

Analysis of the Reliability and Validity of the Evaluation of College Students' General Ability

Meijuan Shan*

Eurasia University, Xi 'an, China

* 282050955@qq.com

Abstract. At present, it is difficult to evaluate students' general ability at home and abroad. Based on the history and research of general education evaluation, this paper defines general education evaluation as “general ability” evaluation of college students. On the basis of literature research, combined with the data obtained from the open questionnaire and interview research, this paper puts forward the theoretical conception of the four capacity structures of “critical thinking and innovative thinking ability, expression and communication ability, analysis and comprehensive information ability and team cooperation ability”, and designs the questionnaire for the assessment of college students' general ability. The results of the questionnaire survey on 2673 college students in a university in xi 'an show that the four elements of competence are reasonable, and the reliability and validity of the self-compiled questionnaire on college students' general literacy ability meet the measurement requirements.

Keywords: *evaluate students' general ability, questionnaire survey, interview research*

1. A Brief Introduction to the General Education Objectives and Evaluations Carried out in Chinese and American Universities

In 1995, China began to gradually promote the curriculum reform of general education, itself experiencing a process of continuous improvement. The evaluation of general education is still in the initial stage, there is no widely recognized general education evaluation system, or even no real general education evaluation plan. Only a few scholars in China have discussed the evaluation of general education theoretically. The evaluation of general education is the most difficult kind of evaluation project because it is scattered and lacks clear focus. Colleges and universities have their own clear mission, and the goal of general education is to lay down in greater detail on the basis of their mission what kind of person they want to make students and what kind of abilities they want to make students. Because different universities have different training objectives, the general education objectives are expressed differently. Schools with good general education evaluation generally have a clear description of their general education goals, but not all colleges and universities have a corresponding evaluation system.

At present, there is no accurate explanation of general education ability in domestic universities. Some researchers believe that it refers to a common knowledge and basic abilities that are common in different disciplines and industries, including expression ability, critical thinking ability, mathematical reasoning ability, ability to integrate and obtain information. It has the characteristics of universality, mobility, durability and integration. The importance of general ability is not self-evident, but the extent to which students in general education courses have these general ability is not easy to determine. This is because the general ability is a “soft skill”, and it is very difficult to directly measure and evaluate its ability. How to scientifically evaluate it is a major difficulty at home and abroad, and it is difficult to conduct quantitative assessment like evaluating many professional knowledge and skills. At present, domestic colleges and universities lack effective evaluation means and approaches in this regard.

FSU attaches great importance to the cultivation of students' general ability. The university requires students to “effectively identify, obtain and use information” through general education courses. Evaluating information using critical reasoning skills; Interpret and use quantitative data and other information to effectively understand and solve real-life problems; Think and act creatively in problem solving; Communicate effectively with others; Function effectively as a team member. In short, FSU requires students to have critical thinking skills, oral communication skills,

written communication skills, mathematical reasoning skills and other general skills. The university believes that the acquisition of these general abilities can contribute to the success of students' college education and also greatly benefit their future work and life. The cultivation of students' general ability is organized and implemented by the general education committee under the school, and the teachers of the general education course are the main body to carry out the cultivation and assessment of students' general ability. FSU believes that the primary purpose of general education evaluation is to promote the development of teaching, and the ultimate goal is to promote the development of students.

2. The Four-Factor Model for the Assessment of College Students' General Ability and the Design of the Questionnaire

The goal of general education in our school is to cultivate the whole person with positive attitude, independent thinking and broad vision, including the cultivation of knowledge structure, core competence and thinking habit. The questionnaire is designed to evaluate the effect of general education from the perspective of ability. The design of the questionnaire mainly refers to the reference books of measurement, such as new huaxin quality dictionary, and the classic works of general education, such as harvard red book of general education. Since it is very difficult to measure the ability directly, we use self-report method to measure the subjects' general ability, and use likert 6-point scale to score. For example, give a descriptive statement and ask the subject whether the statement fits in with their actual situation. The choices range from 1 to 6. 1 is strongly disagree, and 6 is strongly agree. On the basis of these targeted and practical literature, the following four abilities were determined to be evaluated: critical thinking and innovative thinking, expression and communication skills, analytical and comprehensive information skills, and teamwork skills. There were 48 questions in the final version of the questionnaire design, and the calculation method of the four ability indicators was as follows: first, the scores of the items in each scale were summed up, and then standardized into variables with mean value of 0 and variance of 1.

