

A Quality Evaluation System for Dissertation Based on Fuzzy Analytic Hierarchy Process

Kang Xu^{1*}, Zhenhao Zhang^{2*}, Shuo Wang^{2*}, Yunhui Lv², Jiafei Li², Xiaodan Cai³, Yongchun Cui^{2#}

¹School of Basic Medical Sciences, Capital Medical University, Beijing, China

²Beijing Key Laboratory of Pre-Clinical Research and Evaluation for Cardiovascular Implant Materials, State Key Laboratory of Cardiovascular Disease, National Center for Cardiovascular Diseases, Animal Experimental Center, Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

³Department of Endocrine, Suqian First Hospital, Suqian, China

Email: #cuiyongchun@fuwai.com

How to cite this paper: Xu, K., Zhang, Z.H., Wang, S., Lv, Y.H., Li, J.F., Cai, X.D. and Cui, Y.C. (2024) A Quality Evaluation System for Dissertation Based on Fuzzy Analytic Hierarchy Process. *Applied Mathematics*, 15, 441-454.

<https://doi.org/10.4236/am.2024.157026>

Received: May 29, 2024

Accepted: July 8, 2024

Published: July 11, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A dissertation is a research report or scientific paper written by an author to obtain a certain degree. It reflects postgraduates' research achievements and the educational quality of an institute, even a country. To construct an optimized quality evaluation system for postgraduate dissertation (QESPD), we summarized the influencing factors and invited 10 experienced specialists to rate and prioritize them based on fuzzy analytic hierarchy process. Four primary indicators (innovation, integrity, scientificity and normativity) and 16 sub-indicators were selected to form the evaluation system. The order of primary indicators by weight, was innovation (0.4269), scientificity (0.2807), integrity (0.1728) and normativity (0.1196). The top five sub-dimensions were theoretical originality, scientific value, data reliability, design rationality and evidence credibility. To demonstrate the effectiveness of the proposed system, a case study was performed. In the case study, it was demonstrated that the established two-index-hierarchy QESPD in this study was a more scientific and reasonable evaluation system worthy of promotion and application.

Keywords

Dissertation, Quality Evaluation, Indicator System, Fuzzy Analytic Hierarchy Process, Case Study

1. Introduction

The quality of postgraduate dissertation not only reflects the research ability and

*These authors contribute equally.

#Corresponding author.

academic level of the degree applicants, but also reflects the quality of postgraduate education [1]-[3]. The quality evaluation system for postgraduate dissertation (QESPD) is a baton and a navigation for postgraduate education. In recent years, with the expansion of the enrollment scale of master's students, some scholars have proposed that the quality of graduate theses should be further improved by controlling the personal qualities of graduate student, optimizing the structure of the teaching staff, improving the writing process and building a quality management system [4]. As previous studies reported [5]-[8], researchers had made certain improvements in the education models and methods. However, few researches focused on the quality evaluation of postgraduate dissertation. According to the results of the sample review of graduate papers, the quality of some graduate thesis has problems. For example, non-standard language descriptions, uncomprehensive literature reviews, lack of novelty in the topic selection, and the author's academic attitude is not rigorous enough [9]. How to make a scientific, objective and efficient QESPD is a subject with realistic significance [10]-[13].

In this study, we comprehensively analyzed potential dissertation-quality-influencing factors, fully considered the opinions of experts, and successfully established a novel and applicable two-index-hierarchy QESPD based on a fuzzy analytic hierarchy process (FAHP).

2. Methods

2.1. Organization of a Decision-Making Group (DMG)

To outline an indicator system for the evaluation of postgraduate's dissertation quality, a DMG was organized according to inclusive criteria as previously described [14] [15].

- 1) Doctoral or master's tutor in basic medical research, clinical medical research or graduate education management.
- 2) Working as faculty in higher medical schools or research institutes, associated with graduate education.
- 3) Reviewing more than 20 post graduate dissertation.
- 4) Agree to answer the expert questionnaire.

2.2. Construction of an Indicator System

Based on the comments of DMG and reported indicators in recent literature that could reflect the postgraduate dissertation quality [12] [14] [16] [17], a two-index-hierarchy indicator system was constructed.

2.3. Determination of Weights and Ranks of Indicators

- Build a pair-wise comparison (PWC) matrix

Regarding the postgraduate dissertation validity evaluation framework, the DMG should complete pair-wise comparisons between primary and secondary indicators in the two-index-hierarchy indicator system. So, first they were re-

quired to compare the importance of each indicator with the adjacent indicators of the same level (see **Appendix A**). Then, a matrix \tilde{X} was achieved according to PWC [15] [18]-[20].

$$\tilde{x}_{ij} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{n1} & \tilde{x}_{n2} & \cdots & \tilde{x}_{nn} \end{bmatrix} \quad (1)$$

- Consistency check of the matrices

Based on the results of pair-wise comparisons, matrices were established, and the consistency checks of the matrices was performed by computing the consistency ratio (CR):

$$\begin{aligned} CR &= CI/RI \\ CI &= (\lambda_{\max} - n)/(n-1) \end{aligned} \quad (2)$$

where: λ_{\max} means the largest eigen value of PWC. “ CI ” means the consistency index, “ RI ” means the random index, and “ n ” indicates the number of criteria that would be judged against (*i.e.*, matrix size) [15].

