Review on Methods of Comprehensive Evaluation Index System's Effectiveness

Yiping Wang

Management Science and Engineering Department
Shanghai University
Shanghai, China
wypfighting13@126.com

Abstract—The evaluation index system is an important part of the comprehensive evaluation, and its validity will directly determine the quality of comprehensive evaluation，so it is very necessary to assess its effectiveness. Nowadays , with the in-depth research on comprehensive evaluation, more and more attention are paid to the index system's effectiveness by the scholars, and some methods have been proposed, but it is still lack of the system, full-scale research. Based on the researchs, this paper will analyze the existing methods of the index system's effectiveness ----- Reliability & Validity Method, Coefficient of Validity Method, Structural Equation Model Method, Signal/Noise Ratio Method; RST Method, and hope to summarize their advantages, disadvantages and applicable occasions, so as to make the research of the index system effectiveness more perfect.

Keywords- Effectiveness of the Index System; Reliability & Validity Method; Coefficient of Validity Method; Structural Equation Model Method; Signal/Noise Ratio Method; RST Method.

I. INRUODUCTION AND BACKGROUD

At present, multiple index comprehensive evaluations have been applied to many fields, such as enterprise management, agriculture, medicine, transportation and so on, which have multiple attribute characteristics, even it can be said that almost all areas are involved in comprehensive evaluation. For different areas, a large number of the comprehensive evaluation index systems have been summarized by the authoritative specialist, according to their knowledge reserves and the level of understanding. However, on one hand, because the evaluation objects in real life has high complexity, if the main focuses are different, the definitions of evaluation objects are different, or the settings of evaluation goal are different, and so on, all will affect the final establishment of evaluation index system; on the other hand, researchers often found, different scholars always establish different evaluation index system on the same research object, when they choose a different method, and even results contrast is enormous. At this time, a unified criterion is lacked.

Therefore, , for an evaluation index system, how to judge its reliability? Or for different index system, which one is more reliable? That is to say, how to measure the effectiveness of the evaluation index system is particularly important [1-3]. In recent years, scholars has began to study on the effectiveness of index system , and has put forward some evaluation methods, such as Xiaohua Jiang [4] and Can Li [5-6] put forward Reliability & Validity Method; Suicheng Li [7] and Jun yun [8] come up with Coefficient of Validity Method; Flavio, B. [9] propose Structural Equation Model Method; Taguchi [10-11] doctor present Signal/Noise Ratio Method; Yun Fu[12-13] and Dilong Xu [14] put up with RST Method and so on. Each method has its own characteristics and applies range, but it is still lack of system and comprehensive research, and the study is still in exploration stage. Therefore, digging deeper into the existing methods, to analysis out their respective advantages, disadvantages and applies range, is especially important. Comparison and analysis on this issue are as below.

II. RELIABILITY & VALIDITY METHOD

For the effectiveness of index system based on Questionnaire, has always been judged by the Reliability and Validity method. Reliability is credibility, refers to the consistency or stability of comprehensive evaluation, which reflects the degree of control non-system error for comprehensive evaluation. The higher the reliability is, the better consistency and the higher credibility of the comprehensive evaluation is. Validity refers to the degree of the comprehensive evaluation reflects evaluation purpose, which reflects the degree of control system error for comprehensive evaluation. The higher the validity is, the more accurate and the more close to the purpose of the comprehensive is.

Commonly, researchers can use test-retest reliability, parallel-forms reliability, split-half reliability and internal consistency reliability to measure reliability of the index system; and researchers can use face validity, content validity, criterion-related validity, and construct validity to judge validity of the index system. Moreover, researchers can use the help of SPSS software to judge the reliability and validity of the index system. To be specific, in SPSS software, you can use intra-class correlation coefficient ICC, Kappa coefficient, Pearson correlation coefficient, Kendall and Spearman rank correlation coefficient and Cronbach's Alpha coefficient to measure the reliability and validity of index system. Reliability and validity method is relatively simple and convenient to the scholar. But its applicable range is small, it just can be used to judge the effectiveness of the questionnaire evaluation index system,

and it cannot be applied to the non-questionnaire evaluation index system at present.

