PROCEEDINGS OF SCIENCE

Analysis of College Students' Spirit Index Based on Structural Equation Model

Weiqing Cao[1a], Xiaofei Li[2b], Xiangzhao Kong[3c]
School of Science, Beijing University of Posts and Telecommunications
Beijing, 100876, China
Email:[a]10927kjj@126.com;[b]417812026@qq.com;[c]740166066@qq.com

Jianguo Chen[d], HeMing Xia[e], YaJie Li[f]
School of Science, Beijing University of Posts and Telecommunications
Beijing, 100876, China
Email:[d]41027825383@qq.com;[e]hemingxhm@163.com;[f]lyj7712@163.com

In order to measure the happiness of Chinese college students specifically, we build a new indicator in our study, Spirit Index, using data mining and the structure equation model. With the principal component analysis method, we get 9 factors as the observed variables from 18 original observations. Then we select an optimum structure equation model after considerable trying in AMOS software and present the outcomes in data comparative forms as well as the Chernoff faces. After analyzing, we find that the fitness affects students' Spirit Index more significantly than mind does. Relevant departments may need to pay attention to the physical fitness education for Chinese college students in the meantime to improve the psychological conditions. In addition, the Spirit Index is a scientific indicator which can be used in the assessments of students' happiness.

CENet2017
22-23 July 2017
Shanghai, China

1Speaker
2Corresponding Author
3The study is supported by the Research Innovation Fund for College Students of Beijing University of Posts and Telecommunications and Project number is 207.

PoS(CENet2017)075

1.Introduction

Happiness is something subjective and invisible, so many people try to embody it as a specific number which we name Spirit Index. To put it simply, happiness is a kind of subjective psychological feeling of life. Happiness research is a hot issue in today's society, and the study of happiness of college students is of great theoretical and practical significance. College students' happiness is not only the necessary basis for quality-oriented education, but also the key to quality training. Meanwhile, students' creativity and the development of various comprehensive abilities are more closely related to happiness.

Researches conducted by the international scholars on happiness has experienced a brief description of the population statistics related to the subjective well-being project. Scholars through the study of ways to be happy under the multicultural environment establish the related theory mainly focused on the researching methods to explore. The cutoff criteria for goodness or fitness indicators and chi-square test is given by Wen Zhonglin, Hau KitTai and Herbert W Marsh(2004).The relationship among job, life, income distribution and subjective well-being based on SEM is studied by Sunfeng(2007). And the research on the well-being of residents ,medical staff ,workers and college students is given by Gu Shuyu(2014) ,Jin Weihua(2015), Li Hongxuan(2014) and Shisi(2014).

At present, the researches of the subjective well-being tend to research on methods diversification, tools localization, and process rapidness. Through this, we use more scientific methods of mathematics to analyze and summarize the mathematical rules behind the happiness. We explore the measurement of happiness of college students, and get a more realistic mathematical model. The main content of the whole project is through the investigation to decide which is more effective to improve college students' happinesd, the subjective factors that affect college students' sense of happiness or the external objective factors. The core issue However, we are not content to simplily analyze the reasons. We should further study the method of enhancing college students' sense of happiness, so as to make the research to be of ore practical significance. To investigate the college students in different universities of the status quo of happiness, we make reasonable designs of the questionnaire, including questionnaire investigation, collecting and sorting data. By analyzing the questionnaire collected data and applying the theory of the structural equation modeling analysis (SEM), we get the factors influencing college students' sense of happiness. Meanwhile, we conclude that college students' happiness approximately obeys the laws of the function. We analyze problems about happiness using the theory of structural equation, so that we can improve the effective utilization of data, making the model closer to reality and more accurate. Finally, we use AMOS software to explore Spirit Index about the relationship between the variables, creating a model to test the interaction between the variables and the reasons. We can obtain more accurate and abundant comprehensive analysis results than using factor analysis or regression analysis alone. We use psychological theory and the hierarchy of needs theory, optimization of the survey questions and the demographic questionnaire design, making the investigation more profound elaborated to the reaction of the people's well-being.