- Weights determination

To calculate the weights of indicators, 1/9-9 scaling method (**Table 1**) was used as the scoring principle.

Table 1. 1/9-9 scaling method.

| Score scale | Meaning (A vs. B) |
|--------------------------------|------------------------------------------------|
| 9 | Absolutely more important (AMI) |
| 7 | Very strongly more important (VSMI) |
| 5 | Strongly more important (SMI) |
| 3 | Weakly ore important (WMI) |
| 1 | Equally important (EI) |
| 1/3 | Weakly less important (WLI) |
| 1/5 | Strongly less important (SLI) |
| 1/7 | Very strongly less important (VS LI) |
| 1/9 | Absolutely less important (ALI) |
| 2, 4, 6, 8, 1/2, 1/4, 1/6, 1/8 | The median of the two adjacent judgments above |

The above table was used to convert the qualitative factors in the two-index-hierarchy indicator system to fuzzy numbers. All of the primary indicators and the secondary indicators were paired and compared with each other respectively (see **Appendix B** for detailed data).

Based on the fuzzy numbers, a PWC matrix between criteria was established to calculate the indicators’ weights using FAHP method [21]-[25]. The weights were identified as follows [26].

$$\begin{aligned}
\alpha_j &= \left[\prod_{i=1}^n l_{ij} \right]^{1/n}, \beta_j = \left[\prod_{i=1}^n m_{ij} \right]^{1/n} \\
\gamma_j &= \left[\prod_{i=1}^n n_{ij} \right]^{1/n}, \delta_j = \left[\prod_{i=1}^n s_{ij} \right]^{1/n} \\
\alpha &= \sum_{j=1}^n \alpha_j, \beta = \sum_{j=1}^n \beta_j, \gamma = \sum_{j=1}^n \gamma_j, \delta = \sum_{j=1}^n \delta_j
\end{aligned} \tag{3}$$

Then, using a hierarchy model identified by the FAHP approach, the extracted indicators of postgraduate dissertation quality were ranked [27].

3. Results

3.1. Demographic Characteristics of DMG

Ten experts qualified as doctoral tutor in basic medical research ($n = 6$) and clinical medicine ($n = 4$) were invited to participate in this study. Half male and half female, most belonged to the 35 - 55 age group. 4 experts had reviewed more than 50 postgraduate research theses (40%), 5 experts had reviewed 30 - 50 postgraduate research theses (50%) and 1 had reviewed more than 20 postgraduate research theses (10%) (Table 2).

3.2. Extracting the Influencing Indicators of Postgraduate Dissertation Quality

Based on a comprehensive consideration on associated literature [28] [29] and DMG's comments, we extracted those indicators that can be used to effectively evaluate the postgraduate dissertation quality, and established a two-index-hierarchy indicator system, including primary and secondary dimensions.

As shown in Figure 1, the primary dimensions include innovation, integrity, scientificity and normativity. With respect to the DMG opinion, we further developed secondary indicators for each primary dimension. "Innovation" included scientific progress comprehensiveness, topic novelty, theoretical originality, and scientific value. "Integrity" included prominent argument, evidence credibility, rigorous argumentation and solid conclusion. "Scientificity" indicators included design rationality, basic knowledge accuracy, method feasibility and data reliability. "Normativity" included accurate, coherent and fluent writing, standard typographic format, professional language expression and normative/accurate citations.

Based on literature review and DMG's opinions, the two-level evaluation system was established. The primary indicator includes 4 items, and the secondary indicator includes 16 items. DMG means decision-making group.

3.3. Determining the Weights and Important Coefficients of Postgraduate Dissertation Quality Indicators

According to the calculated weight of each indicator, primary indicators were innovation (0.4269), scientificity (0.2807), integrity (0.1728) and normativity (0.1196) in descending order of weight. The five highest weighted secondary indicators included theoretical originality, scientific value, data reliability, design

