JOURNAL OF BALTIC SCIENCE EDUCATION
ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

MULTI-INDEX AND HIERARCHICAL COMPREHENSIVE EVALUATION SYSTEM FOR TRAINING QUALITY OF SCIENCE AND ENGINEERING POSTGRADUATES

Peitong Duan,
Huijun Niu,
Jiawen Xiang,
Caiqin Han

Abstract: *It is essential to establish a multi-dimensional postgraduate quality evaluation system for student assessment and training. This study aimed to explore the construction of the multi-index and hierarchical comprehensive evaluation system for postgraduate training in science and engineering based on the Context, Input, Process, Product (CIPP) model using Analytic Hierarchy Process. It involved 756 postgraduates in physics and engineering who were randomly selected via the Internet. Data were collected from the questionnaire about postgraduates' basic information. After collection, Factor Analysis was used to verify the rationality of the design of second-level and third-level indicators, and adjust the corresponding weights. On this basis, Cluster Analysis was used to classify the training quality of the postgraduates based on their scores on academic ability, basic quality, and social ability indicators. The results revealed that the index system includes 4 first-level indicators,12 second-level indicators and 36 third-level indicators, and different weights being assigned to the indicators according to their influence on the training quality of postgraduates in science and engineering. This study also provides some reference for the quality of science and engineering postgraduate training in Chinese universities by proposing relevant measures, which could be interesting also for international audience.*

Keywords: *Analytic Hierarchy Process, CIPP model, multivariate statistical analysis, postgraduate quality training*

Peitong Duan, Huijun Niu, Jiawen Xiang, Caiqin Han
Jiangsu Normal University, China

Introduction

Postgraduate education shoulders the important mission of high-level talent training and innovation and creation, which is an essential cornerstone of China's development and social progress (State Council of China, 2020). Postgraduate education plays a key role in cultivating talent with cutting-edge knowledge, innovative abilities, critical thinking, and high-level skills as part of the national innovation strategy (Commission on the Future of Graduate Education, 2015). Quality evaluation of postgraduate education is a comprehensive examination for the postgraduates, which has a positive impact on the management mode of postgraduates, the construction of the supervisory teams, scientific research, and classroom teaching. Different countries have different requirements for the quality training of postgraduate education. For example, postgraduate education at the United States universities must be validated by two authoritative assessment organizations, the Council on Postsecondary Accreditation (COPA) of the United States and Federal Ministry of Education, Office of Qualifications and Institutional Assessment, before bachelor's, master's, and doctoral degrees can be awarded. Russia has a complete system for assessing the quality of postgraduate training in higher education. The Japanese government has enacted a series of decrees to assess postgraduate education, which effectively ensures the quality of higher education. The German postgraduate education system has always put a high value on quality, and it is very strict regarding degree awarding and teaching. Each university in Britain has formulated strict regulations on the awarding of degrees. Regardless of the differences in national conditions, modern postgraduate education in the world today values the quality of degree conferral and the evaluation of postgraduate training (Zhang & Guo, 2009).

https://doi.org/10.33225/jbse/22.21.408

In recent years, the number of people applying for postgraduate school in China has been increasing year by year. Approximately 510,000 more people applied for postgraduate school in 2020 than 2019, and 360,000 more people applied in 2021 than 2020. While in 2022, 4.57 million people took the "Unified National Graduate Entrance Examination", an increase of 800,000 over 2021 (Ministry of Education, 2020a). Growing demand for postgraduate studies, driven by social and cultural factors, has forced Chinese universities to rapidly expand postgraduate degrees in a short period of time, making it difficult for universities to accommodate both quality and growth demands. But the school system has to develop students' specific literacies, including science, technology and engineering literacy in the frame of STEM literacy (Cencelj et al., 2020). Furthermore, only about 23% of Chinese students are successful after an entrance examination and interview (Ministry of Education, 2019). The Chinese postgraduates in science and engineering are a group of high-level talent who bears the double mission of high-end talent supply and science and technology innovation, which is crucial for helping the development of new engineering. Therefore, it is necessary to reconsider the goals, models, and content of postgraduate education (Quality Assurance in Postgraduate Education, 2010).

Since the beginning of 21st century, the development of politics, economy and social culture has exerted more and more influence on the higher education environment. The changing environment has, to some extent, influenced attitudes, goals and strategies of higher education training. In order to ensure that the quality development of higher education is in line with student development and academic standards, United Nations Educational, Scientific and Cultural Organization (UNESCO, 2004) and International Network of Quality Assurance Agencies in Higher Education (INQAAHE, 2005) have begun to attach importance to the use of quality assurance evaluation methods. The evaluation process is no longer limited to administrators but is increasingly involving other practitioners and no longer a single factor evaluation involving only school policies or student performances. Zaki (2020) evaluated the teaching performance of teachers majoring in English teaching by methods of quantitative evaluation models and online questionnaires. It pointed out that Quality Assurance Bureau (QAB) plays a positive role in improving teaching quality. Okpa et al. (2020) used the Student Participation Quality Assurance Management Questionnaire (SPQAMQ) to evaluate the student participation in decision-making and management aspects of the higher education evaluation process, pointing out the importance and necessity of student participation in evaluation and management. It can be seen that the evaluation of student educational quality should be a multi-subject and multi-factor evaluation system.

The main methods of quality evaluation are internal evaluation, which evaluates exams, defenses, and learning engagement, and external evaluation, which evaluates externally responsible institutions. Rosa et al. (2016) surveyed and analyzed the attitudes of faculty and general faculty in universities from the perspective of internal evaluation (Internal Quality Assurance) on the four aspects of quality evaluation: goals, culture, obedience, and consistency, and concluded that timely internal evaluation helps to promote faculty and general staff awareness and understanding of the teaching and learning process and to improve critical and reflective evaluation. Goran and Anna (2019) conducted external quality evaluation with the help of the Institutional Evaluation Program (IEP) and used the results of this evaluation as a basis for improvement of peer review. Both evaluation approaches emphasize not only the importance of quality assurance evaluation, but also the feedback and retrospective nature of evaluation and the significance of teaching quality development.

Based on the development trend of evaluation with multiple factors and perspectives, some scholars have proposed multi-dimensional evaluation models. Izci et al. (2020) emphasized the consistency of evaluation indicators with the learning process and established an evaluation model including four indicators, that is, teacher's perspective, task characteristics and details, authentic implementation, and authentic training. Markus and Philipp (2018) designed three least squares regression models for evaluating administrators, individual functions, and integrated perspectives to verify their validity by evaluating the goals, strategies, and methods of higher education in this school. These evaluation approaches have taken into account the evaluation from different stages of teaching and learning, but the evaluation elements and indicators are still relatively few. Moreover, quality evaluation is a continuous and changing process, and various indicators need to be constantly reviewed and considered, and less attention has been paid to the revision and calibration aspects in these evaluation models.

For China, with the rapid development of domestic science and technology and the emergence of new fields and technologies, the demand of society for innovative and compound talents who can quickly adapt to the development of science and technology and industrial reform is increasing (Ye et al., 2021). The establishment of an enterprise-centered, market-oriented, in-depth integration of technology innovation system is one

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

of the trends of science and engineering postgraduate education in China. The postgraduate training mode in China has shifted from the traditional and single disciplinary teaching mode to the ''dual-track" mode of integrating multiple educational resources. The educational administrators of various universities are combining interdisciplinary and cross-disciplinary teaching contents and establishing and perfecting the "dual tutor" employment system. The "dual tutor system" refers to the selection and recruitment of one supervisor each from the enrollment unit and the relevant practice cultivation unit of the university during the cultivation process of professional degree master students, and the guidance of the supervisor of the enrollment unit is the main and the practice base. The main responsibilities of the on-campus supervisor are to cultivate students' learning ability and enhance their development space, while the off-campus supervisors' aim is to guide students to improve their practical operation ability, increase their practical cognitive ability, and adapt to their future work. On the basis of the continuous deepening of the quality assurance evaluation system of higher education in China, in order to adapt to the diversified training strategies of Chinese universities, meet the policy orientation of innovative resource sharing and complementary advantages, and improve the evaluation efficiency, the National Education Commission of China (NECC) puts forward the requirements of multi-factor, multi-angle and quantitative evaluation on the evaluation of the quality of science and engineering postgraduate training.