III. COEFFICIENT OF VALIDITY METHOD

Suicheng Li proposed the Coefficient of validity method, and they consider that the Coefficient of validity method could measure the deviation degree of understanding different evaluation objects face to the same evaluation object. Besides, the greater the deviation degree is, the bigger divergence among the evaluators will be, and then researchers can infer that the weaker the validity of the index system is.

The specific algorithm of the Coefficient of validity method is as below:

Assuming that:

➢ "H" as a set of the evaluation index system, $H = \{h_1, h_2, \cdots h_n\}$;

➢ "M" as the number of evaluation subject;

➢ "Y_j" as a set of the score of every evaluation index, $Y_j = \{y_{1j}, y_{2j}, \cdots, y_{nj}\}$;

➢ "j" as the jth subject of evaluation, $j = 1, 2, \ldots, n$;

➢ "$\overline{y}_i$" as the average score of the index "h_i";

➢ "S" as the highest score of the index of the index "h_i".

Therefore, the validity coefficient of index system is "D":

$$D = \left(\sum_{i=1}^{n} \frac{\sum_{j=1}^{M} |\overline{y}_i - y_{ij}|}{M} * S \right) \Big/ n \qquad (1)$$

Coefficient of validity method is heavily dependent on experts to judging the validity of the evaluation index system, so the subjectivity of the result is stronger, and the objectivity is weaker. That's to say, it is lack of object. But is relatively simple and convenient to user, so it can be used to make a preliminary judgment for the effectiveness of the index system.

IV. STRUCTURAL EQUATION MODEL METHOD

Structural Equation Model (SEM) is an empirical research method to evaluate the consistency between theoretically models and empirical data, and it is a new statistical method to test the reasonableness of a model's assumptions. Here, the confirmatory factor analysis is mainly used to verify the validity of evaluation index system, by comparing the model fitting parameters to achieve it. Chaodong Zhang had used the structural Equation Model method to test the effectiveness of the city competitiveness of China.

The process of assess the validity of index system using the Structural Equation Model method is as Figure 1:
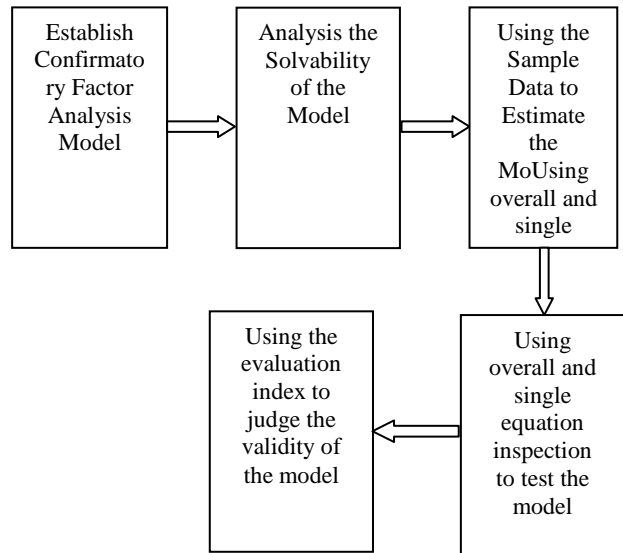

Figure 1. The process of assess the validity of index system using SEM method

Generally speaking, structural equation evaluation index include standardized residual root mean square (SRMR), non-normed fit index (NNFI), normed fit index (NFI), comparative fit index (CFI), etc. And the evaluation index of structural equation evaluation can be calculated by computer program, so the operation is relatively simple and convenient for user; but it is still mainly for a questionnaire evaluation index system or expert's judgement index system, so the relative subjectivity is stronger and its applicable range is also small.

V. SIGNAL/NOISE RATIO METHOD

Dr. Taguchi proposed the concept of Signal/Noise Ratio (SNR), and points out that SNR can be used to measure the system function. Therefore, for the evaluation index system, researchers can use Signal/Noise Ratio to measure the effectiveness of the whole index system, by comparing the size of SNR of the different evaluation index system. Researchers can also contrast the validity of the various index system.

