. For a scale, the relative importance of the common factor \( F_j \) can be measured by \( (g_j^2)h \), which is called the contribution rate of the common factor \( F_j \) to \( X \). The purpose of factor analysis is to find the solution of the factor analysis model from the covariance matrix \( \Sigma \) or the correlation matrix \( R \) of the original random vector, that is, find the load matrix \( A \) and the characteristic variance matrix \( D \), and give the common factor to give an explanation of the actual background.

4. Experimental results and analysis

Using PISA2015 data analysis, specific methods should be used to test the data before factor analysis to determine whether the secondary data is suitable for factor analysis. Through the KMO and Bartlett sphericity test, the results show that the KMO value is 0.604, and the saliency test approximates the chi-square value corresponding to the significance probability of 0.000, reaching a significant level, that is, there is a strong correlation between the variables, indicating the data. Suitable for factor analysis.

| Component | Initial Eigenvalues | Extraction Sum of Squared Loadings |
|-----------|---------------------|-----------------------------------|
|           | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1         | 3.139 | 19.616 | 19.616 | 3.139 | 19.616 | 19.616 |
| 2         | 1.925 | 12.031 | 31.647 | 1.925 | 12.031 | 31.647 |
| 3         | 1.716 | 10.726 | 42.374 | 1.716 | 10.726 | 42.374 |
| 4         | 1.295 | 8.096  | 50.470 |
| 5         | 1.114 | 6.965  | 57.435 |
| 6         | 1.069 | 6.683  | 64.118 |
| 7         | 0.979 | 6.119  | 70.237 |
| 8         | 0.883 | 5.521  | 75.758 |
| 9         | 0.783 | 4.895  | 80.654 |
| 10        | 0.757 | 4.732  | 85.386 |
| 11        | 0.660 | 4.122  | 89.508 |
| 12        | 0.533 | 3.334  | 92.842 |
| 13        | 0.455 | 2.843  | 95.685 |
| 14        | 0.318 | 1.985  | 97.670 |
| 15        | 0.311 | 1.945  | 99.615 |
| 16        | 0.062 | 0.385  | 100.000 |

Extraction Method: Principal Component Analysis.
4.1 Total variance of interpretation

Because of the extraction factor requires eigenvalues greater than 1, three variables are extracted from the original 16 variables as the common factor. The larger the cumulative variance contribution rate of the common factor, the more information is reflected in the original indicator. According to the above table, when the first three common factors are extracted, the cumulative variance contribution rate reaches 42.374%, indicating that these three common factors can reflect 42.374% of all variables, and can better replace 16 indicators. In order to explain the factors more conveniently, the three common factors extracted are represented by F1, F2, and F3, respectively, and the three common factors are used to characterize the student background, the school background feature, and the three dimensions of the support variable. The common factor F1 has a greater correlation in the male, grade, family social status, ordinary junior high school, ordinary high school, and academic performance, and the coefficient is greater than 0.5. Variables reflect student background characteristics from 6 different perspectives, so F1 can be named as student background features.

The public factor F2 has a greater correlation in urban schools, school scale, class size, teacher-student ratio, public school, repeat student ratio, and disciplinary atmosphere, with coefficients greater than 0.5. Therefore F2 can be named as a school background feature.

The common factor F3 reflects the degree of support in parental emotional support, teacher support, and teacher injustice, with a coefficient greater than 0.5. Therefore F3 can be named as a support variable.

4.2 Scree Plot

![Scree Plot](image)

**Figure 1.** Extract characteristic variable Scree Plot

As shown in the Figure 1, the abscissa is the number of factors, and the ordinate is the eigenvalue. It can be seen from the gravel map that the eigenvalues of the first three factors are high, and the contribution rate to the original variables is the largest; at the third common factor In the future, the eigenvalues are all less than 1, which can be neglected gravel, indicating that the three common factors extracted are more reasonable.
4.3 Component score coefficient matrix

| Component Score Coefficient Matrix |
|------------------------------------|
| Component | 1 | 2 | 3 |
| boys      | .025 | .213 | .079 |
| grade     | .066 | -.003 | -.018 |
| family social status | .309 | -.008 | .030 |
| ordinary junior high school | .024 | .150 | .075 |
| ordinary high school | -.016 | -.067 | -.096 |
| academic performance | .280 | .001 | .011 |
| city school | -.002 | .189 | .080 |
| school scale | .019 | .285 | -.344 |
| class size | .018 | .200 | -.370 |
| teacher-student ratio | -.022 | -.102 | .274 |
| public school | -.050 | .330 | .143 |
| repeat student ratio | -.020 | -.004 | .328 |
| disciplinary atmosphere | .267 | .043 | .024 |
| parent emotional support | -.031 | -.043 | -.195 |
| teacher support | -.005 | .406 | .237 |
| teacher injustice | .250 | -.050 | .022 |

The obtained three common factors were used as the dependent variables for regression analysis, and the factor score matrix was obtained. The factor score reflects factor analysis. The factor scores can be used to analyze the problems studied more intuitively.

Although the original variable information exists in the three common factors, based on the rigorous research attitude, the situation of student bullying cannot be reflected by a single independent factor. In order to be accurate, a comprehensive study of student bullying, The rate of contribution is weighted based on the total variance.