CIPP Model

CIPP is an evaluation model covering the early, middle, and later period of activities. It was first proposed by Stufflebeam, and has the characteristics of decision, guidance, and feedback (Stufflebeam & Shinkfield, 1966). After CIPP model was proposed, Ohara and Pickcard (1985) built it into a simulation model based on computer programming ideas, which laid a foundation for its wide application. Subsequently, CIPP model has been widely applied in education focusing on the evaluation of school programs, school policies, curriculum implementation and so on.

In the aspect of school projects evaluation, Jumari and Suwandi (2020) described to the evaluation of the CIPP model on early childhood friendly programs and argued that the results of the CIPP evaluation are expected to contribute to Islamic educational institutions as best practices for implementing child-friendly school programs. Prasetiyo et al. (2020) found that the application of the CIPP model to an environmentally friendly program is a fully inclusive school. The results showed that the program was successful in disseminating motivation and environmental care to students, laying the foundation for sustainable development. Agustina and Mukhtaruddin (2019) described to the use of CIPP model for evaluation in Integrated English Learning (IEL) program, provided useful ideas for improving the program, motivating teachers to do better and more in the teaching and learning process. Hurmaini (2015) conducted an evaluation of evaluation of the campus social practice program, made some improvements to address the shortcomings exposed by the context, input, process, and product evaluation.

Meanwhile, some researchers have evaluated policies using the CIPP model, such as Jamil and Iqbal (2020) who evaluated the implementation of a vocational training program for women's entrepreneurial skills, and the findings recommended revising curriculum content, providing a better environment, and developing effective teaching strategies. Edwards (2016) used the CIPP model for the integration of agricultural literacy development education policy while enabling change agents to support teacher participants and encourage them to use agriculture as a context for teaching and learning.

In terms of evaluating learning strategies, the CIPP model was used to assess conceptual framing of learning styles, which had an impact on institutional assessment and continuous improvement in higher education institutions (Chinta et al., 2016). Tokmak et al. (2013) used the CIPP model to evaluate and redesign an online master's degree program. The results showed that most students were satisfied with the new version of the program.

On the part of curriculum evaluation, Al-Shanawani (2019) used the CIPP model to evaluate the kindergarten self-learning curriculum and suggested that all aspects of preschool children's education should be studied and evaluated on an ongoing and comprehensive basis. Akpur et al. (2016) evaluation of the design of the English preparatory curriculum at Yildiz Technical University, the findings revealed that teachers and students were generally positive about the curriculum. Powell and Conrad (2015) utilized the CIPP model to develop, among other things, an integrated service-learning component for a university health course, and the

https://doi.org/10.33225/jbse/22.21.408

results indicated that students felt that service-learning enhanced health promotion knowledge, reinforced topics, and increased student ownership of the course. As seen from the current achievement, the CIPP model has been widely used in the field of education and achieved good results in teaching evaluation. Therefore, it is important to build a quality evaluation system for postgraduate training in science and engineering based on CIPP model.

Evaluation System Construction

In the evaluation of training quality of science and engineering, postgraduates based on CIPP model, students' enrollment basis, training objectives and training mechanism are part of the training background. It is a part of context evaluation that can be carried out at the beginning of semester. The input of teacher resources, external resources and policy support of universities in cultivation, are the basic resources and conditions of teaching implementation. It is a part of input evaluation that can be carried out at the beginning of semester. The types of courses, the richness of activities and the specific implementation are the main parts of teaching process, parts of process evaluation, which can be carried out once a month. The basic requirements, academic ability and social ability of the students are three parts of product evaluation. It can be carried out in the mid-term and the final term because of the parts of product evaluation from which we can observe that the evaluation indexes of science and engineering postgraduate quality training are consistent with the four elements of CIPP model, which can be used to construct the evaluation system. Figure 1 shows the flow chart of establishing the weight of engineering postgraduate quality index based on CIPP model.

First, the main factors related to the objectives were extracted from the teaching activities according to the requirements of the engineering postgraduate training program and CIPP evaluation model, and the first-level, second-level and third-level indicators were established. The weights of each first-level and second-level index are calculated by using Analytic Hierarchy Process. Next, as a result of the third-level indicators, a questionnaire was developed based on expert opinions and training program requirements, and 756 postgraduates were collected for empirical analysis. After collecting the data, Factor Analysis was carried out in order to isolate common factors to assess the validity of the content of the second and third level indicators, as well as the weights of the second-level indicators. Based on the extraction of the common factors, the Bx indicator weights and the scores of the student samples on the common factors were calculated. The Bx index weights and the subjective weights of the Bx indexes are used to determine comprehensive weights of the Bx indexes, after which the formula for calculating the scores using the CIPP evaluation model is derived from the Ax index weights and the comprehensive weights of the Bx indexes in order to construct the final evaluation system for postgraduates in science and engineering. Finally, based on this, Cluster Analysis was used to classify the quality of students' training based on their scores on second-level indicators in order to understand the differences between students and their performance on each influencing factor.

In this model, a total of three analysis methods were selected, namely, Analytic Hierarchy Process, Factor Analysis, and Cluster Analysis. Among them, the Analytic Hierarchy Process includes four steps to calculate the weight values of each first-level and second-level indicator (Saaty, 1993). Firstly, constructing the structure of the hierarchical order of the indicator system, it was divided into three levels, namely, the target level, the criterion level, and the implementation level. Secondly, the expert discussion method was used to determine the judgment matrix as well as the relative importance of each issue using the Delphi method. Thirdly, the relative weights of each layer were calculated, and the judgment matrix was listed according to the importance level of 1-9 (Seal, 1964).

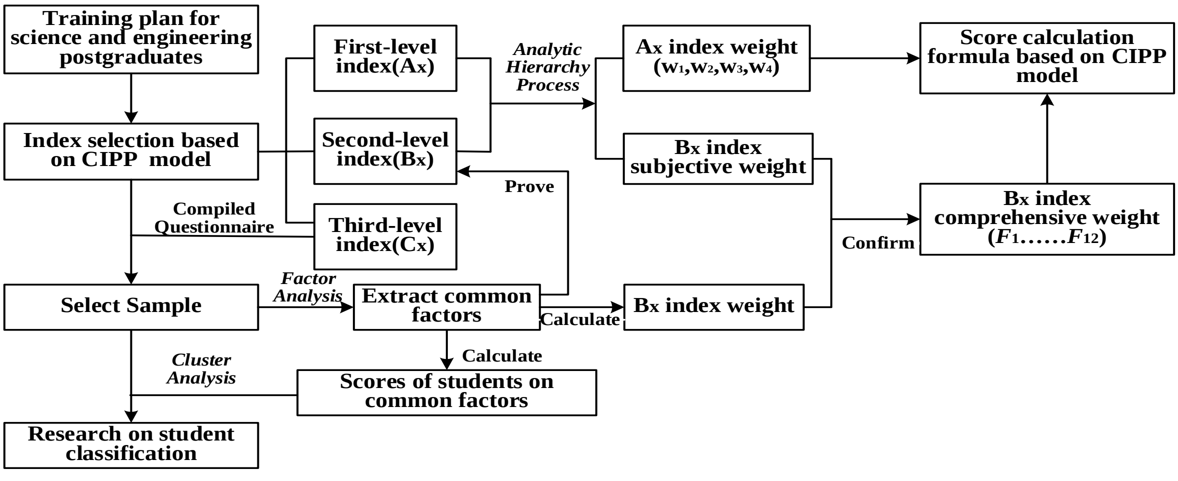
In order to detect whether the selected indicators are suitable for Factor Analysis, KMO (Kaiser Meyer Olkin) and Bartlett's tests were performed using SPSS (Statistical Package for the Social Sciences) 23.0 software, and usually a KMO value greater than 0.5 can be considered for Factor Analysis.

Cluster Analysis includes hierarchical clustering, two-step analysis, and K-means clustering (Jang & Hitchcock, 2021). Since we were not sure how many specific groups to classify in this study, we chose the scheme of using a combination of systematic clustering and K-means clustering, firstly using systematic Cluster Analysis to determine the number of groups, and then using K-means clustering to obtain specific grouping information.