2.Data and Model

2.1 Model Hypothesis

According to the needs of this project, this paper designed the questionnaire which consisted of identity attributes of the survey object and 18 evaluation indicators, with reference to GWB (General Well-Being Schedules). There are both positive and negative questions in the questionnaire, and the higher the final score is, the more important the indicator is. Research subjects are all college students. The survey was conducted in two phases: firstly, 200 questionnaires were sent out and 198 valid ones were retrieved. The questionnaire was revised according to the results. Then the questionnaire was distributed in the second stage when 589 copies were distributed, deleting 6 foreign data (our project focuses on Chinese college students). Finally we checked the internal consistency reliability of the questionnaire.

Internal consistency reliability, also known as internal consistency coefficient, is used to measure the consensus degree of the multiple measurement indicators. The internal consistency of the Cronbachα test scale is commonly used in academia. Hair, Anderson, Taehan (1988) pointed out that the internal consistency coefficient greater than 0.7 indicated that the reliability of the scale was high. In the exploratory study, the internal consistency coefficient might be less than 0.7, but greater than 0.6. In our study, the Cronbachα coefficient is 0.777(shown in Table1), exceeding the acceptable level of 0.7, indicating that each concept of the scale has high credibility[1].

2.2Model Establishment

2.2.1 Model Hypotheses

A. There are sufficient relevance in these variables.

B. Our data conforms to the multivariate normal distribution.

C. There is a causal relationship in a set of latent variables which can be represented by a set of observed variables.

2.2.2 Variable Establishment

In the beginning, we try to straightly consider the results of the questionnaire as the indicator variables, but the model based on these is not qualified. We think that we ought to make these data more effective so that we can use the factor analysis to build a better model[2]. According to Table 1, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy is 0.900, and the Sig. is 0.000, so our data is suitable for the factor analysis.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		
Bartlett's Test of Sphericity	Approx. Chi-Square	
	df	
	Sig.	

Table 1: KMO and Bartlett's Test

Factor analysis is a multivariate statistical method used to find a small set of unobserved variables (also called latent variables, or factors) which can account for the covariance among a larger set of observed variables (also called manifest variables). However, in the specific operation process, there is no fixed method or standard to determine the number of public factors and .(factor names)

We can use SPSS to carry out EFA (exploratory factor analysis) with the principal

component analysis method to select the main component in order to determine the number of common factors. A total of m principal components are chosen. And in this literature,there gives two commonly used standards: one is to choose the main components while the cumulative contribution rate is more than 70% and the other is to select the components whose eigenvalue is greater than a certain value (the general selection is 1 in SPSS) . However, it is shown that the former is easier to take too many principal components while the latter often get small number of the main component. So we combine the two methods in order to receive the best results of the research in the the application process. Because in the SEM, a potential variable must be defined by two or more indicator variables , and finally 18 observations are combined into 9 common factors[3].

Component	Extradition Sums of Squared Loadings			Ro		
	Total	% of Variance	Cumulative %		Total	% of Variance
1	6.199	34.437	34.437		2.046	11.366
2	1.989	11.048	45.485		2.025	11.247
3	1.285	7.141	52.626		1.913	10.629
4	1.209	6.716	59.342		1.611	9.288
5	0.796	4.423	63.765		1.649	9.160
6	0.772	4.292	68.057		1.562	8.679
7	0.722	4.011	72.068		1.163	6.461
8	0.655	3.636	75.704		1.135	6.305
9	0.581	3.230	78.934		1.055	5.859

Table 2: Cumulative Contribution Rate

Evaluation Index	1	2	3	4	5	6	7	8	9
Q9	0.696					0.486			
Q13	0.716								
Q19	0.806								
Q11		0.839							
Q14		0.670							
Q16			0.778						
Q18			0.453						
Q20			0.611						
Q17				0.588					
Q23				0.651					
Q24				0.740					
Q8					0.852				
Q10					0.585				
Q7						0.777			
Q12						0.695			
Q21							0.931		
Q15								0.904	
Q22									0.935