Critical thinking is the process of rational, reflective thinking that goes into deciding what to believe and what not to believe. Critical thinking is the ability and quality to judge the correctness of an argument by questioning its assumptions and logical processes. This questionnaire mainly focuses on the self-confidence of critical thinking, the ability of analysis and the ability of systematic thinking, and designs a series of statements describing specific behaviors. Innovation ability refers to the ability to constantly improve working and learning methods, to come up with new theories, new methods, new ideas and new measures of social and economic value, and to create new products or produce new results, unfettered by stereotypes and past experience. Specifically, we break down the ability to innovate into four aspects: curiosity, flexibility, openness to new ideas, and willingness to question. A total of 15 problems were designed.

Expression and communication skills include verbal and written communication skills. Oral expression ability refers to the ability to express one's own opinions, opinions or opinions appropriately, smoothly and accurately in different occasions by means of language, body movements, etc. Written expression ability refers to the ability to convey information accurately, clearly, concisely and completely in written form according to different work requirements. According to the above definition, this questionnaire designed 13 questions from the four dimensions of logic of expression, good at expressing complex problems, good at choosing appropriate expression methods according to audience/reader when expressing, and good at writing.

Analysis and comprehensive information ability refers to the ability to master a variety of information integrated together, with a certain standard to master the information for analysis, and for effective identification. Its behavioral characteristics include good grasp of key, establishment of judgment standards, and ability to respond in a timely manner. In this questionnaire, 9 questions were designed to describe the behavioral characteristics related to judgment ability from different perspectives, and the subjects were asked to fill out the questionnaire.

Team cooperation ability refers to the ability to build on the basis of the team, give play to team spirit, complement and help each other to achieve the maximum work efficiency of the team. For team members, they should not only have individual ability, but also the ability to do their best in different positions and coordinate and cooperate with other members. This questionnaire was designed with 11 questions.

To sum up, there are 48 questions in the final version of the questionnaire design. The calculation method of the four ability indicators is as follows: first, add up the scores of the items in the respective scales (the reverse scoring questions have been changed to the forward scoring questions), and then normalize them into variables with mean value of 0 and variance of 1.

3. Questionnaire Test and Results

3.1 Questionnaire Evaluation Results

At the end of November 2018, questionnaires were sent to undergraduates of grade 2018 in a university in xi 'an, covering students of various majors in nine schools of the university, and 2,673 valid questionnaires were collected. The subsequent reliability and validity tests were all based on the data of these samples. In the process of answering the questions, the students did not ask any questions about the language expression and the understanding of the questions.

As can be seen from the data charts, more than 67% undergraduates of grade 2018 in the test university believe that they have basically acquired the abilities of critical thinking and innovative thinking, expression and communication, analysis and comprehensive information, and teamwork through part of the general courses.

3.2 Reliability Analysis

Reliability refers to the degree of consistency or stability of a measurement. Consistency is the test of whether items on the same scale measure the same concept or trait. Stability is mainly used to test whether the same questionnaire is used to repeatedly measure the same group of subjects at different times and whether the measured results are consistent. This questionnaire test did not repeat the measurement many times, mainly from the internal consistency to investigate the reliability of the questionnaire. The index used is the Cronbach Coefficient. This method compares the data of any item in the questionnaire with the data of all other items and is very cautious about the internal consistency estimation of the scale. In this study, SPSS software was used to calculate the Alpha value of each subscale and the Alpha value of the whole set of questionnaires. After some instability problems were deleted, the final Alpha coefficient reached a very good level. Table 1 reports each set of scales and Alpha values in detail.