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Figure 1
Process of Establishing the Quality Index Weights for Science and Engineering Postgraduates



Research Focus

Based on the above research, this current study constructed an evaluation index system and determined its weights from four aspects: Context, Input, Process and Product, based on CIPP model and combined with the training mode of science and engineering postgraduates, and designed a strategy to evaluate the quality of science and engineering postgraduates' cultivation from both qualitative and quantitative analysis. Therefore, this study aimed to address the following questions:

1. Is it reasonable to determine the multi-index and hierarchical comprehensive evaluation system for postgraduate training in science and engineering based on the CIPP model? And does each weight need to be adjusted in some way through the results of empirical analysis?
2. Is there any possibility to construct a quantitative evaluation method for the quality training of science and engineering postgraduate based on the CIPP model?

Research Methodology

General Background

Analytic Hierarchy Process is a decision-making method that decomposes the elements always related to decision-making into levels such as objectives, criteria, and options, on the basis of which qualitative and quantitative analysis is performed. Its main feature is the hierarchy of the decision-making process according to the individual's thinking and mental patterns, which allows to summarize multi-factor problems or complex multi-criteria problems into a hierarchical structure. Saaty (1993) pointed out that the hierarchy is a representative structure of complex problems in a multi-level hierarchical structure.

Factor Analysis refers to the study of statistical techniques that extract common factors from a population of variables. Factor Analysis can identify hidden representative factors among many variables. Grouping variables with the same essence into one factor reduces the numbers of variables and also tests the hypothesis that variables are related (Spearman, 1904).

Cluster Analysis refers to the analytical process of grouping a collection of physical or abstract objects into multiple classes consisting of similar objects (Jang & Hitchcock, 2021). It is an important human behavior. The goal of Cluster Analysis is to collect data to classify on the basis of similarity.

Participants

756 postgraduates with a master's degree in physics and engineering were randomly selected as the objects via the internet, including five majors in optics, optical engineering, theoretical physics, condensed matter, and electronic engineering, who volunteered to participate.

https://doi.org/10.33225/jbse/22.21.408

Design of Evaluation Indicators

Based on CIPP Model, context evaluation of science and engineering postgraduates takes postgraduate training plan as the core, including three second-level indexes: enrollment basis, cultivation objectives and training mechanism. Among them, enrollment basis is the assessment of students' professional knowledge mastery, cultivation objectives and mechanism are the assessment of postgraduate training direction, and the latter two indicators are more closely related to the future development of postgraduates. Input evaluation takes educational resources as the core, including three second-level indexes: teacher input, external resources, and organizational guarantee. Among them, the teacher input index is to explore students' satisfaction with the qualification and arrangement of teachers in their schools, and the external resources and organizational guarantee index is to investigate students' satisfaction with the research conditions and research funding support, which is more important because teachers are directly involved in students' learning. Process evaluation takes teaching process as the core, including three second-level indexes, namely multi courses, the richness of teaching activities and practical process. The diversity of courses and teaching activities is the basis for the multifaceted cultivation of students, and the practical process indicators are the evaluation of students' commitment to learning, and these three indicators are of similar importance. Product evaluation takes the postgraduate training results as the core, including the basic quality, academic ability, and social ability. Among them, basic quality indicators mainly evaluate students' academic completion and thesis writing ability, which are the most important compared with the other two indicators.

Accordingly, the quality training evaluation index of science and engineering postgraduates based on CIPP model was designed, and the specific content is shown in Table 1. Using A_1-A_4 to denote 4 first-level indicators, B_1-B_{12} to denote 12 second-level indicators, and C_1-C_{36} to represent 36 third-level indicators. Among these first-level indexes, the content of context evaluation is mainly the basic environment, allocation of educational resources belongs to input evaluation, the main content of process evaluation is the teaching and learning process, and academic achievement belongs to product evaluation. Using the Analytic Hierarchy Process, experts with many years of rich teaching management experience were invited to construct a judgment matrix, the characteristic roots and weights of the matrix were calculated by MATLAB software, on which the consistency test was conducted. The weights of context, input, process, and product evaluation in the first-level indicators were obtained as 0.11, 0.17, 0.34 and 0.38, respectively, which means that if the total evaluation score is 100, the scores of these four evaluation elements are 11, 17, 34 and 38, respectively. The weights are determined in the Appendix 1. It is concluded that the weights of the four indexes in the first-level index were 0.106, 0.167, 0.345, 0.381 and the weight of three indicators in the second-level index for context evaluation were 0.163, 0.296, 0.540. The weight of three indexes for input evaluation were 0.413, 0.327 and 0.260. The weight of three indexes for process evaluation were 0.333, 0.333 and 0.333 and product evaluation weights of three indexes were 0.327, 0.413 and 0.260, respectively. Since there are too many third-level indicators, they don't all have independent weights, and each indicator has the same weight.

Table 1
Evaluation Index and Weight of Engineering Postgraduate Quality Training Based on CIPP Model

First-level index	Second-level index	Third-level index (Measurable quota)
Context evaluation(A1) (0.106)	Enrollment basis(B1) (0.163)	Postgraduate entrance score of specialized courses(C1) Postgraduate entrance interview score(C2) Cross-disciplinary situation of students(C3)
	Cultivation objective(B2) (0.296)	The training of postgraduate scientific research ability(C4) The cultivation of postgraduates' criticism and innovation ability(C5) Interdisciplinary postgraduate training(C6)
	Training mechanism(B3) (0.540)	The construction of industry-university-research collaborative innovation mechanism(C7) The construction of cooperative training mechanism in colleges and universities(C8) The training of interdisciplinary cross-integration(C9)

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

First-level index	Second-level index	Third-level index (Measurable quota)
Input evaluation(A2) (0.167)	Teacher input(B4) (0.413)	On-campus supervisor qualification(C10) Enterprise supervisor qualification(C11) Qualifications of industry peripheral supervisor (C12)
	External resources(B5) (0.327)	Factory in related fields and richness of enterprise resources(C13) Laboratory instrument configuration(C14) Richness of library resources(C15)
	Organizational guarantee(B6) (0.260)	The funding of school education(C16) The reward system of postgraduate learning outcome(C17) The administrative system of postgraduate(C18)
Process evaluation(A3) (0.345)	Multi courses(B7) (0.333)	Basic professional knowledge in this field(C19) Ability to use foreign languages, mathematics and computer software(C20) Social humanities courses, ideological and political courses(C21)
	The richness of teaching activities(B8) (0.333)	Frontier development curriculum(C22) Domestic and foreign academic exchanges(C23) Practice design comprehensive training(C24)
	Practice process(B9) (0.333)	The degree of teacher-student interaction(C25) The use of teaching methods and strategies by teachers(C26) Postgraduate classroom participation(C27)
Product evaluation(A4) (0.381)	Basic quality(B10) (0.327)	Academic integrity(C28) Grade Point Average (GPA) score(C29) Postgraduate dissertation grades(C30)
	Academic ability (B11) (0.413)	Postgraduate publications(C31) Postgraduate patent applications(C32) Honors and awards received at school(C33)
	Social ability(B12) (0.260)	Participate in academic conferences and have results(C34) Participate in other social competitions(C35) Employ and set up a business(C36)

Instrument

A questionnaire about the basic information for postgraduates to evaluate the quality of postgraduate training in science and engineering (Appendix 2). On the basis of the above evaluation indicators, the questionnaire was prepared according to the content of the third-level indexes, which includes 36 items for evaluating students on context, input, process, and product evaluation within the whole questionnaire, including multiple-choice questions that set five options A-E, corresponding to 4-0 points and whether questions. The quality of postgraduate training was evaluated according to the high score, taking the product evaluation as an example, as shown in Table 2.

Table 2
Take the Product Evaluation in the Questionnaire as an example

Product evaluation——Academic achievements(A_4)					
Third-level index (Measurable quota)			Options		
			A (1 point)	B (0 point)	
Comply with academic integrity(C_{28})					
GPA above 3.5(C_{29})					
Participate in academic conferences and have results(C_{34})					
Participate in other social competitions(C_{35})					
Employ and set up a business(C_{36})					
A	B	C	D	E	

https://doi.org/10.33225/jbse/22.21.408

Product evaluation——Academic achievements(A_4)					
Postgraduate dissertation grades(C_{30})	Above 90	80-89	70-79	60-69	Below 60
Postgraduate papers publications(C_{31})	SCI (Q1, Q2)	SCI (Q3, Q4) EI	Core Journals	General Journal	Not have
Honors and awards received at school(C_{33})	Three or more awards	Two awards	Second and above academic scholar-ship	Third academic scholarship	Not have
Postgraduate patent applications(C_{32})	Invention patent granted	Utility model patent and ap-pearance patent granted	software copyright obtained	Invention pat-ent applied	Not have

The product evaluation for postgraduates consists of a total of nine questions, divided into five whether questions and four multiple-choice questions. Among the yes/no questions, there are five dimensions: whether the student has finished the dissertation, whether the GPA of the student is higher than 3.5, whether the student has participated and succeeded in academic conferences, whether the student has participated in other social competitions, and whether the student is employed, which are tests of the basic ability of the student's study. In addition, the four multiple-choice questions examine the students' comprehensive quality ability from the score of the dissertation, the publication of the dissertation, the award and honor of the university, and the patent application. Among them, for Chinese postgraduates, publishing SCI, EI and Chinese core papers as well as applying for invention patents, utility model patent and appearance patent, are significant achievements during the students' study, which is a comprehensive evaluation of the students' academic and comprehensive abilities.