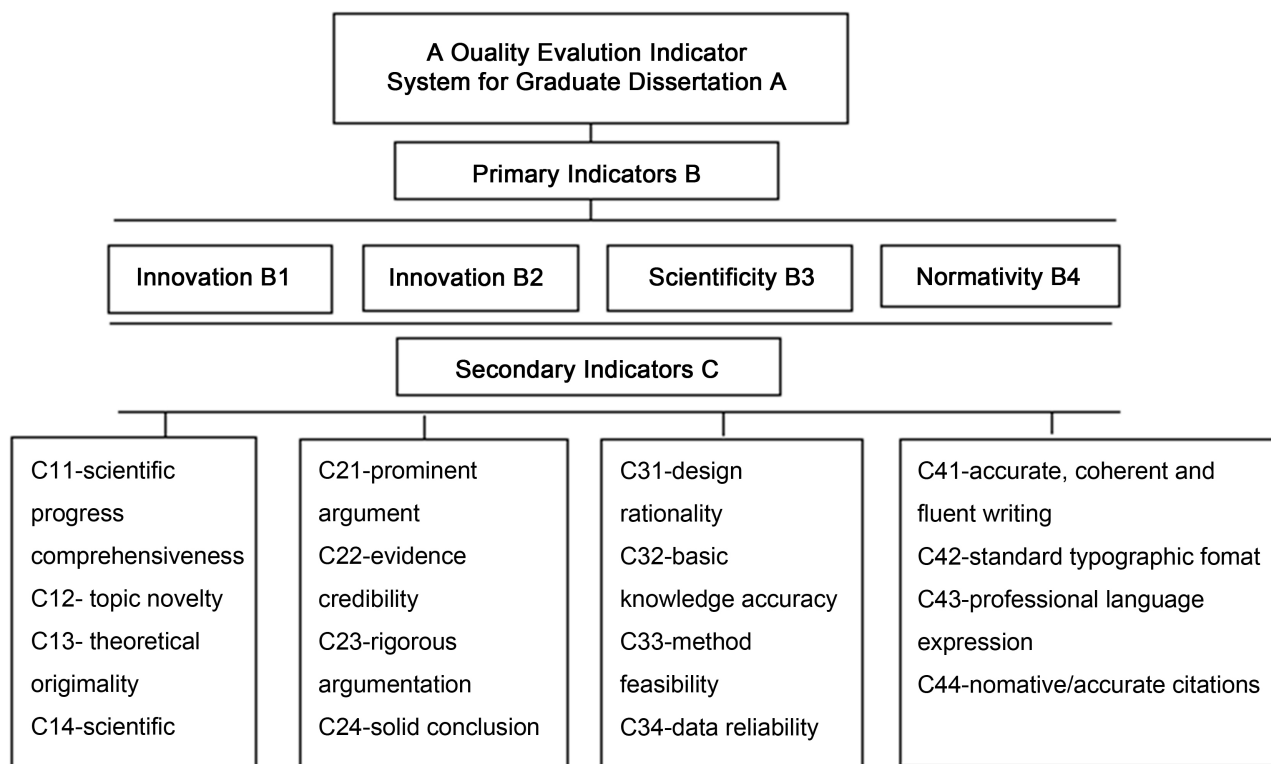


Figure 1. The two-level evaluation indicator system.

Table 2. Descriptive demographic characteristics of specialists.

| | Item | Number of responders | Percentage (%) |
|---------------------------------|---------------------|----------------------|----------------|
| Gender | Male | 5 | 50 |
| | Female | 5 | 50 |
| Age | 25 - 35 | 1 | 10 |
| | 35 - 55 | 8 | 80 |
| | More than 50 | 1 | 10 |
| Number of reviewed dissertation | 20 - 30 | 1 | 10 |
| | 30 - 50 | 5 | 50 |
| | More than 50 | 4 | 40 |
| Professional title | Professor | 6 | 60 |
| | Associate professor | 4 | 40 |

rationality, evidence credibility (**Table 3**).

To verify the consistency and validity of expert scoring results (see **Appendix B**), we used consistency ratio (CR) to judge variations between PWCs. CR value less than or equal to 0.1 indicates that the opinions of experts are in good consistency; CR value greater than 0.1 indicated poor consistency [27]. In this study, CR value was equal to 0.01, suggesting that the confidence level of experts response was over 90%. The scoring results provided by experts were consistent and reliable.

3.4. Postgraduate Dissertation Quality Grading

In order to facilitate the extended application of QESPD, a quality grading method was used in this study. First, experts calculated the formula calculation results (FCR) through the formula in **Table 4**. Second, experts obtained the actual effectiveness value (AEV) of each indicator by multiplying their FCR value by corresponding weight in **Table 3**. Finally, a total of four postgraduate dissertation quality levels were suggested, and represented as “A (excellent), B (good), C (medium) and D (poor)” (see **Table 5** for details). If required, this quality evaluation level representation method was not only suitable for primary indicators, but also independently ranked secondary indicators.