Table 1. Cronbach's Alpha coefficient

	Cronbach's Alpha	The title number
A. ability to think critically and creatively	0.857	15
B. presentation and communication skills	0.880	13
C. Ability to analyze and synthesize information	0.923	9
D. teamwork ability	0.916	11

An Alpha value of 0.6-0.8 indicates good reliability, while an Alpha value of 0.8-1 indicates very good reliability. The closer it is to 1, the better the reliability is. Therefore, the questionnaire has good reliability in terms of overall consistency and subscale consistency in the above four dimensions.

3.3 Validity Analysis

The validity discussed in surveying refers to the degree to which the measuring tool can accurately measure the trait to be measured. Validity can be divided into three main types: content validity, criterion validity and structure validity.

Content validity, also known as logical validity, refers to the suitability and consistency between the measurement object and the measurement content. Content validity is often evaluated by combining logical analysis with statistical analysis. The logical analysis is usually made by the researcher or expert according to their own professional knowledge and experience to judge whether the topic meets the purpose and requirements of the measurement. The design of this questionnaire carefully refers to the existing questionnaires and all kinds of researches on general education and ability development, which are logical to some extent, but lack of expert argumentation, which will be further improved.

Criterion validity is to measure the validity of the newly designed scale by taking the existing scale as a criterion. If the correlation between the measurement results of the new scale and the measurement results of the existing scale is between 0.4 and 0.8, the validity of the new scale is considered to be ideal. The key to this test is to have a set of scales that are valid and can be used as criteria for reference. Since the general education ability development questionnaire designed by us has not found a questionnaire that can match the coverage, we cannot test the validity of the criterion.