Reliability and Validity

The collected data were entered into SPSS 23.0 software, the information was normalized into dimensionless data, and the reliability and validity of the questionnaire were verified using KMO statistical test and Bartlett's test. In addition, the significance level of the Bartlett's test results was set at $p = .05$. The results of reliability analysis are shown in Table 3.

Table 3
Reliability Statistics of Questionnaire

KMO and Bartlett's Test					
		Context	Input	Process	Product
KMO		.906	.925	.919	.765
Bartlett's Test	*p*	< .05	< .05	< .05	< .05

As seen in Table 3, the KMO values of context evaluation, input evaluation, process evaluation, and product evaluation were .906, .925, .919, and .765, respectively. These analyses and calculations indicated that all items are in positions above .500 and sig values $p < .05$, which means there is a strong correlation between indicators and Factor Analysis can be conducted.

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Research Results

Rationality of Index Design

Factor Analysis was conducted through the experimental data to obtain the variance contribution rate of each factor to the final results. The 12 common factors selected for subsequent analysis, which had a variance contribution greater than 85% or an eigenvalue greater than (or close to) 1, are denoted by F_1-F_{12}. The variance contribution rate of Postgraduates' Academic Achievement Evaluation is shown in Table 4. It showed that eigenvalues of the first two factors were greater than 1, and the eigenvalues of the third factor were approximately 1, which indicated that each part had three factors that could well describe the specific meaning of the corresponding nine indicators. These factors can be regarded as common factors corresponding to the number of second-level indexes, indicating that the setting of third-level indicators is reasonable.

Table 4
Eigenvalue and Variance Contribution Rate of Postgraduates' Academic Achievement (Product) Evaluation

	Eigenvalues of matrix	Percentage of variance (%)	Cumulative contribution value (%)
		Context evaluation	
F_1	6.544	72.706	72.706
F_2	0.781	8.673	81.379
F_3	0.571	6.341	87.720
		Input evaluation	
F_4	6.724	74.708	74.708
F_5	0.594	6.597	81.305
F_6	0.392	4.352	85.658
		Process evaluation	
F_7	6.283	69.811	69.811
F_8	0.662	7.358	77.169
F_9	0.477	5.305	82.474
		Product evaluation	
F_{10}	2.921	32.453	32.453
F_{11}	1.280	14.222	46.675
F_{12}	0.995	11.052	57.727

According to rotation load matrix in the Factor Analysis, the influence of each variable on each factor can be analyzed. The rotated load matrix is the factor load matrix obtained by deriving a deterministic solution from the uncertain solution that gives a meaningful interpretation of the common factors and transforming the initial factor load matrix according to the simple structure criterion. The specific coefficients are shown in Table 5. On the basis of coefficient of each factor on the common factor, the influence degree of each factor on the common factor was judged. In light of the requirements of Factor Analysis, the closer the KMO value is to 1, the more suitable for Factor Analysis, more than 0.8 is also more suitable, 0.6 is more general, and below 0.5 is not suitable for Factor Analysis, so the factors with absolute values of coefficients greater than 0.5 are selected and the common factors are extracted and compared with evaluation content of the second-level index to test rationality of the design of the second-level index.

Taking Factor Analysis rotation load matrix of the product evaluation as an example, as can be seen that the coefficients of common factor F_{10} in " paper publication""patent application" and "national scholarship acquisition" were 0.854, 0.754 and 0.726 respectively, which mainly reflects the students' scientific research ability, innovation and expansion ability, engineering technology and practice ability. Therefore, the common factor F_{10} was named as "academic ability factor". The coefficients of common factor F_{11} on "academic integrity score" "GPA score" and

https://doi.org/10.33225/jbse/22.21.408

"dissertation score" were 0.591, 0.674 and 0.669, reflecting the basic abilities of postgraduates such as knowledge learning, literature reading and thesis writing. Accordingly, the common factor F_{11} was named as "basic quality factor". The coefficients of common factor F_{12} in "academic conference participation and awards""other social competition awards" and "employment and entrepreneurship" were 0.687, 0.600 and 0.704. These awards are mainly based on entrepreneurship competitions, academic communication, and social activities in other fields, reflecting the students' communication and presentation skills, team cooperation ability and the ability to act. Therefore, the common factor F_{12} was named as "social ability factor". As a result, it shows that the name design of three second-level indexes is in line with the data analysis, and the design is more reasonable.

Table 5

Factor Analysis of Postgraduate Academic Achievement (Product) Evaluation Rotating Load Matrix

Third-level index	Component		
	F_{10}	F_{11}	F_{12}
Comply with academic integrity(C_{28})	0.150	0.591	0.264
GPA above 3.5(C_{29})	0.530	0.674	0. 061
Results of postgraduate dissertation(C_{30})	0.605	0.669	-0.002
Situation of papers published(C_{31})	0.854	0.114	0.061
Situation of patent applications(C_{32})	0.754	-0.116	-0.189
Honors and awards received at School(C_{33})	0.726	-0.080	0.078
Participate in academic conferences and have results(C_{34})	0.344	0.176	-0.687
Participate in other social competitions(C_{35})	-0.233	0.374	0.600
Employ and set up a business(C_{36})	0.272	0.254	0.704

In the evaluation system of quality training of postgraduates in science and engineering based on the CIPP model, the second-level indexes were designed to correspond to three factors, and based on this, the third-level indicators that can be easily measured were prepared. Through the Analytic Hierarchy Process, the weights of the second-level indexes were established in the initial establishment of the evaluation system. On this basis, the scores of the experimental samples on each third-level indexes were collected through the questionnaire method, and the common factors affecting the scores of these third-level indexes were identified by using Factor Analysis. The coefficient and naming of third-level index on corresponding common factor are summarized in Table 6.

Table 6

Coefficient and Naming of Third-Level Index on Corresponding Common Factor

Third-level index				Common factor
	C_4	C_5	C_6	
F_1	0.777	0.745	0.862	Alignment of training objectives with student needs
	C_1	C_2	C_3	
F_2	0.562	0.823	0.877	Student's academic foundation factor
	C_7	C_8	C_9	
F_3	0.719	0.771	0.869	Culture mechanism factor
	C_{10}	C_{11}	C_{12}	
F_4	0.635	0.718	0.843	School organization guarantee factor

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

	Third-level index			Common factor
	C_{13}	C_{14}	C_{15}	
F_5	0.724	0.583	0.774	External resources factor
	C_{16}	C_{17}	C_{18}	
F_6	0.816	0.814	0.537	Teacher resources factor
	C_{19}	C_{20}	C_{21}	
F_7	0.720	0.840	0.758	Activity richness factor
	C_{22}	C_{23}	C_{24}	
F_8	0.815	0.856	0.876	Factor of curriculum diversity degree
	C_{25}	C_{26}	C_{27}	
F_9	0.801	0.702	0.982	Practical factor
	C_{31}	C_{32}	C_{33}	
F_{10}	0.854	0.754	0.726	Academic ability factor
	C_{28}	C_{29}	C_{30}	
F_{11}	0.591	0.674	0.669	Quality foundation factor
	C_{34}	C_{35}	C_{36}	
F_{12}	0.687	0.600	0.704	Social ability factor

It shows that the common factors affecting the experimental sample data can be divided into four categories based on the four elements of the CIPP model, with three factors in each category, and the specific contents are consistent with the preliminary construction of the second-level index indicating that the establishment of second-level indexes is more reasonable. For example, in terms of product evaluation, three second-level indexes were designed in the initial construction of second-level indexes, and through Factor Analysis, it is found that the common factor of "participation and awards in academic conferences," "awards in other social competitions," and "employment and entrepreneurship" can be summarized as "social competence factor", which is in line with the previous design.