Table 3: Factor Load Capacity

As for factor names, taking factor loadings and information about our questionnaire into consideration, we name the 9 factors in turn as : self-control, negative level, exhausted level, energy, reason level, general mood, health worries, sobriety level and pressure, which are shown in Table 4 as below :

Y1	Self-control
X2	Negative level
X3	Exhausted level
X4	Energy
X5	Reason level
Y6	General mood
X7	Health worries
X8	Sobriety level
Y9	Pressure

Table 4: Factors

We set variables Y1, Y6 and Y9 as the observed variables of the internal latent variable, Spirit Index. Variables X2, X3, X5 and X8 are set as the observed variables of the external latent

variable, mind. And variables X4, X7 (are part of) the other external latent variable, fitness. The variable X in Table 4 is the observed variable of the latent variable, and the variable Y represents the observed variable of the internal variable.

2.2.3 Model Structure

The SEM is divided into the measurement model and the structural model. If the matrix equation is used, there are:Measurement model:

$$Y = \Lambda_y + \varepsilon,\ X = \Lambda_x + \delta \tag{2.1}$$

Structural model:

$$\eta = \Gamma\xi + \zeta \tag{2.2}$$

and are the load matrix, which represents the path coefficient matrix of the connecting observed variable and the latent variable with the measurements of the X and Y variables respectively. η is the vector of the internal variable and ξ is the vector of the external variable. Meanwhile is the error vector of the internal variable, and Γ is the regression coefficient matrix of the external variable due to the internal variable.

We draw our SEM in AMOS and it can be seen in Figure 1:

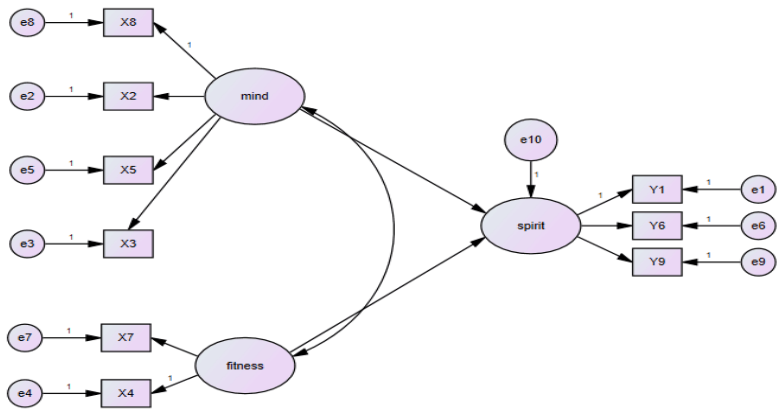

Figure 1: The Initial Structural Equation Model

In the output, all regression coefficients are significant at the significance level of 0.05, and the overall was examined by the significance level. At the same time, all the error variances are positive, which means estimated regression coefficient that were put by model does not surpass the acceptable range. Nevertheless, there are still some demerits in it.

2.2.4 Outlier Processing and Model Modification

We can see that the c.r. is too high, which means that some single-variables may have outliers. It also means that these single-variables violate the assumption that initial variables conform to the normal distribution. In the observation farthest from the centroid Table, the p2 is more reasonable to judge whether a single-variable violates the assumption than p1. Therefore, we delete 32 outliers that p2<0.05[4], which makes the model better.

However, we find some value like AGFI and CN are still unqualified, so we need to do further model modification. In the process of model fitness test, we modify the hypothesis model that is the release of certain assumptions. In the revised index value of the report,

according to the MI index value, we take the model adaptability and the simplicity of the model into consideration, and delete some error terms from the fixed parameter to the free parameter, which introduces the covariant relationship. The results can be better. And the analysis of the index variables, corresponding to the error variables that increase the covariant relationship, also proves that it has a certain correlation.