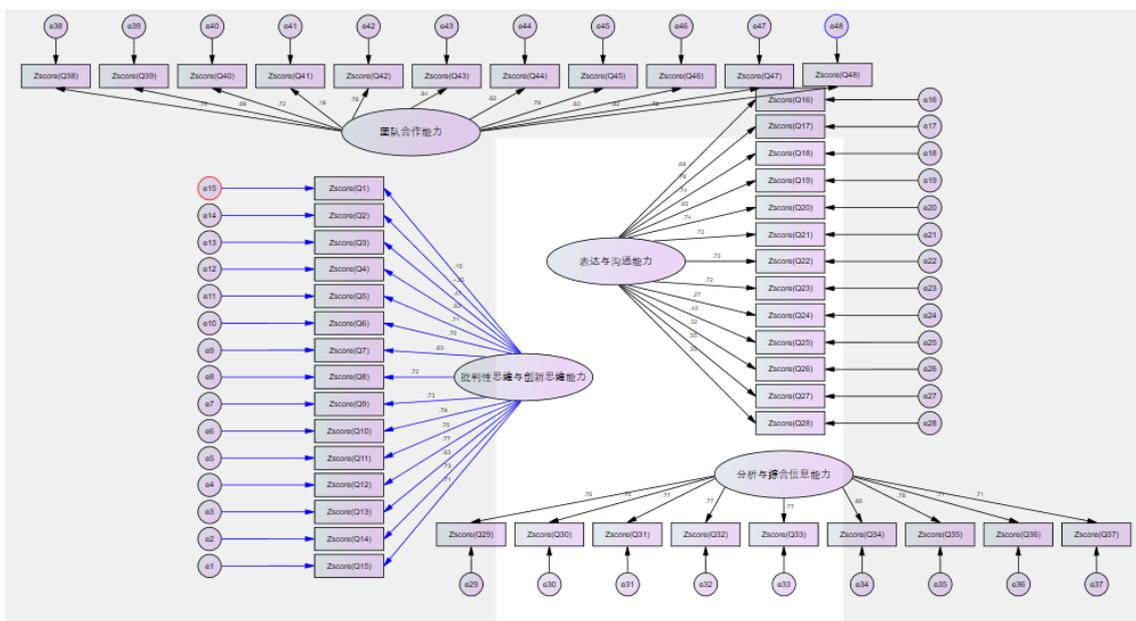


Figure 1. diagram

Structural validity, also known as theoretical validity, refers to the degree to which the measurement tool reflects the internal structure of concepts and propositions. In other words, if the questionnaire results can measure their theoretical characteristics and make the survey results consistent with theoretical expectations, the data will be considered to have structural validity. Structural validity is usually tested by confirmatory factor analysis and correlation coefficients between potential variables. This project mainly uses Amos software to conduct confirmatory factor analysis and correlation coefficient between potential variables to test the structural validity. Figure 1 shows the relationship between the four general abilities (potential variables) and measurable variables. Table 2 shows the standardized load coefficient estimates for each item, and it can be seen that most estimates are above 0.5 and are significant at the 0.001 level (significance $P < 0.01$, denoted by ***). Table 3 reports the goodness of fit test results of the questionnaire developed for the verification factor analysis of college students' general ability assessment.

Table 2. load factor estimates for confirmatory factor analysis

		Estimate	S.E.	C.R.	P
ZQ15 <---	Ability to think critically and creatively	1.000			
ZQ14 <---	Ability to think critically and creatively	1.035	028.	36.482	* * *
ZQ13 <---	Ability to think critically and creatively	886.	028.	31.262	* * *
ZQ12 <---	Ability to think critically and creatively	1.083	028.	38.158	* * *
ZQ11 <---	Ability to think critically and creatively	1.061	028.	37.408	* * *
ZQ10 <---	Ability to think critically and creatively	1.117	028.	39.369	* * *
ZQ9 <---	Ability to think critically and creatively	1.023	028.	36.076	* * *
ZQ8 <---	Ability to think critically and creatively	1.008	028.	35.560	* * *
ZQ7 <---	Ability to think critically and creatively	881.	028.	31.107	* * *
ZQ6 <---	Ability to think critically and creatively	983.	028.	34.691	* * *
ZQ5 <---	Ability to think critically and creatively	1.005	028.	35.458	* * *
ZQ4 <---	Ability to think critically and creatively	882.	028.	31.149	* * *
ZQ3 <---	Ability to think critically and creatively	665.	028.	23.499	* * *
ZQ1 <---	Ability to think critically and creatively	205.	028.	7.233	* * *
ZQ17 <---	Presentation and communication skills	1.124	031.	36.707	* * *
ZQ18 <---	Presentation and communication skills	1.067	030.	35.004	* * *
ZQ19 <---	Presentation and communication skills	942.	030.	31.166	* * *
ZQ20 <---	Presentation and communication skills	1.073	030.	35.188	* * *
ZQ21 <---	Presentation and communication skills	1.041	030.	34.210	* * *
ZQ22 <---	Presentation and communication skills	1.059	030.	34.760	* * *
ZQ23 <---	Presentation and communication skills	1.039	030.	34.145	* * *
ZQ24 <---	Presentation and communication skills	388.	030.	13.155	* * *
ZQ25 <---	Presentation and communication skills	605.	030.	20.362	* * *
ZQ26 <---	Presentation and communication skills	467.	030.	15.781	* * *
ZQ27 <---	Presentation and communication skills	797.	030.	26.581	* * *
ZQ28 <---	Presentation and communication skills	471.	030.	15.913	* * *
ZQ36 <---	Ability to analyze and synthesize information	1.003	028.	35.385	* * *
ZQ35 <---	Ability to analyze and synthesize information	1.092	028.	38.463	* * *
ZQ34 <---	Ability to analyze and synthesize information	1.124	028.	39.586	* * *
ZQ33 <---	Ability to analyze and synthesize information	1.078	028.	37.988	* * *
ZQ32 <---	Ability to analyze and synthesize information	1.086	028.	38.279	* * *
ZQ31 <---	Ability to analyze and synthesize information	1.089	028.	38.387	* * *
ZQ30 <---	Ability to analyze and synthesize information	1.072	028.	37.772	* * *
ZQ29 <---	Ability to analyze and synthesize information	1.067	028.	37.616	* * *
ZQ40 <---	Teamwork ability	1.029	029.	35.863	* * *
ZQ41 <---	Teamwork ability	265.	028.	9.347	* * *
ZQ42 <---	Teamwork ability	1.115	029.	38.793	* * *
ZQ43 <---	Teamwork ability	1.190	029.	41.308	* * *
ZQ44 <---	Teamwork ability	1.174	029.	40.769	* * *
ZQ45 <---	Teamwork ability	1.124	029.	39.072	* * *
ZQ46 <---	Teamwork ability	1.188	029.	41.246	* * *
ZQ47 <---	Teamwork ability	1.172	029.	40.705	* * *
ZQ48 <---	Teamwork ability	1.109	029.	38.590	* * *