Weight Adjustment

The scores of the student on the 12 common factors were used as variables and again downscaled, which allowed the scores of each second-level index for the first-level index to be derived as the objective weights of the three second-level indexes. The factor weight of postgraduate quality training evaluation is shown in Table 7.

Table 7
Factor Weight of Postgraduate Quality Training Evaluation

	Corresponding weights of the second-level index		
	B_1	B_2	B_3
Context	0.100	0.291	0.609
	B_4	B_5	B_6
Input	0.414	0.341	0.245
	B_7	B_8	B_9
Process	0.267	0.354	0.379
	B_{10}	B_{11}	B_{12}
Product	0.316	0.436	0.248

https://doi.org/10.33225/jbse/22.21.408

We can see that the scores of each second-level index for the first-level index, which are the objective weights of the three second-level indexes, 0.316, 0.436 and 0.248, respectively. Through comparing the subjective weight (0.327, 0.413, 0.260) formulated by Analytic Hierarchy Process with the objective weight, it is found that the difference was not significant, and the previous subjective judgment was more reasonable. In addition, through the weighing method of objective weight and subjective weight proposed by Song and Wang (2003), the comprehensive weights were defined as the arithmetic mean value of subjective and objective weight.

The weight adjustment of the second-level index of the product evaluation is (0.322, 0.424, 0.254). Using the same method, it can be concluded that the weights of second-level indexes of context evaluation are (0.132, 0.293, 0.575), the weights of second-level indexes of input evaluation are (0.413, 0.334, 0.253), and the weights of second-level indexes of process evaluation are (0.300, 0.344, 0.356).

Quantitative Evaluation Mode and Improvement

After determining the weight of first and second-level indexes of the evaluation system, the initial calculation formula of student quality evaluation score was obtained, assuming that the total score is 100 points. In the following equation, F_{ps} stands for the product evaluation score of the common factors.

$$
\begin{aligned}
F_{ps}=&17\left[0.132\left(C_1+C_2+C_3\right)+0.293\left(C_4+C_5+C_6\right)+0.254\left(C_7+C_8+C_9\right)\right]\\
&+11\left[0.413\left(C_{10}+C_{11}+C_{12}\right)+0.334\left(C_{13}+C_{14}+C_{15}\right)+0.253\left(C_{16}+C_{17}+C_{18}\right)\right]\\
&+34\left[0.300\left(C_{19}+C_{20}+C_{21}\right)+0.344\left(C_{22}+C_{23}+C_{24}\right)+0.356\left(C_{25}+C_{26}+C_{27}\right)\right]\\
&+38\left[0.322\left(C_{28}+C_{29}+C_{30}\right)+0.424\left(C_{31}+C_{32}+C_{33}\right)+0.254\left(C_{34}+C_{35}+C_{36}\right)\right]
\end{aligned}
$$

At this moment, C_i needs to be in the unified range, and the use has limitations. In order to expand the scope of use, data need to be normalized before calculation and become dimensionless data. Therefore, the coefficient of each three-level index on corresponding common factor can be calculated according to existing data through Factor Analysis. For example, the factor score matrix of the product evaluation is shown in Table 8.

Table 8
Factor Score Coefficient Matrix of Postgraduate Academic Achievement (Product) Evaluation

Third-level index	Component		
	F_{10}	F_{11}	F_{12}
Comply with academic integrity(C_{28})	-0.072	0.410	-0.115
GPA above 3.5(C_{29})	0.136	0.200	0.017
Results of postgraduate dissertation(C_{30})	0.084	0.384	0.047
Situation of papers published(C_{31})	0.329	-0.007	-0.037
Situation of patent applications(C_{32})	0.339	-0.164	0.142
Honors and awards received at school(C_{33})	0.329	-0.153	-0.104
Participate in academic conferences and have results(C_{34})	0.059	0.125	0.703
Participate in other social competitions(C_{35})	-0.268	0.509	0.278
Employ and set up a business(C_{36})	0.074	0.180	-0.570

Table 8 shows linear combination relationship between common factors F_{10}, F_{11} and F_{12} and each three-level index, and the scores of F_{10}, F_{11} and F_{12} can be calculated respectively according to the expression of the factor score function, the score equation of first-level index is established and score coefficients of other indicators can also be obtained through investigation and Factor Analysis (Appendix 3).

To sum up, a more reasonable evaluation system of science and engineering postgraduate training quality

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

based on CIPP model is established by Factor Analysis on the existing basis. When evaluating the quality of engineering postgraduate training, we can intuitively understand the situation of postgraduate quality training by comparing the comprehensive score. By analyzing and comparing the scores of each first-level index or second-level index, we can accurately understand the implementation of each measure and training results and put forward effective suggestions.

Student Classification

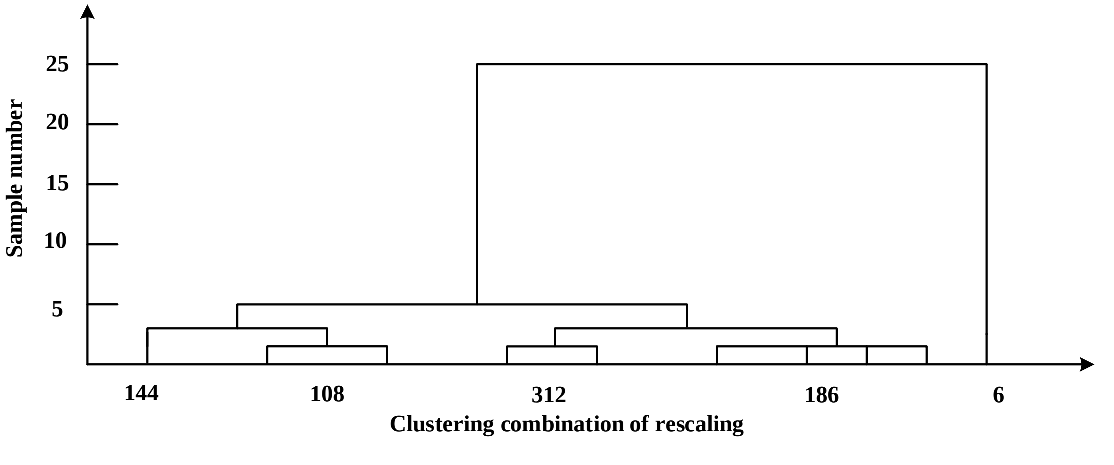
According to the results of Factor Analysis, the scores of each student in academic ability (F_{10}), basic quality (F_{11}), and social ability (F_{12}) are partially presented in Table 9.

Table 9
756 Students' Scores on the Second-Level Index

	F_{10}	F_{11}	F_{12}
1	-1.170	-0.669	0.874
2	1.976	-0.075	1.324
3	1.259	-0.258	-0.696
…	…	…	…
753	-1.185	-0.101	-0.686
754	-1.090	-0.742	-0.784
756	-1.640	-1.749	0.762

Based on the scores of middle school students in F_{10}, F_{11} and F_{12} in Table 9, a pedigree chart describing the classification steps was obtained, as shown in Figure 2.

Figure 2
Hierarchical Cluster Analysis Pedigree Chart



The students can be divided into five categories by observing the pedigree, and then the K-means clustering is used to obtain the grouping data, as shown in Table 10.

https://doi.org/10.33225/jbse/22.21.408

Table 10
Cluster Analysis Results of 756 Students

Factor \ Group (Number)	Group I (144)	Group II (108)	Group III (312)	Group IV (186)	Group V (6)
F_{10} (Academic ability)	-1.117	0.978	0.616	-0.706	-0.952
F_{11} (Basic quality)	-0.877	0.122	-0.094	0.989	-6.930
F_{12} (Social ability)	-0.074	1.700	-0.644	0.088	1.979

It can be seen from Table 10 that there are 144 students in Group I, whose abilities in all aspects are at the average level, and they are slightly deficient in the social ability (F_{12}). There are 108 students in Group II, whose learning degree is relatively good overall, and they have published papers in journals such as SCI. Most of them have patents that are being applied for or have been issued. Some students have applied for national scholarships and actively participated in various national conferences, with stronger abilities in all aspects. There are 312 students in group III, who had a relatively average degree of study, scored above 80 points in dissertations during their master's degree studies, no journal papers published or only published in general journals, and they participated in fewer academic conferences and exchanges. They had mastered basic professional knowledge and scientific research ability. A total of 186 students in the fourth category who have published papers in journals such as EI and Chinese core. Some students have patents being applied for or issued and have a good degree of study. There are 6 students in Group V who had no experience of applying for patents or publishing papers during their study but had participated in competitions held in other fields of society, and received awards, and had performed slightly in terms of academic ability and basic quality.