Below is the specific algorithm of the Signal/Noise Ratio method:

Defining that:

➢ $j = 1, 2, \ldots, n$;

➢ $Z_{ij} = (Z_{1j}, Z_{2j}, \cdots Z_{ij}) = (X_{ij} - m_i) \Big/ s_i$;

➢ "X_{ij}" reflects ith index of jth sample;

➢ "m_i" reflects the average of ith index of normal samples;

➢ "s_i" reflects standard deviation of ith index of normal samples;

➢ "'" reflects Transpose vector;

➢ "C^{-1}" reflects inverse matrix of correlation matrix of normal samples;

> ➤ "D_j^2" as Mahalanobis-Distance (MD) of jth abnormal samples;
> ➤ "η" as the SNR of the index system.

First, researchers figure up the Mahalanobis-Distance (MD) of jth abnormal samples:

$$MD_j = D_j^2 = \frac{1}{k} Z_{ij}^{'} C^{-1} Z_{ij} \qquad (2)$$

And then researchers can calculate the the SNR"η" of the index system:

$$\eta = -10\log\left[\frac{1}{t}\sum_{j=1}^{t}\left(\frac{1}{D_j^2}\right)\right] \qquad (3)$$

When use the Signal/Noise Ratio method, researchers should
set a threshold value first, according to our experience, and then compare the value of SNR with the threshold value, if the former is greater than the latter, researchers can think the evaluation index system is effective; when there are multiple evaluation index system, the greater the SNR value is, the more effective the evaluation index system is.
SNR method can judge the validity of the whole evaluation index system, and can compare the effectiveness among multiple index system, but its calculation is relatively complicated.

VI.   RST METHOD

Yun Fu proposes the RST method, and uses the RST method to measure the effectiveness of the index system. To be specific, he proposes that by calculating Redundancy Degree (RD), Sensitivity Degree (SD) and Tendency Degree (TD) three angles, to examine independence, redundant, space universality and time consistency of index system, and further study the effectiveness of the index system.

Below is the specific algorithm of the RST method:

First, for an evaluation object "Q", amusing that:
> ➤ "H" as a set of the evaluation index, $H = \{y_1, y_2, \cdots, y_m\}$, index system $Y \subset H$;
> ➤ "D" as a set of the evaluation index system, $D = \{Y^1, Y^2, \cdots, Y^p\}$;
> ➤ $Y^L = \{y_1^L, y_2^L \cdots, y_{n_L}^L\}$, $L = 1, 2, \cdots, P$,
> ➤ "n_L" as the number of index of the index system "L".
> ➤ For index system "L", "W^L" reflects evaluation results, $W^L = (w_1, w_2, \cdots, w_n)$;
> ➤ "r_{ij}" as the correlation coefficient of index system Y^L.

Based on the above assumption, researchers can figure up the Redundancy Degree of the index system:

$$RD = \frac{\sum_{i=1}^{m_L}\sum_{j=1}^{m_L}|r_{ij}| - m_L}{m_L^2 - m_L} \qquad (4)$$

And then, researchers can get the Sensitivity Degree of the index system:

$$SD = \frac{1}{m_L}\sum_{i=1}^{m_L} SD_i \qquad (5)$$

Also, researchers can work out the Tendency Degree of the index system:

$$TD^{L_1,L_2} = \sum_{t=2}^{m} \varepsilon_{L_2,L_1} / (m-1) \qquad (6)$$

At last, researchers can base on the value of RD, SD, and TD to judge the effectiveness of the index system. That's to say, for an index system, just when $RD \leq 0.5, SD \leq 5$, researchers can consider the index system is valid; when the number of index system is more than 2, the premise $TD \geq 0.8$ should be met first, and then compare the RD and SD to judge which is more effective.

RST method can assess the validity of multiple index systems, and has a complete quantitative assessment criteria. However, when the number of the index system is more than 2, the calculation is very complex, so it is hard to be used by scholar.