Factor 1: \( F_1 = .025X_{\text{boys}} + .066X_{\text{grade}} + .309X_{\text{family social status}} + .024X_{\text{ordinary junior high school}} + -.016X_{\text{ordinary high school}} + .280X_{\text{academic performance}} + -.002X_{\text{city school}} + .019X_{\text{school scale}} + .018X_{\text{class size}} + -.022X_{\text{teacher-student ratio}} + -.050X_{\text{public school}} + -.020X_{\text{repeat student ratio}} + .267X_{\text{disciplinary atmosphere}} + -.031X_{\text{parental emotional support}} + -.005X_{\text{teacher support}} + .250X_{\text{teacher injustice}} \)

Factor 2: \( F_2 = .213X_{\text{boys}} + -.003X_{\text{grade}} + -.008X_{\text{family social status}} + .150X_{\text{ordinary junior high school}} + .067X_{\text{ordinary high school}} + .001X_{\text{academic performance}} + .189X_{\text{city school}} + .285X_{\text{school scale}} + .200X_{\text{class size}} + .102X_{\text{teacher-student ratio}} + .330X_{\text{public school}} + -.004X_{\text{repeat student ratio}} + .043X_{\text{disciplinary atmosphere}} + .043X_{\text{parental emotional support}} + .406X_{\text{teacher support}} + .050X_{\text{teacher injustice}} \)

Factor 3: \( F_3 = .079X_{\text{boys}} + .018X_{\text{grade}} + .030X_{\text{family social status}} + .075X_{\text{ordinary junior high school}} + .096X_{\text{ordinary high school}} + .011X_{\text{academic performance}} + .080X_{\text{city school}} + .344X_{\text{school scale}} + .370X_{\text{class size}} + .274X_{\text{teacher-student ratio}} + .143X_{\text{public school}} + .328X_{\text{repeat student ratio}} + .024X_{\text{disciplinary atmosphere}} + .195X_{\text{parental emotional support}} + .237X_{\text{teacher support}} + .022X_{\text{teacher injustice}} \)

In the light of the variance contribution rate of each factor, the following common factor function have access to:
\[ Y = 19.616/42.374F1 + 12.031/42.374F2 + 10.726/42.374F3 \]
5. Conclusions

In this paper, aiming at the analysis of the factors affecting the bullying of young people in school, the horizontal factor analysis model is applied. Firstly, the common factor is extracted from the original variables, and then the rationality of the extracted common factors is judged by the gravel map. The factor's component score coefficient matrix identifies the factors that most influence the bullying of adolescents. The experimental results show that from the perspective of horizontal dimension analysis, the factors affecting students' bullying are as follows:

F1>F2>F3, student background characteristics> school background characteristics> emotional support, student background characteristics are composed of boys, grades, family socioeconomic status and students' academic performance, which have a significant impact on students' bullying. In view of this, school administrators must establish an effective early warning system for campus bullying. The first thing to do is to pay attention to and understand the groups with special characteristics and distinctive characteristics, including potential bullies and bullies, and fully grasp and analyze students during educational counseling. The family growth environment and family education concept, and detailed records, the establishment of real-time tracking, evidence-based bullying case data and database, in order to achieve "precise intervention" on campus bullying. Based on this paper, the factor analysis model of horizontal dimension is adopted. Therefore, the factors affecting campus bullying will be explored through longitudinal dimension analysis.

References
[1] Formulating a special law against bullying [EB/OL].http://xinhuanet.com/legal/2018-04/17/c_1122692250.htm,2018-04-17.
[2] OECD. PISA 2015 Results (Volume III) :Students' Well-Being[R]. Paris: OECD Publishing, 2017.
[3] Smith PK, Morita Y, Junyer-Tas J, Olweus D, Catalana R, Slee P. The Nature of School Bullying: A cross national perspective[M]. London, United Kingdom: Routledge, 1999:1.
[4] Sarah Woods, Dieter Wolke. Direct and relational bullying among primary school children and academic achievement[J]. Journal of School Psychology, 2004, 42(2): 135-155.
[5] RINTAKERTTU KALTIALA-HEINO, MATTI RIMPELÄ, PÄIVI RANTANEN, ARJA RIMPELÄ. Bullying at school—an indicator of adolescents at risk for mental disorders[J]. Journal of Adolescence, 2000, 23(6): 661-674.
[6] Delprato, M., Akyeampong, K., & Dunne, M. The impact of bullying on students' learning in Latin America: A matching approach for 15 countries[J]. International Journal of Educational Development, 2017, 52 (2) : 37-57.
[7] Huang Liang. The situation and influencing factors of school bullying in 15-year-old students in China — Based on the analysis of PISA2015 data of four provinces and cities in China[J]. Educational Science Research, 2017(11): 36-42.
[8] Chen Chunjin , Zhi Tingjin. The Influencing Factors of Campus Bullying and the Construction of Long-term Prevention Mechanism — Based on the Analysis of 2015 Bullying Behavior Measurement Data of Youth Campus[J]. Educational Science Research, 2017, 37(20): 31-41.
[9] Wang Jian, Li Chunmei, Xie Fei, Melissa Deadmond, Gabriela Vieyra. Analysis of Factors Affecting School Bullying in High School Students in Jiangxi Provinces[J]. Chinese School Health, 2018, 39(12): 1814-1817.
[10] Ru Fuxia, Yin Yingqun. Analysis on Characteristics and Influential Factors of Bullying Behavior of Middle School Students in Jiangxi Provinces [J]. Chinese School Health, 2018, 39(11): 1655-1660.
[11] Li Ling, Zhang Pingjuan, Wang Han. Research on the Causes of School Bullying and Countermeasures in Middle School Students from the Perspective of Social Control Theory[J]. China’s electrification education, 2018(08): 1-8.
[12] Smith, P.K., & Brain, P. Bullying in schools: Lessons from two decades of research[J]. Aggressive behavior, 2000, 26(1): 1-9.