Table 3. Weight and prioritization of indicators and sub-indicators using FAHP.

| Primary indicators | Weight of Primary indicators | Sub-indicators | Weight of sub-indicators | Priority |
|--------------------|------------------------------|-------------------------------------------|--------------------------|----------|
| Innovation | 0.4269 | scientific progress comprehensiveness C11 | 0.0122 | 14 |
| | | topic novelty C12 | 0.0731 | 6 |
| | | theoretical originality C13 | 0.2044 | 1 |
| | | scientific value C14 | 0.1372 | 2 |
| Integrity | 0.1728 | prominent argument C21 | 0.0235 | 12 |
| | | evidence credibility C22 | 0.0877 | 5 |
| | | rigorous argumentation logic C23 | 0.0097 | 15 |
| | | solid conclusion C24 | 0.0519 | 8 |
| Scientificity | 0.2807 | design rationality C31 | 0.0916 | 4 |
| | | basic knowledge accuracy C32 | 0.0196 | 13 |
| | | method feasibility C33 | 0.0594 | 7 |
| | | data reliability C34 | 0.1101 | 3 |
| Normativity | 0.1196 | accurate, coherent and fluent writing C41 | 0.0083 | 16 |
| | | standard typographic format C42 | 0.046 | 9 |
| | | professional expression C43 | 0.0387 | 10 |
| | | normative and accurate citations C44 | 0.0266 | 11 |
| Total | 1.0 | | 1.0 | |

3.5. Case Study

To verify the applicability of the proposed QESPD, a case study was performed. First, 50 doctoral research theses (reviewed by the authors during 2010-2019) in PUMC were selected, including 25 doctoral theses in basic medicine and 25 doctoral theses in clinical medicine. Then, 3 experts were invited to re-evaluate the quality of them using the quality evaluation system of postgraduate dissertation constructed in this study.

Table 4. Calculation of different index items for dissertation quality grading.

| Primary indicators | Secondary indicators | Formula | FCR value interpretation |
|--------------------|-------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| Innovation | scientific progress comprehensiveness C11 | <p>FCR = A/4</p> <p>A value may equal to 1, 2, 3 or 4</p> <p>1: The advance of a chosen research area at home and abroad is not fully described.</p> <p>2: The advance of a chosen research area at home and abroad is fully described.</p> <p>3: The advance of a chosen research area at home and abroad is fully described and commented appropriately.</p> <p>4: The advance of a chosen research area at home and abroad is fully described, commented appropriately and put forward one's own views.</p> | 0 < FCR ≤ 1, the closer to 1, the better |
| | topic novelty C12 | <p>FCR = A/3</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Represent the direction of the development of theory or technology in a scientific field.</p> <p>2: Contribute to the upgrading of industrial technology and achieve leapfrog development.</p> <p>3: Lead the formation and development of the country's future emerging industries.</p> | 0 < FCR ≤ 1, the closer to 1, the better |
| | theoretical originality C13 | <p>FCR = A/3</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Follow-up innovation: Based on the research of others, make some necessary extensions or changes to develop something new.</p> <p>2: Integrated innovation: refer to the combination of existing technologies to create a new product or new technology, or to introduce mature technologies in one field into another field, so that it can create new changes.</p> <p>3: Primitive innovation: An invention or discovery with independent intellectual property rights.</p> | 0 < FCR ≤ 1, the closer to 1, the better |
| | scientific value C14 | <p>FCR = A/3</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Based on the research of others, through the necessary extensions and changes, the solution of existing problems was optimized.</p> <p>2: By introducing or reorganizing technology, inventing new things or new methods, and expanding or changing methods to solve incomplete problems.</p> <p>3: Primitive innovation can solve the major problems that urgently need to be solved in current social practice.</p> | 0 < FCR ≤ 1, the closer to 1, the better |
| Integrity | clear point of view C21 | <p>FCR value may equal to 0 or 1</p> <p>0: Argument is fuzzy.</p> <p>1: Argument is clear.</p> | |