VII.   CONCLUSION

Through the above comparison research, researchers analyze the five methods in detail, and it is not difficult to find that in the research on the effectiveness of the comprehensive evaluation index system, the existing methods all have their own advantages, of course, they all have their limitations. Besides, researchers can also get their applicable occasions.
The result of the analysis is shown as Figure 2:

Method	Advantage	Disadvantage
Reliability &Validity Method	*Simple operation * Easy to master	*Application range is relatively small
Coefficient of Validity Method	*Operation is relatively simple * Easier to grasp	*Subjectivity is Very strong, and lack of objectivity
Structural Equation Model Method	*With the aid of computer program, the operation is relatively simple * Relatively easy to learn	*Subjectivity is Very strong, and lack of objectivity *measurable indicator is too much
Signal/Noise Ratio Method	*Has a only judgment standard * Application range is very large	*complicated operate *Hard to use
RST Method	* Theory perfection *Large application range	*excessive calculation * difficult to master

Figure 2. the result of the analysis of the five methods

To sum up, Reliability & Validity method, Coefficient of validity method and Structural Equation Model method can be in the same bracket, because they all can be calculated by some software, like SPSS software and so on, that's to say, they have relatively simple operation and easy mastering. Therefore, researchers can consider that these methods have a relatively high application degree. However, their standards are still not clear, so their application scope is relatively small. However, researchers can use these methods to get an elementary judge, or if a rough result is merely needed, researchers also can use these methods first. To the contrary, researchers can put Signal/Noise Ratio method and RST method in a same sort. Because they have a relatively complete evaluation criteria and indicators, to be further, their applicable condition is relatively large, so these methods are relatively mature and perfect in theory. But as these methods are relied on statistics theory, so their operation are relatively complicated, they are difficult to be master.

VIII. CONCLUSION

Therefore, in different occasions, use which kind of method, sometimes, researchers should consider use some of them at the same time, and researchers should according to the object, choose the appropriate method. Due to personal understanding and the research level are limited, at this stage, the study of the effectiveness of the multi-index comprehensive evaluation index system is as above. In the later study, researchers will continue to dig and summarizes the new method, to enrich and optimize the theoretical system of the comprehensive evaluation index system.

REFERENCES

[1] Weihua Z. (2004).Multi-index comprehensive evaluation of the quality problem. Statistics and Decision. 12: 126-127

[2] Qing Z. (2005). The multi-index comprehensive evaluation of failure factor analysis and model. Statistics and Decision. 3: 6-8

[3] Qing Z. (2004). Multi-index comprehensive evaluation of quality analysis model and its empirical research. Shanghai Management Science. 3:33-34

[4] Jiang, X.H., and Shen, Z.Z. (2010). The reliability and validity of the questionnaire analysis. Modern Preventive Medicine. 37,429-431

[5] Li, C. and Xin, L. (2008)Research on the method of the questionnaire reliability and validity. Chinese Journal of Health Statistics 25,541-544

[6] Jiening, P.(2013). Education evaluation index system of the reliability and validity of research - based on the perspective of evaluation of yuan.Education Exploration. 3: 72-73

[7] Li, S.,C., and Chen J. D., and Zhao, H. G. (2001). Studying on the method of appraising qualitative decision indication system. Systems Engineering-theory & Practice. 9,23-29

[8] Yun, J., and Li, Y. (2009). The effectiveness and optimization of the project evaluation index system. Commercial Times. 9, 47-48

[9] Flavio, B., and Pasquale, S.(2014). Structural equation model for the evaluation of social initiatives on customer behavior. Procedia Economics and Finance, 17, 211 – 220

[10] Taguchi, S. (2000). Mahalanobis-Taguchi System. ASI Taguchi Symposium,

[11] Taguchi, G., and Chowdhury, S., and Wu. Y. (2001).The Mahalanobis-Taguchi System. New York: Mcgraw-Hill

[12] Liverman D M, Hanson M E, Bromn B J, et al.(1998). Global sustainability: Toward measurement. Environmental Management, 12,133-143

[13] Fu, Y., and Liu, Y.J. (2009).Using RST evaluation method to assess the effectiveness of index system. Management Review.7,91-95,112.