Therefore, we get a more reliable model as follow:

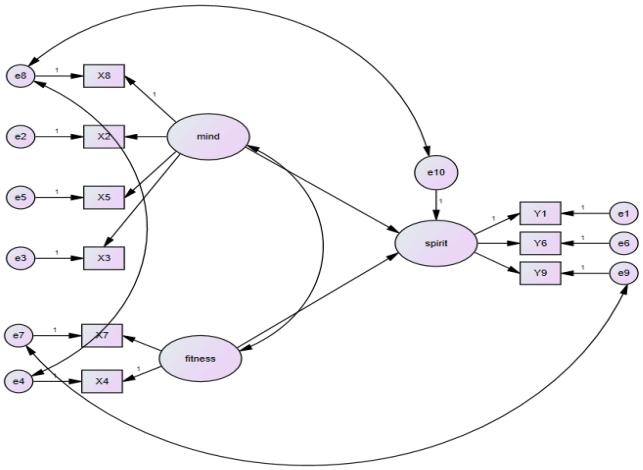

Figure 2: The Final Structural Squation Model

2.2.5 Model Test

At first, we use the P-P plot for its normality test. We use the SPSS software to draw the P-P plots of each factor and we find that each factor obeys the normal distribution. In the general factor analysis, as long as the sample size is large enough, and each variable obeys the normal distribution, the joint distribution of these variables is also normal. As a result, we think that it satisfies the hypothesis about multivariate normal distribution.

Moreover, we can use the AVE (Average Variance Extracted) to represent the convergent validity of the model. AVE reflects each how much of the latent variables explained variance from all questions in the latent variables. When the AVE value is greater than 0.50, it indicates that the latent variables have good convergent. In our model, the AVE value is 0.5311, which means that our model has good convergence.

We can also use CR (Construct Reliability) to test the internal consistency of the model. When the CR value is greater than 0.70, it indicates that the model has good construct validity. In our model, the CR value is 0.8764, which means that our model has strong construct validity[5].

3.Outcome Analysis

3.1 Output Interpretation

The partial output of our model is presented in Table 5.

	Default Model	Saturated Model	Independence Model
CMIN/DF	4.190	-	87.980
GFI	0.967	1.000	0.434
AGFI	0.928	-	0.293
RMSEA	0.076	-	0.336
TLI	0.949	-	0.000
CN(0.01)	244	-	15

Table 5: Model Fit

Due to our large size of sample, we don't refer to the CMIN (Chi-square value) and its p-value that may have relatively large errors. The CIMN/DF is 4.190<5, which meets the requirement. As for other indicators,such as the GFI (goodness of fit index), the AGFI (adjust goodness of fit index) and the TLI (Tucker-Lewis Index), all need to be higher than 0.90. Meanwhile, the RMSEA (Root Mean Square Error of Approximation) is supposed to be lower than 0.80 and the HOELTER 0.01 represents that the Critical N should be larger than 200. According to the table, all of the indicators meet the requirements, which means that our model and data are well fitted[6].

3.2 Practical Analysis

3.2.1 Analysis of Factors

From Figure3, the factor load of the factor, mind, is 0.25, and the factor load of the factor, fitness, is 0.46. In other words, fitness affects the Spirit Index more significantly than mind. We can see from Table 6, the path coefficient of different factors is different. For example, the Energy affects Spirit Index dramatically and the Sobriety level affects Spirit Index more slightly than others. We can improve the well-being of college concretely from these aspects.

Fitness		M
factors	path coefficient	factors
Negative level	0.93	Health worries
Sobriety level	0.23	Energy
Reason level	0.83	
Exhausted level	0.87	

Tble 6: Ingredient List

What we think is realistic in because what Chinese college students are most likely to worry about is almost only study and its influence is always limited. However, as for fitness, the disease damage to them is unbearable and annoying. One who just gets a cold will become more irritable and inpatient, so he/she naturally doesn't feel happy at all. Of course, we don't exclude the situation that some students who are unhappy even behave extremely due to the mental factors. However, these are several individual cases which can't represent all the students. We also take it into consideration, so the factor load of mind is obviously greater than 0 and just lower than fitness.