Continued

| | | | |
|---------------|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|
| | evidence credibility, full workload C22 | <p>FCR = $A/3$</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Less work to support the argument.</p> <p>2: Full workload, but not enough to support the argument.</p> <p>3: Full workload to support the argument.</p> | $0 < \text{FCR} \leq 1$, the closer to 1, the better |
| | rigorous argumentation C23 | <p>FCR value may equal to 0 or 1</p> <p>0: Poor logic, failed to prove the point.</p> <p>1: Clear thinking, proper method, enough arguments, strong logic, can prove the point.</p> | |
| | solid conclusion C24 | <p>FCR value may equal to 0 or 1</p> <p>0: Failed to give a solid conclusion.</p> <p>1: Give a solid conclusion.</p> | |
| | design rationality C31 | <p>FCR value may equal to 0 or 1</p> <p>0: Study design is unreasonable.</p> <p>1: Study design is scientific and reasonable, which can ensure scientific research is carried out in an orderly manner.</p> | |
| Scientificity | basic knowledge accuracy C32 | <p>FCR value may equal to 0 or 1</p> <p>0: Cannot perform accurate data analysis and problem summary.</p> <p>1: Accurate basic theory and profound theoretical skills.</p> | |
| | method feasibility C33 | <p>FCR = A/B</p> <p>A = The number of research methods that are reasonable, accurate, and feasible.</p> <p>B = Total number of methods.</p> | $0 < \text{FCR} \leq 1$, the closer to 1, the better |
| | data reliability C34 | <p>FCR value may equal to 0 or 1</p> <p>0: Untrustworthy result.</p> <p>1: The research results are accurate, objective, traceable, and repeatable.</p> | |
| Normativity | accurate, coherent and fluent writing C41 | <p>FCR = $A/3$</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Three or more typographical errors in chapters and paragraphs.</p> <p>2: One-two typographical errors in chapters and paragraphs.</p> <p>3: Rigorous structure with clear chapters and paragraphs.</p> | $0 < \text{FCR} \leq 1$, the closer to 1, the better |
| | standard typographic format C42 | <p>FCR = $A/3$</p> <p>A value may equal to 1, 2 or 3</p> <p>1: Three or more points are contrary to relevant rules and standards.</p> <p>2: One-two points are contrary to relevant rules and standards.</p> <p>3: Each part or link complies with relevant rules and standards.</p> | $0 < \text{FCR} \leq 1$, the closer to 1, the better |

Continued

| | | |
|--------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| professional language expression C43 | <p>FCR = A/3 A value may equal to 1, 2 or 3 1: Three or more places did not follow the internationally accepted academic norms of the subject. 2: One-two places did not follow the internationally accepted academic norms of the subject. 3: Strictly follow the internationally accepted academic norms of this discipline.</p> | 0 < FCR ≤ 1, the closer to 1, the better |
| normative and accurate citations C44 | <p>FCR = A/3 A value may equal to 1, 2 or 3 1: References are accurate and meet format requirements. 2: References are accurate and meet format requirements. Every important point adequately referenced. 3: References are accurate and meet format requirements. Every important point adequately referenced. References are reasonably up to date.</p> | 0 < FCR ≤ 1, the closer to 1, the better |

Table 5. The presentation of quality rating results.

| Rating | Actual effective value (AEV) | Markers |
|-----------|------------------------------|---------|
| Excellent | AEV ≥ 0.9 | A |
| Good | 0.7 ≤ AEV < 0.9 | B |
| Medium | 0.5 ≤ AEV < 0.7 | C |
| Poor | AEV < 0.5 | D |

According to the experts' comments, the FCR of each index was calculated; combined with primary indicators and sub-indicators' weight, the comprehensive implementation effectiveness assessment set was obtained. After normalization processing and further assignment, final comprehensive postgraduate dissertation quality AEV for the 50 doctoral theses showed that a total of 18 (36%, 10 academic and 8 professional) doctoral theses, with a AEV score of excellent (AEV ≥ 0.9), rated A; a total of 26 (52%) doctoral theses, with a AEV score of good (0.7 ≤ AEV < 0.9), rated B; a total of 5 (10%) doctoral theses, with a AEV score of medium (0.5 ≤ AEV < 0.7), rated C; 1 (2%) doctoral thesis, with a AEV score of poor (AEV < 0.5), rated D.

The evaluation result of our proposed QESPD was similar with the review result of the current version of dissertation assessment book (DAB) in PUMC (**Appendix C**). Compared with the current version of DAB, our proposed QESPD focuses more on the problems existing in postgraduate dissertation. We found the common problems in reviewed thesis were that the width and depth of the literature was not proper, and the extraction of scientific questions was not precise enough.

4. Discussion and Conclusion

The degree of social development and civilization is increasingly associated with the quality of education both in the developing countries and in the developed countries. With the globalization of scientific technology, new challenges have appeared in the quality control and management of higher education [30]-[33]. Higher education institutes have a responsibility to ensure that the quality of higher education meets internationally acceptable standards [34].

A good scientific dissertation needs to have three basic characteristics [35]-[37]: 1) objective prescriptiveness, which means that it should meet the requirements of degree authorizer; 2) subjectivity, which means that it should meet the various requirements of country, society, schools and families for students' mental training and personality development; 3) knowledge transformation, which means that it should meet the requirements of social and economic benefits.

In this study, we used a FAHP method to construct a novel QESPD. FAHP adopted in this study is a multi-criteria decision-making method [38], which has been widely used to evaluate the efficacy of various medical diagnostic techniques and treatment protocols. However, no studies have been done with the postgraduate dissertation quality evaluation. FAHP is an extension of traditional analytic hierarchy process (AHP) and can overcome the deficiency of AHP. It fully reveals human's fuzzy opinions [39] and avoids considering only distinct judgments of decision makers [40] [41]. Here, we demonstrated that FAHP is also suitable for constructing a scientific and effective postgraduate dissertation quality evaluation indicator system.