Discussion

Discussion on Model Application

In this study, according to the CIPP evaluation model and the requirements of postgraduate training in science and engineering, the process of constructing a quality evaluation system for postgraduate training in science and engineering is designed, and the rationality of indicators is tested as well as empirical analysis is conducted by Analytic Hierarchy Process, Factor Analysis and Cluster Analysis. The results showed the reasonableness of the indexes, and then designed the way to evaluate the quality of postgraduates in science and engineering. Based on the analysis results, effective measures to promote the development of postgraduates in science and engineering can be actively proposed, and the specific measures are as follows.

Establish a high level of dual type mentor group faculty. Cultivating innovative talents with practical experience has become one of the basic goals of higher education talent cultivation (Qi & Peng, 2018). Postgraduates are required to have certain basic theories as well as strong practical skills and innovation ability. In Chinese universities, there are a large number of teachers with solid basic theories to serve as on-campus supervisors for postgraduates in science and engineering, but the on-campus supervisors have less practical experience. This can make up for the disadvantage of having less practical experience for the on-campus instructors, and the off-campus instructors can provide rich cases to guide the students. Colleges and universities can establish industry-university-research cooperation with enterprises and appoint enterprise teachers and on-campus supervisors to jointly guide the research and practice of postgraduates.

Reconstruct the existing resource system of postgraduate education. Postgraduate cultivation needs to rely on relevant resources outside the education system and reconstruct the existing postgraduate education resource system (Cheng & Wang, 2010). In order to improve postgraduate training, enterprises with similar majors can be invited to participate in training postgraduates, combining the supervisor's subject with enterprise, performing innovative research in conjunction with the actual situation, cooperating with enterprises in industry-university-research, and establishing practice bases. In addition to providing relevant training, postgraduates should be encouraged to engage in work under the cultivation of cooperative enterprise units in combination with their supervisors' topics. Under the guidance of professionals, enterprise project research and development can also be accomplished. By combining postgraduate innovation training with practical application, these resources can

https://doi.org/10.33225/jbse/22.21.408

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

be integrated into the postgraduate training system. It is also possible to broaden the horizons of postgraduates through the university's foreign exchanges and cooperation, so that they can understand the research dynamics and frontiers of their disciplines and improve their innovation ability.

Improve the construction of postgraduate curriculum system. Take the cultivation objectives and degree requirements as the fundamental basis for the design of the curriculum system, implement the postgraduate cultivation objectives and degree requirements completely, and pay attention to the systematic design and overall optimization of the curriculum system (Ministry of Education, 2020b). It insists on taking ability cultivation as the core and innovation ability cultivation as the focus, broadening the knowledge base and cultivating humanistic literacy, strengthening the integration and articulation of curriculum systems in different cultivation stages, designing the curriculum scientifically.

Enhance the ability of postgraduates to conduct scientific research. Throughout the three years of postgraduate training, basic research training is provided to postgraduates in order to enhance their research abilities (Wu et al., 2019). As a result, on the one hand, all the training time can be utilized, while on the other hand, students' research abilities can be strengthened. On this basis, the project serves as a guide to enable postgraduates to integrate into research activities as soon as possible and maximize the impact of cultivating postgraduates' research abilities.

Discussion on Student Classification Using K-means Clustering

In this study, the students were divided into five groups by first using systematic Cluster Analysis to obtain a lineage diagram describing the classification steps, and then using the K-means clustering to obtain specific grouping information. The results of the Cluster Analysis showed that the students in group II had the best academic ability, basic quality, and social ability among the five groups.

As seen that the mechanism of industry-university-research collaboration has begun to take shape and achieve results in the cultivation of science and engineering in Chinese universities (Yang & Li, 2012). About 60% of students have access to more advanced hardware equipment in this learning environment, have a wide range of opportunity for communication and selection, and can be more clear about the combination of experimental purposes and social needs. At the same time, factories and enterprises can help students to clarify scientific research objectives and find the combination of their own needs and social needs, which improves the applicability of research results. It is conducive to the improvement of postgraduates' interest in scientific research, scientific inquiry ability and hands-on practical ability (Shen & Zhang, 2015).

About 50% of the students have applied for patents and published corresponding papers, among which only 20% of them are the first applicants, and the remaining students mostly apply with their supervisors or off-campus supervisors, which shows that the dual tutorial system has played its role in promoting the development of students' innovative ability and increasing the diversity of supervisors' guidance. At the same time, off-campus supervisors, such as some high-level engineers and experts, can explain emerging technologies and methods used in practical production for students, guide students to carry out research about the social production practice and work needs, and stimulate postgraduates' thinking and innovation through practice.

Nearly 65% of the students have attended more than three academic conferences with their supervsors or project teams, and the university also frequently invites experts in relevant disciplines to hold academic lectures and exchange events. They create a strong academic atmosphere for students, keep them abreast of the latest scientific research achievements, and expand their academic horizons. This helps students to grasp professional knowledge at the macro level and further study issues of interest to them. However, there are still some students who seldom participate in academic exchange activities because of their courses, experiments, or social activities. Therefore, schools should establish a rich online learning platform to bring students closer to academic communication. For example, public teaching platforms, learning resource forums and MOOC courses.

About 60.3% of students' social skills needs to be improved, and the poor integration of academic research and social communication activities is one of the important reasons for this unbalanced development (Barton-Arwood et al., 2005). Therefore, while increasing social exchange activities, schools should integrate academic research into them, such as holding university exchange meetings, providing opportunities for students to communicate with different schools and students of different majors, strengthening cooperation between colleges and departments, and sharing teaching resources. On the other hand, it motivates students to obtain help from researchers in related fields and peer learning possesses a stronger incentive effect, and social skills such as oral and written expressions have also been cultivated on the other.

https://doi.org/10.33225/jbse/22.21.408

Conclusions and Implications

In China, with the continuous expansion of the enrollment scale of postgraduates, the mode of postgraduate training and the quality of training have become more and more widely concerned by the society. However, there are problems in the process of postgraduate training in China, such as unreasonable distribution mechanism and unsound management and evaluation system, so there is a real need for reform of the postgraduate training system. In this study, based on the four elements of CIPP model, the indexes for evaluating the quality of postgraduate training in engineering are designed, and the weights of the first and second level indexes are calculated by using Analytic Hierarchy Process. Secondly, the rationality of the content and weight design of the second-level index corresponding to the product evaluation is proved through empirical analysis, and the weights of the second-level index are adjusted. Finally, based on the results of Factor Analysis, the quantitative evaluation of the quality of engineering postgraduate training was designed. The index system can basically reflect a variety of factors affecting the quality of postgraduate training in science and engineering, which can provide some reference for each postgraduate training and play a realistic role in the development of scientific research abilities and comprehensive abilities of postgraduate students in science and engineering.

However, compared with other disciplines, engineering postgraduates focus more attention on the application of scientific research results, the development of scientific and technological productivity and serving the development of modern factory enterprises, namely, they have to meet both the external demands such as the development of the main industry and scientific and technological innovation, and the internal demands of the development of students' academic and social abilities. With such complex training needs, evaluators need to design a more comprehensive and precise evaluation scheme. The weights of the second-level index given in the paper are influenced both by the subjective perceptions of the evaluator and the selected research sample, and their coefficients are not directly applicable to the evaluation of other schools. In practice, the objective weights need to be recalculated based on the actual situation of the school and combined with the subjective weights to arrive at the final second-level index weights. At the same time, it is necessary to collect information about the experimental samples to verify the degree of matching between the designed second-level and third-level indicators and the actual situation of the experimental samples to make the evaluation more accurate.

Acknowledgements

The authors would like to thank a Project Funded the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) for the generous financial support. This research was founded by Jiangsu Normal University Postgraduate Training Quality Project-Education Teaching Reform and Research Key Project "Research on Quality Evaluation of Engineering Postgraduate Training Based on CIPP Model" (Grant No. JGKTZ201921), and Jiangsu Modern Educational Technology Research Project "Construction and Application of Adaptive Cognitive Diagnostic Evaluation System for Physics Learning" (Grant No. 2021-R-91911). This research was also supported by the Educational Research Project of the Teaching Guidance Sub-committee of Optoelectronic Information Science and Engineering Specialty of the Teaching Guidance Committee of Electronic Information Specialty in Colleges and Universities of the Ministry of Education Fund of P.R. China (Grant No. 2020XGK26 and 2020SYL41).