Table 3. fitting index

Fitting index	F test P values	χ^2 (df)	X/df2	RMSEA	CFI	NFI	IFI	AIC
The results of	0.000	23478.525 1080	21.748	0.088	0.729	0.720	0.729	2352.000

4. Conclusion

It can be seen from the above analysis that the reliability and validity of the general competency assessment questionnaire constructed according to the general education theory and the ability development theory are very good. In the limited questionnaire test conditions and 59.4% of the sample size, also achieved good test results. The questionnaire design, however, after all, from the perspective of students' personal feeling and evaluation of indirect access to information, still need to improve the follow-up questionnaire development and data analysis, one is hoped to be performed within a larger sample measured again on the validity, the second is, by student interviews, directly to examine students' comprehensive ability. This research needs further exploration and improvement.

5. Reference and Inspiration

First, on the basis of investigation and research, it is clear that each course should focus on cultivating general ability and form a consensus with students. It is suggested that teachers should carry out surveys on industrial enterprises and other employers from time to time in the future to further understand their requirements on graduates' comprehensive quality and vocational general ability. Teachers should re-examine the content structure of the course and the training objectives of the course to make clear which general ability the course focuses on.

Second, we should innovate the mechanism and form to strengthen the practical training of students' general ability. For a long time, the training and testing of students' general vocational ability are often carried out in the form of classroom discussion, project demonstration, impromptu speech, oral test, written test, etc. But in general, due to the influence of large class teaching and other factors, the courses give students less opportunities and times of practical training. It is suggested that on the basis of trying small class teaching, effectively extending students' general ability training from classroom to extracurricular through project tasks.

Third, according to the needs of students' general ability training, reform the content of current students' practical training programs. For example, teachers can ask students to read the designated original literature according to the teaching objectives and the contents of the learning unit, and cultivate students' general ability through quantitative reading of the original literature, questioning, inquiry, discussion, and irregular testing.

Fourth, the gradual reform is adopted in the assessment of students' general ability. General capability assessment is a complex project, evaluating the general ability of students should be based on the circumstance of the university, inherit the former feasible evaluation measures, through the designed scientific evaluation and the score, adhere to the direct evaluation and indirect evaluation.

References

- [1]. Li manli. Zhang yu. Ouyang jue. Development of questionnaire for assessment of college students' general literacy ability. Education research of tsinghua university, 2012,(4).
- [2]. Dai jing. The reference and enlightenment of FSU's evaluation of general education ability -- taking "the foundation of university humanities" and "American history" as examples. Journal of changsha college of civil affairs, 2016(12).

- [3]. Li manli. Zhang yu. Ouyang jue. Research on implementation effect of college students' general education curriculum. *Research on education development*, 2014(13).
- [4]. Liang meiyi. Mission and practice of general education of the Chinese university of Hong Kong. *Journal of the national school of administration*, 2005(10).
- [5]. Tang chunxia. Overview and enlightenment of the teaching quality control system of ferris state university. *Changsha vocational college of civil affairs*, 2013,(4).