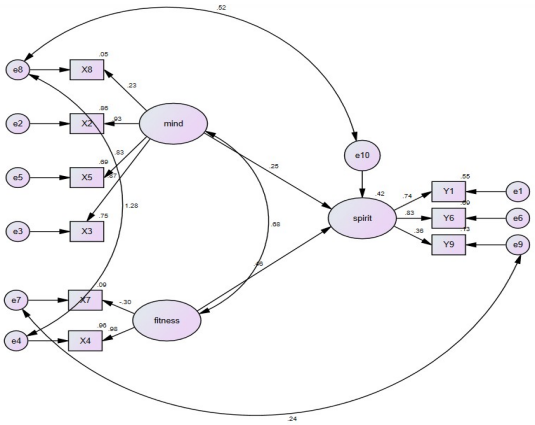

Figure 3: The Final Structural Equation Model

3.2.2 Analysis of Spirit Index

According to the SEM, we don't only know which factor affects the SI (Spirit Index) most but also get an algorithm for it. The formula is shown as below：

$$SI=\gamma_1/(4\times2)\times((X_8-e_8)/\lambda_8+(X_2-e_2)/\lambda_2+(X_5-e_5)/\lambda_5+(X_3-e_3)/\lambda_3)+\gamma_2/(2\times2)\times((X_7-e_7)/\lambda_7+(X$$

In this formula, the path coefficient is represented between the external latent variable and the internal latent variable which indicates the standard error.

We investigate some subjects of our questionnaire, finding that the SI values are almost entirely consistent with their actual conditions, which means that SI can definitely show the grade of one's happiness. In calculation, we get the average of SI which is 5.67. The standard deviation is 1.86, showing that SI of Chinese college students fluctuate greatly and most of them distribute evenly. In reality, it's impossible that everyone is in the same SI level. Moreover, the maximum value is 11.74 while the minimum value is -1.23, which also shows the psychological discrepancy between the individuals. According to the scatter plot about students' SI value, we divide the SI into 5 grades : very happy, relatively happy, general, relatively unhappy and extremely unhappy. They are presented in following Table 7. The majority of the students are in general grade, while there are still several students who seem to be situated in an extreme condition. They may need some psychological help.

Grade	Data Range	Number of People
very happy	>10	11
relatively happy	7-10	101
general	4-7	354
relatively unhappy	2-4	7
extremely unhappy	<2	13

Table 7: Levels of Happiness

As seen in Figure 4, to embody the Spirit Index, R painting 3 Chernoff faces are used to represent the three main typical grades: unhappy (1), general (2) and happy (3) which makes the result of our project clear and visual.

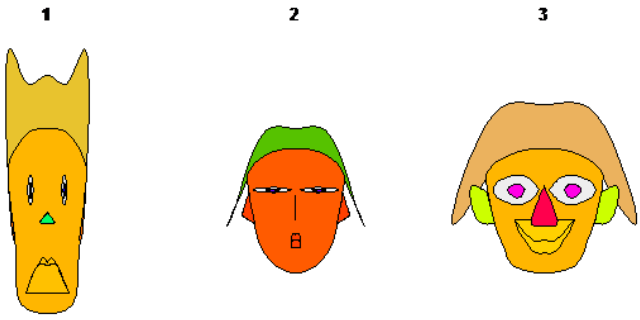

Figure 4: The Chernoff Faces

Through our study, we can get one person's SI after he/she finishes the questionnaire, and then know about his/her happiness grade specifically as well as which aspects mainly influence it. That exactly counts when the students need some advice, which makes our study really more valuable.

3.3 Multiple-Group Analysis

We use multiple-group analysis with the aim of analysing about the discrepancies among different kinds of people such as female and male. We doubt that whether there are significant differences between each group.

First, we use the gender as the variable. The p-value of measurement weights is 0.024<0.05. However, according to the critical ratios, we can see the measure coefficients are less than 1.96. The absolute value between a1_1 and a1_2 is 0.163， and the absolute value between a2_1 and a2_2 are 0.427, 1.770, 0.432, 0.951 and 0.842 respectively. Therefore, we can conclude that there are not significant differences in the factor load[7]. In other words, the measurement model has the same significant level for both female and male.