Declaration of Interest

Authors declare no competing interest.

References

Agustina, N. Q., & Mukhtaruddin, F. (2019). The CIPP model-based evaluation on Integrated English Learning (IEL) program at language center. *English Language Teaching Educational Journal*, *2*(1), 22-31. https://doi.org/10.12928/eltej.v2i1.1043

Akpur, U., Alci, B., & Karatas, H. (2016). Evaluation of the curriculum of English preparatory classes at Yildiz Technical University using CIPP model. *Educational Research and Reviews*, *11*(7), 466-473. https://doi.org/10.5897/err2016.2638

Al-Shanawani, H. M. (2019). Evaluation of self-learning curriculum for kindergarten using Stufflebeam's CIPP model. *SAGE Open*, *9*(1), 1-13. https://doi.org/10.1177/2158244018822380

Barton-Arwood, S., Murrow, L., Lane, K., & Jolivette, K. (2005). Project improve: Improving teachers' ability to address students' social needs. *Education & Treatment of Children*, *28*(4), 430-443. https://www.jstor.org/stable/42899863

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Cencelj, Z., Aberšek, B., Flogie, A., & Kordigel-Aberšek, M. (2020). Metacognitive model for developing science, technology and engineering functional literacy. *Journal of Baltic Science Education*, *19*(2), 220-233. https://doi.org/10.33225/jbse/20.19.220

Cheng, S. H., & Wang, C. Y. (2010). Graduate student training model: Reality and future. *Academic Degrees & Graduate Education,3*, 50-53.

Chinta, R., Kebritchi, M., & Ellias, J. (1987). A conceptual framework for evaluating higher education institutions. *International Journal of Educational Management*, *30*(6), 989-1002. https://doi.org/10.1108/ijem-09-2015-0120

Council of Graduate Schools and Educational Testing Service. (2015). The pathward: The future of graduate education in the United States. http://www.cgsnet.org

Edwards, E.B. (2016). *Dig into Learning: A program evaluation of an agricultural literacy innovation*. ProQuest LLC, Ed.D. Dissertation, Gardner-Webb University.

Everitt, B. (1980). Cluster analysis. *Quality & Quantity: International Journal of Methodology*, *14*(1), 75- https://doi.org/100.10.1007/BF00154794

Goran, D. & Anna, G. (2019). Impact evaluation of external quality assurance by the Cardoso Institutional Evaluation Programme. *Quality in Higher Education*, *25*(2), 208-224. https://doi.org/10.1080/13538322.2019.1635304

Hurmaini, M. (2015). Evaluation on social internship program of Iain Sultan Thaha Saifuddin Jambi students: Using Context, Input, Process and Product model (CIPP model). *Journal of Education & Practice*, *6*(1), 56-62. https://doi.org/10.15548/jt.v22i1.114

International Network of Quality Assurance Agencies in Higher Education. (2005). *Guidelines of good practice*. Wellington.

Izci, K., Muslu, N., Burcks, S. M., & Siegel, M. A. (2018). Exploring effectiveness of classroom assessments for students' learning in high school chemistry. *Research in Science Education*, *5*(50), 1885-1916. https://doi.org/10.1007/s11165-018-9757-0

Jamil, S. B., & Iqbal, Z. (2020). Evaluation of vocational training programs for developing entrepreneurial skills among women in the Punjab. *Bulletin of Education and Research*, *42*(2), 197-218. https://www.researchgate.net/publication/347437390_Evaluation_of_Vocational_Training_Programs_for_Developing_Entrepreneurial_Skills_among_Women_in_the_Punjab_Evaluation_of_Vocational_Training_Programs_198

Jang, J., & Hitchcock, D. B. (2021). Model-based cluster analysis of democracies. *Journal of Data Science*, *10*(2), 297-319. https://doi.org/10.6339/jds.201204_10(2).0009

Jumari, J., & Suwandi, S. (2020). Evaluation of child-friendly schools program in Islamic schools using the CIPP model. *Dinamika Ilmu*, *20*(2), 323-341. https://doi.org/10.21093/di.v20i2.2580

Kittel, J., & Bill, R. (1981). *Unallocated loss adjustment expense reserves implications for property-casualty insurance*. Casualty Actuarial Society Discussion Paper Program.

Li, H., Shen, J., Zhao, D., & Li, Y. (2021). Construction and empirical analysis of an online teaching quality evaluation system in colleges and universities. *Higher Vocational Education Exploration*, *20*(4), 46-54.

Markus, S. & Philipp, P. (2018). Assessing quality assurance in higher education: Quality managers' perceptions of effectiveness. *European Journal of Higher Education*, *8*(3), 258-271. https://doi.org/10.1080/21568235.2018.1474777

Ministry of Education of the People's Republic of China. (2020a). The Ministry of Education once again deployed around the coordination of prevention and control of the epidemic and the organization of the study exams reminded the majority of candidates to carefully prepare for the exams. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202012/t20201223_507096.html

Ministry of Education of the People's Republic of China. (2020b). *Promoting graduate education to a new level*. http://www.moe.gov.cn/srcsite/A26/jcj_kcjcgh/201404/t20140408_167226.html

Ohara, T. (1979). *Computerized simulator based on the CIPP evaluation model*. Amsterdam University Press.

Okpa, O. E., Odigwe, F. N., Emeribe, K. V., & Obule, E. V. (2020). Students' participation in quality assurance management and tertiary institutions effectiveness in Nigeria. *International Education Studies*, *13*(12), 41-47. https://doi.org/10.5539/ies.v13n12p41

Powell, B., & Conrad, E. (2015). Utilizing the CIPP model as a means to develop an integrated service-learning component in a university health course. *Journal of Health Education Teaching*, *6*(1), 21-32. https://files.eric.ed.gov/fulltext/EJ1085221.pdf

Prasetiyo, W. H., Ishak, N. A., Basit, A., Dewantara, J. A., & Muhibbin, A. (2020). Caring for the environment in an inclusive school: The Adiwiyata green school program in Indonesia. *Issues in Educational Research*, *30*(3), 1040-1057.

Qi, F. G., & Peng, J. (2018). Innovation and practice research of cultivation model for master's degree in engineering. *Journal of Higher Education,3*, 31-33.

Quality Assurance in Postgraduate Education. (2010). The changing nature of postgraduate education. https://files.eric.ed.gov/fulltext/ED512345.pdf

Rosa, M. J., Cardoso, S., Videira, P., & Amaral, A. (2016). Internal quality assurance: A new culture or added bureaucracy? *Assessment & Evaluation in Higher Education*, *44*(2), 249-262. https://doi.org/10.1080/02602938.2018.1494818

Saaty, T. L. (1999). Basic theory of the analytic hierarchy process: How to make a decision. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, *1993*(4), 395-423. https://doi.org/10.1287/inte.24.6.19

Seal, H. L. (1964). *Multivariate statistical analysis for biologists*. Methuen.

Shen, Y., & Zhang, Z. (2015). Analysis of the measures top universities in America taken to improve the postgraduates' innovative and practical ability. *Journal of Tianjin University (Social Sciences)*, *17*(5), 454-458.

Song, H. Z., & Wang, Z. J. (2003). The weight between objective weight value and subjective weight value. *Technoeconomics & Management Research*, *3*, 62.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201-292. https://doi.org/10.1037/h0065005

Stufflebeam, D. L., & Shinkfield, A. J. (1966). *Systematic evaluation*. Boston.

https://doi.org/10.33225/jbse/22.21.408

Tokmak, H. S., Baturay H. M., & Fadde, P. (2013). Applying the context, input, process, product evaluation model for evaluation, research, and redesign of an online master's program. *International Review of Research in Open and Distance Learning*, *14*(3), 273-293. https://doi.org/10.19173/irrodl.v14i3.1485

United Nations Educational, Scientific and Cultural Organization. (2004). *Indicators of quality and facilitating academic mobility through quality assurance agencies in the Asia-Pacific region*. Bangkok: UNESCO and Thailand National Accreditation Council.