The QESPD established here consists of two-index hierarchies. In the primary dimension, innovation is the most important with the highest weight values. Innovation includes four secondary indicators: scientific progress comprehensiveness, topic novelty, theoretical originality and scientific value. Innovations in research results fall into three categories: 1) If the study was based on the research of others, only made some necessary extensions or changes to develop something new, the results were called "follow-up innovation"; 2) If the study made a combination of existing technologies to create a new product or new technology, or to introduce mature technologies in one field into another field, so that it can create new changes, the results were called "integrated innovation"; 3) If the study was an invention or discovery with independent intellectual property rights, the results were called "primitive innovation".

In addition, in previous studies, the scientificity of academic papers mainly focused on research methods and data collection. The research methods used in the postgraduate thesis should be scientific, reasonable, and reliable, and the data collection should be rigorous, objective and accurate. In this study, scientificity included "design rationality", "basic knowledge accuracy", "method feasibility", "data reliability". Our results showed that sub-indicator "data reliability" had the highest weighting, suggesting that reliable results were essential for the

scientificity of graduate thesis.

The comprehensiveness of scientific progress reflects a better ability to check literature data, grasp the underlying trends and the focus in the chosen field, analyze the existing problems and explore corresponding solutions for focus or hot issues [42].

The top five secondary indicators in the QESPD include “theoretical originality”, “scientific value”, “data reliability”, “design rationality”, “evidence credibility” in order. Among them, theoretical originality and scientific value belong to the primary dimension “innovation” as mentioned above; data reliability and design rationality belong to the primary dimension “scientificity”; evidence credibility belongs to primary dimension “integrity”.

The theoretical originality is the most important secondary indicator for the quality evaluation of postgraduate dissertation. The “follow-up” innovative results contribute to optimizing solutions to existing problems through extending other people’s research results. “Integrated” innovative results provide new things or new methods to solve unsolved problems by introducing or reorganizing existing technologies. “Primitive” innovative results can solve the major problems that urgently need to be solved in current field.

As previous study described, only when there are sufficient evidences and strong arguments, a solid conclusion can be reached [36]. If the evidence was insufficient or the logic was confusing, the article will not be persuasive, and no solid conclusion can be reached. So, evidence credibility are necessary. Study design should match article objective. References should be accurate and meet format requirements. Every important point is adequately referenced and reasonably up to date.

For the convenience of popularization and application of this system, this study also proposed the use “A-D” to represent postgraduate dissertation quality evaluation results. As a case study, this study used the system to assess the postgraduate dissertation quality of 50 doctoral theses in PUMC. The results showed that compared with the current version of DAB, our proposed QESPD focuses more on the problems existing in postgraduate dissertation. The evaluation index system constructed in this study can further complement and optimize the existing quality assessment book for postgraduate dissertation and provide necessary supplements for it.

Graduate education is an important part of higher education. Our study provided standards and norms for postgraduate thesis evaluation by constructing a two-index hierarchies quality assessment indicator system, which will help to standardize academic research and ensure the quality and reputation of higher education. A standardized evaluation scheme ensures that each paper is evaluated based on the same criteria and processes, which helps to maintain the fairness and impartiality of graduate thesis evaluation. In short, the graduate thesis quality evaluation indicator system plays an important role in standardizing academic research, promoting fairness and justice, guiding research performance, improving education quality, and laying a solid foundation for career

development.

However, there are certain limitations in this study, for example, all of the decision makers (DMs) are from the same region—Beijing. Expert opinions may be affected by many factors, such as cultural background, socioeconomics level. In the future, we will expand the number of DMs and invite experts from different regions to obtain more comprehensive information.

