Application of statistical analysis method in scale validity test

Gang Wang, Huaren Zhou, Jia Yao
Basic Department, Army Engineering University of PLA, Nanjing, Jiangsu 211101, China
627557126@qq.com

Abstract. In order to facilitate research and improve application efficiency, scales based on various functions have been designed and developed. However, in the process of using the scale, people often ignore the validity of the scale for a specific purpose, resulting in the loss of significance of subsequent studies. This paper briefly introduces the application of correlation analysis and factor analysis in scale validity test, which is enlightening to improve people's understanding of statistical analysis methods.

1. Introduction
China has entered the era of big data. People begin to pay attention to data, because data can objectively reflect some attributes, which is helpful to grasp the evolution and law of things. In this situation, scales based on various functions have been designed and developed to facilitate research and improve application efficiency. However, it should be noted that there are still many irregularities in the use of the scale. One of the most prominent points is that people tend to ignore the validity of the scale when it is used for a specific purpose.

Take the mental health test SCL-90 scale for example. It was first introduced and compiled by Wang Zhengyu of Shanghai Mental Health Center in the 1980s, and has been widely used to test mental health status in China since then. But many scholars have found that SCL-90 scale factor structure unstable[1-2]. European scholars studying cannot get the 9 factors of SCL-90 scale[3-4]. In 2006, Xie Hua found that the scale had about 15 factors through the test sample of college students[5]. Wang Zihuan conducted an empirical analysis on the data of Chinese working women in 2017, and concluded that the validity of the scale structure was poor[6]. Therefore, when we prepare to use this scale for group mental health test, we must first test its validity. If the validity fails to meet the test standard, the quality of the test result must not be guaranteed, and the subsequent research will lose its significance. Below, this paper will take SCL-90 scale as an example to briefly introduce the application of some statistical analysis methods in scale validity test.

2. Check ideas and steps
2.1. Inspection idea
Firstly, it should be clarified that the validity of the scale refers to the validity degree of the scale measurement, that is, the extent to which the scale reflects the characteristics of the test object.

(1) Make clear the purpose of using the scale. Here, the SCL-90 scale will be used to test the mental health status of freshmen of XX University. Since the scale has been widely used, no details will be given here.
(2) Collect samples. It is found that there are more than 6000 freshmen in "XX University". The random sampling method will be adopted here, and the effective sample size should reach at least 1200 people.

(3) Test the validity of the scale. Correlation analysis and factor analysis will be used to test the validity of the scale.

(4) Standard comparison.

(5) Follow-up research and analysis.

2.2. Step Implementation

Generally, the validity of the scale is divided into content validity and structure validity. Only when both of them pass the test can the quality of the test content of the scale be guaranteed.

(1) Collect samples. A random sampling method was adopted to issue 1,800 questionnaires to freshmen of "XX University" from September 2019 to October 2019, and 1519 copies were recovered. Screening criteria: ① complete basic information; ② complete and standard answers. Finally, 1264 valid questionnaires were obtained, with an effective rate of 83.21%. The composition of subjects is shown in Table 1.

| Attribute | Gender | Age | Number | Proportion |
|-----------|--------|-----|--------|------------|
|           | Male   |    | 493    | 39%        |
|           | Female |    | 771    | 61%        |
|           | 17     |    | 51     | 4%         |
|           | 18     |    | 608    | 48.2%      |
|           | 19     |    | 541    | 42.8%      |
|           | 20     |    | 64     | 5%         |

(2) Test the validity of the content of the scale. The correlation coefficient between each factor and the total score can be compared and analysed. If the correlation coefficient between the factor and the total score is significantly greater than the correlation coefficient between factors, the content validity is good. Pearson correlation coefficient is introduced as a reference index for comparison. The calculation formula is shown in (1).

\[
\rho_{X_iY_j} = \frac{n \sum X_i Y_j - \sum X_i \sum Y_j}{\sqrt{n \sum X_i^2 - (\sum X_i)^2} \sqrt{n \sum Y_j^2 - (\sum Y_j)^2}}
\]

As shown in formula (1), the correlation coefficient between each factor and the total score is between 0.883 and 0.955, showing a high correlation, and the correlation coefficient between each factor and the total score is greater than the correlation coefficient between factors, indicating that the content validity of the scale is good. The correlation coefficient matrix results of each factor and the total score are shown in Table 2.

| Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 | Y8 | Y9 | Y10 | Z  |
|----|----|----|----|----|----|----|----|----|-----|----|
| 1.00 | | | | | | | | | | |
| 0.79 | 1.00 | | | | | | | | | |
| 0.79 | 0.83 | 1.00 | | | | | | | | |
| 0.82 | 0.86 | 0.87 | 1.00 | | | | | | | |
| 0.86 | 0.84 | 0.85 | 0.90 | 1.00 | | | | | | |
| 0.76 | 0.76 | 0.80 | 0.83 | 0.82 | 1.00 | | | | | |
| 0.82 | 0.75 | 0.82 | 0.84 | 0.86 | 0.80 | 1.00 | | | | |
| 0.79 | 0.78 | 0.84 | 0.85 | 0.85 | 0.82 | 0.82 | 1.00 | | | |
| 0.82 | 0.81 | 0.85 | 0.88 | 0.89 | 0.82 | 0.85 | 0.86 | 1.00 | | |
| 0.84 | 0.82 | 0.82 | 0.86 | 0.87 | 0.8 | 0.83 | 0.82 | 0.85 | 1.00 | |
| 0.91 | 0.90 | 0.92 | 0.96 | 0.95 | 0.88 | 0.90 | 0.90 | 0.93 | 0.92 | 1.00 |

(3) Test the validity of the scale structure. First, KMO test and Bartlett spherical test were used to verify the applicability of factor analysis [7]. The results are shown in Table 3. KMO values of each
factor subscale and total table were greater than 0.8, and Bartlett spherical test was significant, indicating that it was suitable for factor analysis.

Table 3. KMO test and Bartlett test for sphericity

| Factor | KMO  | Bartlett(X)² | P     |
|--------|------|--------------|-------|
| Y1     | 0.946| 7511.177     | <0.001|
| Y2     | 0.931| 4661.706     | <0.001|
| Y3     | 0.907| 3758.890     | <0.001|
| Y4     | 0.951| 8180.227     | <0.001|
| Y5     | 0.925| 5078.129     | <0.001|
| Y6     | 0.863| 2539.167     | <0.001|
| Y7     | 0.902| 3562.913     | <0.001|
| Y8     | 0.877| 2343.385     | <0.001|
| Y9     | 0.932| 4346.338     | <0.001|
| Y10    | 0.870| 2408.553     | <0.001|
| Total  | 0.982| 80979.139    | <0.001|

In order to explore the reasonable structure of the scale, the author hopes to start from the correlation between the original items, so that the degree of item correlation can be explained by common factors as far as possible. Therefore, the principal axis factor method is adopted for the study[8]. Here, “eigenvalue>1” criteria, gravel diagram and parallel analysis are also used for comprehensive analysis[9]. It was found that only 1 factor was extracted from each factor subscale, while 6 factors were extracted from the total table, with a cumulative interpretation of 52.323%. According to the design structure of the original scale, the author retained 10 factors, and the cumulative interpretation was 57.130%. The parallel analysis results were drawn by R software[8], as shown in Figure 1. The analysis of the number of factors is shown in Table 4.

Fig.1 Results of parallel analysis of aggregate table data
Table 4. Analysis of number of factors

| Factor | Number Extraction | Eigenvalues | Variance Explanation | Cumulative Variance Explanation |
|--------|-------------------|-------------|-----------------------|---------------------------------|
| Y1     | 1                 | 5.943       | 49.522%               |                                |
| Y2     | 1                 | 4.404       | 44.039%               |                                |
| Y3     | 1                 | 3.808       | 42.316%               |                                |
| Y4     | 1                 | 6.334       | 48.720%               |                                |
| Y5     | 1                 | 4.588       | 45.882%               |                                |
| Y6     | 1                 | 2.893       | 48.215%               |                                |
| Y7     | 1                 | 3.563       | 50.896%               |                                |
| Y8     | 1                 | 2.776       | 46.264%               |                                |
| Y9     | 1                 | 4.273       | 42.726%               |                                |
| Y10    | 1                 | 2.869       | 40.982%               |                                |
| Total  | 1                 | 38.853      | 43.170%               | 43.170%                         |
|        | 2                 | 2.375       | 2.639%                | 45.809%                         |
|        | 3                 | 1.657       | 1.841%                | 47.650%                         |
|        | 4                 | 1.557       | 1.730%                | 49.379%                         |
|        | 5                 | 1.380       | 1.533%                | 50.912%                         |
|        | 6                 | 1.270       | 1.411%                | 52.323%                         |
|        | 7                 | 1.173       | 1.304%                | 53.627%                         |
|        | 8                 | 1.095       | 1.217%                | 54.844%                         |
|        | 9                 | 1.079       | 1.198%                | 56.042%                         |
|        | 10                | 0.979       | 1.087%                | 57.130%                         |