Wu, X. L., Zeng, J. H., Yue, D. L., & Xu, B. (2019). To promote the reform of postgraduate training mode with engineering practice and innovation ability as the core. *Research in Higher Education of Engineering*, *5*, 103-109.

Yang, Q., & Li, B. (2012). Research on the industry-academia-research cooperation mechanism of local university and college—take Changchun University of Science and Technology as an example. *Higher Education Studies*, *2*(3), 88.

Ye, J., Mi., S., & Bi, H. (2021). Constructing core teaching competency indicators for secondary school science teachers in China. *Journal of Baltic Science Education*, *20*(3), 389-406. https://doi.org/10.33225/jbse/21.20.389

Zaki, M. (2020). Academic quality assurance survey in higher education. *International Journal of Higher Education*, *9*(6), 268. https://doi.org/10.5430/ijhe.v9n6p268

Zhao, J., & Niu, G. (2017). Research on the quality management system of science and engineering postgraduate education in universities based on the process approach. *Journal of Tianjin University (Social Sciences)*, *19*(5), 463-470.

Zhang, H. Y., & Guo, Z. D. (2009). Inspiration of foreign universities' teaching quality assessment experience to China. *Education for Chinese After-school*, *6*, 16.

Appendix 1

Determination of weights using Analytic Hierarchy Process

Firstly, the complex problem is decomposed into different elements, and then the relative importance of different elements in the same level is determined by two-by-two comparison, and then the corresponding matrix is listed according to the 1-9 importance scales, and the relative importance weights of elements in each level are obtained after calculation and random consistency test. is assumed to be the importance score of indicator over indicator , and the judgment matrix is listed through a questionnaire survey of several experts:

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/3 \\ 2 & 1 & 1/2 & 1/3 \\ 3 & 2 & 1 & 1 \\ 3 & 3 & 1 & 1 \end{bmatrix}$$

where and, the eigenvectors d_i= [0.106879, 0.167274, 0.344545, 0.381302] and the maximum eigenvalue λ_{max} of the judgment matrix are solved using the sum-product method, and the eigenvectors are the corresponding weights of each index, and then the consistency difference of the judgment matrix A is tested to see if it is within the acceptable range.

$$C.I. = \frac{(\lambda_{\max} - n)}{(n-1)} \tag{1}$$

$$C.R. = \frac{C.I.}{R.I.} \tag{2}$$

According to the formula (1) and (2), the *C.R.* of the judgment matrix A = 0.017 < 0.1, which means that the matrix is more consistent, and after rounding the obtained data *di* to two decimal places, the weights of the context, input, process and product evaluation in the first-level index are derived, which are 0.11, 0.17, 0.34 and 0.38, respectively, that is, if the total evaluation score is 100, then the four evaluation elements are 11, 17, 34 and 38 respectively. The same method is used to calculate the weights of the second-level indicators Bx.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Appendix 2

Items
1. Postgraduate entrance score of specialized courses
2. Postgraduate entrance interview score
3. Interdisciplinary situation of postgraduates
4. Satisfaction with the school policy on the development of postgraduate scientific research skills
5. Satisfaction with the school policy on the cultivation of critical and creative skills of postgraduates
6. Satisfaction with the school policy on the development of interdisciplinary postgraduates
7. Satisfaction with the construction of industry-university-research collaborative innovation mechanism
8. Satisfaction with the construction of cooperative training mechanism in colleges and universities
9. Satisfaction with the interdisciplinary integration training
10. Satisfaction with the qualifications of tutors in their schools
11. Satisfaction with the qualifications of business tutors in their schools
12. Satisfaction with qualifications of external mentors in the school sector
13. Satisfaction with the richness of factories and enterprise' resources in relevant fields in the school
14. Satisfaction with the configuration of laboratory instruments in the school
15. Satisfaction with the richness of the school library resources
16. Satisfaction with funding for school education
17. Satisfaction with the reward system for academic achievements of school postgraduates
18. Satisfaction with the postgraduate administrative system
19. Satisfaction with the cultivation of basic professional knowledge in this field
20. satisfaction with the cultivation of foreign language, mathematics, and computer software skills
21. Satisfaction with the social and humanities and ideological and political courses
22. Satisfaction with the curriculum for cutting-edge development
23. Satisfaction with the setting of academic exchanges at home and abroad
24. Satisfaction with integrated training settings in practice design
25. Satisfaction with teacher-student interaction
26. Satisfaction with teachers' teaching methods and strategies
27. Satisfaction with postgraduates' classroom participation
28. Whether the academic integrity complied or not
29. Whether the GPA (Grade Point Average) score above 3.5 or not
30. Postgraduate degree dissertation scores
31. Postgraduates publications
32. Postgraduates patent applications
33. Postgraduates honors
34. Whether participated in the academic conference or not
35. Whether participated in other social competitions and awards or not
36. Whether employed and set up a business or not

https://doi.org/10.33225/jbse/22.21.408

Appendix 3

Take the product evaluation as an example and calculate the score of the common factor

Table 8 shows linear combination relationship between common factors F_{10}, F_{11} and F_{12} and each three-level index, the score of the common factor can be calculated based on the following expression of the factor score function:

$$Fm = \sum_{1}^{36} D_i \, g \, C_i$$

In the product evaluation, *m* is 10-11, *Di* is the score coefficient of *Ci*, and *Ci* is the measurable index, ps means product evaluation score.

$$F_{10ps} = -0.072C_{28} + 0.136C_{29} + 0.084C_{30} - 0.329C_{31} + 0.339C_{32} + 0.329C_{33} + 0.059C_{34} - 0.268C_{35} + 0.074C_{36}$$

$$F_{11ps} = 0.410C_{28} + 0.200C_{29} + 0.384C_{30} - 0.007C_{31} - 0.164C_{32} - 0.153C_{33} + 0.125C_{34} + 0.509C_{35} + 0.180C_{36}$$

$$F_{12ps} = -0.072C_{28} + 0.136C_{29} + 0.084C_{30} - 0.329C_{31} + 0.339C_{32} + 0.329C_{33} + 0.059C_{34} - 0.268C_{35} + 0.074C_{36}$$

The score equation of first-level index is established. Assuming a total score of 100, *Fs* represents the score of common factors. Weight adjustment of the second-level index of the product evaluation is (0.322, 0.424, 0.254). The total academic achievement is confirmed by

$$F_s = 100 \times 0.38(0.322F_{10ps} + 0.424F_{11ps} + 0.254F_{12ps})$$

Score coefficients of other indicators can also be obtained through investigation and Factor Analysis. Among these, F_{cs} means comprehensive score of the common factors F_{fs} represents foundation score of the common factors, F_{rs} stands for resources score of the common factors F_{ts} means teaching score of the common factors, F_{ps} represents product evaluation score of the common factors. The final evaluation score is as follows:

$$F_{cs} = 17 \times F_{fs} + 11 \times F_{rs} + 34 \times F_{ts} + 38 \times F_{ps}$$

Received: *March 22, 2022* Revised: *April 22, 2022* Accepted: *May 08, 2022*

Cite as: Duan, P.T., Niu, H. J., Xiang, J.W., & Han, C.Q. (2022). Multi-index and hierarchical comprehensive evaluation system for training quality of science and engineering postgraduates. *Journal of Baltic Science Education*, *21*(3), 408-427. https://doi.org/10.33225/jbse/22.21.408

Peitong Duan Lecturer, Deputy Director, National University Science Park, Jiangsu Normal University, No. 246 Jiefang Road, Yunlong District, Xuzhou 221009, Jiangsu, China.
E-mail: 6019990667@jsnu.edu.cn

Huijun Niu Postgraduate Student, School of Physics and Electronic Engineering, Jiangsu Normal University, No.101 Shanghai Road, Tongshan District, Xuzhou 221116, Jiangsu, China.
E-mail: niunhj@163.com

Jiawen Xiang Postgraduate Student, School of Physics and Electronic Engineering, Jiangsu Normal University, No.101 Shanghai Road, Tongshan District, Xuzhou 221116, Jiangsu, China.
E-mail: 1156741602@qq.com

Caiqin Han *(Corresponding author)* PhD, Professor, School of Physics and Electronic Engineering, Jiangsu Normal University, No.101 Shanghai Road, Tongshan District, Xuzhou 221116, Jiangsu, China.
E-mail: hancq@jsnu.edu.cn
ORCID: https://orcid.org/0000-0002-0570-1426