There are contradictory results through two ways, and it can be learned that the measurement weights model shows only the overall phenomenon. That is, the overall difference between the test may be blinded with a specific factor load between groups. Therefore, the use of " Parameter pairing " to examine the individual variables and to determine the significance level is more effective . As for the structural model, the probable level of structural weights is 0.755>0.05, refering to the structural model possesses invariance between the two genders.

What's more, we conduct a multi-group analysis of the differences in the level of student living expenses and the development level of the cities where the universities are. To our surprise, neither the measurement model, the structure model, nor the covariance is concerned, there are significant discrepancy between different levels.

In addition, from the perspective of the analysis about SI, the average value of sub data under the three categories almost reaches the same result, which we present in Table 8.

	Ave	Std.
male	5.670987	1.858
female	5.675811	1.864
metropolis	5.667273	1.856
Minor city	5.675764	1.863
Higher living costs(>1500)	5.675764	1.863
Lower living costs(<1500)	5.676593	1.862
overall	5.674773	1.861

Table 8: The Average and Standard Deviation of SI

This shows that both males or females feel happy unanimously. As for the cost of living level of Chinese college students, no matter it is more than 1500 RMB or below 1500 RMB, it does not influence the students' SI value. Of course, this phenomenon may result from the background that the major living expenses level concentrates between 1000-2000 RMB, which has a very small interval. The standard of living costs for Chinese student is enough in that the average value is 1200 RMB per month. And comparing those students in the metropolis with those in the minor city, we find that the Spirit Index is almost consistent. At first, we guess students in the metropolis may have more pressure and the happiness might decrease. But after analysis , we find that even though school is different with society, the main indicators that measure students' happiness do not have the significant relations with the urban atmosphere[7].

4. Conclusions and Suggestions

In our study, we mainly use the SEM theory and AMOS software to create the indicator, Spirit Index, standing for the grade of happiness of Chinese college students. Meanwhile, we can clearly see the relationship among mind, fitness and the Spirit Index and find that fitness affects one's SI greater. From our study, we can also tell which aspect is more effective in improving happiness, such as pressure or self-control. However, we can't deny that there are probably some unavoidable errors in our study. For instance, we can't assure that every questionnaire finished is truthful. Besides, our data sample is not absolutely random. Furthermore, when the data changes, the coefficients of our model will be different as well. This also means for one data sample in a condition where our model fits the reality well to a great extent.

Moreover, there are some practical suggestions we want to offer. Firstly, the Spirit Index could be used in colleges as a psychological assessment which makes students be aware of their own conditions. Secondly, colleges should provide more sports opportunities for students. For example, arranging more PE class or lowing the price of gym could motivate students to exercise, so they could be less likely to worry about their health leading to unhappiness. Thirdly, the mind also significantly influences one's happiness level. Therefore, it's essential for a college to provide a psychological counseling room or some lessons about psychology in order to help some students who can't manage their stress correctly on their own, and so forth.

In conclusion, the college students are the main force in building the motherland in the future. Relevant governments ought to pay more attention to their psychology which is an increasing serious issue. We hope that our study could contribute to the psychological shaping of Chinese college students.

References

[1]*A Summary of the Application of Structural Equation Model Theory*[J].Journal of Industrial Technological Economics,2014,:5

[2]*S.tructural Equation Model of College Students' Subjective Well-being*. Modern economic information,2014:1001-828X

[3] *Application of Multivariate Statistical Analysis*. Beijing: Peking University Press, 2006: 266-320

[4]*AMOS and Research Methods*[M].Chongqing University Press,2009.

[5]*Structural Equation Model Based on Reliability Coefficient and α Coefficient Analysis*[J].2009.

[6]*Structural Equation Model - Operation and Application of AMOS*[M].Chongqing University Press,2010

[7].2014 International Conference on Information Management and Management Engineering[C].2014