Acknowledgements

The authors wish to thank the advice and support of Dr. Jiajie Li with regards data analysis. This work was supported by [National Natural Science Foundation of China] under Grant [number 81970387; 32300794]; [Beijing Municipal Science & Technology Commission] under Grant [numbers Z161100005016014, Z101107052210004]; [Peking Union Medical College Small-scale Characteristic School Project] under Grant [number 10023201700202, 2017E-JG02]; and [Beijing Key Laboratory of Preclinical Research and Evaluation of Cardiovascular Implant Materials] under Grant [number 2018-PT2-ZR04]; R&D Program of Beijing Municipal Education Commission (KM202310025030). The funding body did not involve in the design of the study or collection, analysis, and interpretation of data.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Blackmur, D. (2005) An Evaluation of the South African Model of MBA Accreditation. *Quality in Higher Education*, **11**, 87-102. <https://doi.org/10.1080/13538320500175001>
- [2] Bertolin, J. and Leite, D. (2008) Quality Evaluation of the Brazilian Higher Education System: Relevance, Diversity, Equity and Effectiveness. *Quality in Higher Education*, **14**, 121-133. <https://doi.org/10.1080/13538320802279980>
- [3] Bloxham, S. (2012) ‘You Can See the Quality in Front of Your Eyes’: Grounding Academic Standards between Rationality and Interpretation. *Quality in Higher Education*, **18**, 185-204. <https://doi.org/10.1080/13538322.2012.711071>
- [4] Min, Z. and Lizhen, Z. (2017) Reflections on the Quality Assurance System of Full-Time Master’s Degree Dissertation. *China Adult Education*, No. 21, 86-88.
- [5] Alexander, F.K. (2000) The Changing Face of Accountability. *The Journal of Higher Education*, **71**, 411-431. <https://doi.org/10.1080/00221546.2000.11778843>
- [6] Altbach, P.G. (2009) The Giants Awake: The Present and Future of Higher Education Systems in China and India. In: Altbach, P.G., Ed., *Higher Education to 2030, Volume 2, Globalisation*, OECD Publishing, 179-204. <https://doi.org/10.1787/9789264075375-8-en>
- [7] Bazargan, A. (2007) Problems of Organising and Reporting Internal and External Evaluation in Developing Countries: The Case of Iran. *Quality in Higher Education*, **13**, 207-214. <https://doi.org/10.1080/13538320701800126>

- [8] UNESCO-UIS (2012) International Standard Classification of Education ISCED 2011.
- [9] Min, L. and Hongjie, C. (2017) Typical Characteristics of Unqualified Academic Master's Degree dissertations—An Analysis Based on the Expert Review of Dissertation Sampling. *Degree and Postgraduate Education*, **6**, 61-67.
- [10] UNESCO (2008) Global Initiative for Quality Assurance Capacity (GIAC). UNESCO, Paris.
- [11] Altbach, P.G., Reisberg, L. and Rumbley, L.E. (2010) Trends in Global Higher Education. BRILL. <https://doi.org/10.1163/9789004406155>
- [12] Fabra, M.E. and Camisón, C. (2009) Direct and Indirect Effects of Education on Job Satisfaction: A Structural Equation Model for the Spanish Case. *Economics of Education Review*, **28**, 600-610. <https://doi.org/10.1016/j.econedurev.2008.12.002>
- [13] OECD (2019) Education at a Glance 2019. OECD Publisher. <https://doi.org/10.1787/f8d7880d-en>
- [14] Tsui, C. (2009) Quality Assurance in Higher Education: A Hong Kong Perspective. Ph.D. Dissertation, University College London.
- [15] Cui, Y., Luo, F., Yang, B., Li, B., Zhang, Q., Das, G., et al. (2019) Construction and Application of Service Quality Evaluation System in the Preclinical Research on Cardiovascular Implant Devices. *BMC Medical Informatics and Decision Making*, **19**, Article No. 37. <https://doi.org/10.1186/s12911-019-0773-4>
- [16] OECD (2015) ISCED 2011 Operational Manual: Guidelines for Classifying National Education Programmes and Related Qualifications. OECD Publishing.
- [17] Beland, L. and Murphy, R. (2016) Ill Communication: Technology, Distraction & Student Performance. *Labour Economics*, **41**, 61-76. <https://doi.org/10.1016/j.labeco.2016.04.004>
- [18] Hung, Y.H., Huang, M.L. and Chen, K.S. (2003) Service Quality Evaluation by Service Quality Performance Matrix. *Total Quality Management & Business Excellence*, **14**, 79-89. <https://doi.org/10.1080/14783360309706>
- [19] Dagger, T.S., Sweeney, J.C. and Johnson, L.W. (2007) A Hierarchical Model of Health Service Quality. *Journal of Service Research*, **10**, 123-142. <https://doi.org/10.1177/1094670507309594>
- [20] Chen, L., Chen, S. and Su, C. (2017) An Innovative Service Quality Evaluation and Improvement Model. *The Service Industries Journal*, **38**, 228-249. <https://doi.org/10.1080/02642069.2017.1389907>
- [21] Yu, K. and Shen, C. (2011) Service Quality Management by Expanded Traditional Importance-Performance Analysis. *Journal of the Chinese Institute of Industrial Engineers*, **28**, 146-154. <https://doi.org/10.1080/10170669.2010.546557>
- [22] Azzopardi, E. and Nash, R. (2013) A Critical Evaluation of Importance-Performance Analysis. *Tourism Management*, **35**, 222-233. <https://doi.org/10.1016/j.tourman.2012.07.007>
- [23] Feng, M., Mangan, J., Wong, C., Xu, M. and Lalwani, C. (2014) Investigating the Different Approaches to Importance-Performance Analysis. *The Service Industries Journal*, **34**, 1021-1041. <https://doi.org/10.1080/02642069.2014.915949>
- [24] Lai, I.K.W. and Hitchcock, M. (2015) A Consideration of Normality in Importance-Performance Analysis. *Current Issues in Tourism*, **18**, 979-1000. <https://doi.org/10.1080/13683500.2014.990423>
- [25] Sever, I. (2015) Importance-Performance Analysis: A Valid Management Tool? *Tourism Management*, **48**, 43-53. <https://doi.org/10.1016/j.tourman.2014.10.022>
- [26] Cronin, J.J. and Taylor, S.A. (1994) Servperf versus Servqual: Reconciling Perfor-