The author analyzed the 10 factors extracted above according to factor loading. The first factor extracted mainly reflects the characteristics of depression, the second factor to the seventh factor are all multi-factor compound terms, the eighth and ninth factors mainly reflect somatization, and the tenth factor is multi-factor compound terms.

(4) Evaluation of scale validity. The results of exploratory factor analysis showed that through the comprehensive factor extraction method constituted by “eigenvalue>1” criteria, gravel diagram and parallel analysis[9]. Only 1 factor is extracted from each subscale, while 6 factors are extracted from the total scale, which also reflects that there may be strong correlation between some subscales, resulting in the occurrence of compound factors of some components in the process of factor extraction. In order to carry out a comparative study, 10 factors were retained, and according to the load of each factor, an attempt was made to explain each factor. It was found that depression was an independent factor, both factors could explain somatization, and multiple factors could be combined for multiple factors. Through the above analysis, it can be concluded that the SCL-90 scale has good content validity, but poor structural validity, and the original 10-factor structure of the scale cannot be obtained. Therefore, meaningful analysis cannot be performed based on the test data.

3. Conclusion
In the era of big data, the frequency of people's use of various scales increases accordingly, so the standardization of operation should be consciously improved. Therefore, mastering basic statistical analysis methods has become an essential skill. This paper briefly introduces the application of correlation analysis and factor analysis in scale validity test. At the same time, in terms of the selection of factor analysis methods, this paper uses the principal axis factor method for exploratory factor analysis. This method pays more attention to the correlation of original items than the principal component method, which is more in line with the concept of common factor extraction, and the results obtained are more accurate. When determining the number of factors, the comprehensive factor extraction method constituted by "eigenvalue>1" criteria, gravel diagram and parallel analysis can make factor extraction results more objective.
Acknowledgments
This work is supported by the National Social Science Foundation of China (17BGJ005).

References
[1] Yue Peng, Song Qiliang, TANG Hongya. Factor Rationality Analysis of SCL-90 [J]. Xizang Science and Technology,2010(09):45-47(in Chinese).
[2] Bian Junshi, Jing Xixue, Zhuang Na. Controversy on the application of symptom Checklist 90 (SCL-90) in mental health research [J]. Chinese Journal of Health Psychology,2008(02):231-233(in Chinese).
[3] Matti Holli. Assessment of psychiatric symptoms using the SCL-90,ACADEMIC DISSERTATION, To be publicly discussed with the assent of the Medical Faculty of the University of Helsinki. uditorium of the Lapinlahti Hospital,Helsinki,28.3.2003:44-46(in English).
[4] Sereda Y, Serhii D. Validity assessment of the symptom checklist SCL-90-R and shortened versions for the general population in Ukraine [J] BMC Psychiatry, 2016, 16 (1) : 300(in English).
[5] Xie Hua, DAI Haizaki. Evaluation of SCL-90 scale [J]. Neurological Diseases and Mental Health,2006(02):156-159(in Chinese).
[6] Wang Zihuan, Yu Wenlan, Shen Zhum, et al. Reliability and validity of self-rating symptom checklist 90 (scl-90) in psychological assessment of professional women in China [J]. Chinese journal of industrial medicine,2017,30(04):247-250(in Chinese).
[7] Henson, R. K. & Roberts, J.K. Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. Educational and Psychological Measurement, 2006 , 66(3):393-416(in English).
[8] Xue Yi. Chen Liping. Statistical Modeling and R Software [M]. Tsinghua University Press, 2006:519-541(in Chinese).
[9] Kong Ming, Bian Ran, Zhang Hcan. Application of parallel Analysis in Exploratory factor Analysis [J]. Psychological Science,2007(04):924-925+918(in Chinese).