- mance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. *Journal of Marketing*, **58**, 125-131. <https://doi.org/10.1177/002224299405800110>
- [27] Saaty, T.L. (1980) *The Analytic Hierarchy Process: Planning, Priority Setting, Resources Allocation*. McGraw.
- [28] Tong, S., Li, X. and Liu, R. (2010) Research on Ways to Enhancing Undergraduate Dissertation Quality. *Shanxi Finance and Economics University*, **9**, 45-53.
- [29] Ivan, L. (2013) Criteria for Evaluating Masters and Doctoral Theses. <https://scientificlanguage.com/methodology/thesiscriteria7>
- [30] Alstete, J.W. (1995) Benchmarking in Higher Education: Adapting Best Practices to Improve Quality. ERIC Digest, 1-6.
- [31] Westerheijden, D.F., Hulpiau, V. and Waeytens, K. (2006) Lines of Change in the Discourse on Quality Assurance: An Overview of Some Studies into What Impacts Improvement. *EAIR 28th Annual Forum 2006*, Rome, 30 August-1 September 2006, 1-19.
- [32] El-Khawas, E. (2007) Accountability and Quality Assurance: New Issues for Academic Inquiry. In: Forest, J.J.F. and Altbach, P.G., Eds., *International Handbook of Higher Education*, Springer, 23-37. https://doi.org/10.1007/978-1-4020-4012-2_3
- [33] Prince, M. and Stella, A. (2007) External Quality Assurance in Higher Education: Making Choices. International Institute for Educational Planning.
- [34] Martin, M. (2007) Cross-Border Higher Education: Regulation, Quality Assurance and Impact. International Institute for Educational Planning.
- [35] Bamber, V. and Anderson, S. (2012) Evaluating Learning and Teaching: Institutional Needs and Individual Practices. *International Journal for Academic Development*, **17**, 5-18. <https://doi.org/10.1080/1360144x.2011.586459>
- [36] Jin, Z., Li, S. and Xu, J. (2014) Quality Evaluation of Undergraduate Thesis as the Goal of Applied Talent's Cultivation—Based on the Self-Assessment Analysis of Accounting Major Theses. *Proceedings of the 2014 International Conference on Education Reform and Modern Management*, Phuket, 11 January 2014, 151-154. <https://doi.org/10.2991/ermm-14.2014.42>
- [37] Gao, Y., Chen, H.J., Shen, W.Q. and Li, M. (2017) Disciplinary Differences in the Quality of Academic Master's Dissertation-Quantitative Analysis Based on the Results of X Province Dissertation Sampling Inspection. *Academic Degrees & Graduate Education*, No. 2, 54-61.
- [38] Shafii, M., Rafiei, S., Abooe, F., Bahrami, M.A., Nouhi, M., Lotfi, F., et al. (2016) Assessment of Service Quality in Teaching Hospitals of Yazd University of Medical Sciences: Using Multi-Criteria Decision Making Techniques. *Osong Public Health and Research Perspectives*, **7**, 239-247. <https://doi.org/10.1016/j.phrp.2016.05.001>
- [39] Chang, C., Liang, C., Shu, K. and Tsai, C. (2015) Key Successful Factors of Knowledge Management for University Students Using E-Portfolios: Approach of Fuzzy Delphi and Fuzzy AHP. *Computer Applications in Engineering Education*, **23**, 673-681. <https://doi.org/10.1002/cae.21640>
- [40] Ma, Z., Shao, C., Ma, S. and Ye, Z. (2011) Constructing Road Safety Performance Indicators Using Fuzzy Delphi Method and Grey Delphi Method. *Expert Systems with Applications*, **38**, 1509-1514. <https://doi.org/10.1016/j.eswa.2010.07.062>
- [41] Gustafsson, F. and Rogers, J.G. (2017) Left Ventricular Assist Device Therapy in Advanced Heart Failure: Patient Selection and Outcomes. *European Journal of Heart Failure*, **19**, 595-602. <https://doi.org/10.1002/ejhf.779>
- [42] OECD (2018) *OECD Handbook for Internationally Comparative Education Statistics 2018: Concepts, Standards, Definitions and Classifications*. OECD Publishing.