

An Empirical Analysis of the Final Examination Papers in Educational Psychology Course for Undergraduates in Different Educational Majors under the Background of Teaching Reform

Tian Sun^{1*}, Yiyan Zhang²

¹Yunnan Technology and Business University, Kunming 650000, Yunnan, China

²Civil Aviation Management Institute of China, Beijing 100102, China

*Corresponding author: Tian Sun, 376208995@qq.com

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: In the context of the continuous advancement of teaching reform, the optimization of the curriculum evaluation system has become a key issue in the field of education. This study focuses on the course of educational psychology, using the final examination papers and scores of the 2022 Physical Education Major and 2024 Early Childhood Education Major undergraduate students at Yunnan Technology and Business University as samples. Basic statistical analysis methods are employed to conduct an in-depth investigation. The results show that the overall difficulty of the exam papers is moderate (0.7), with acceptable reliability (0.78) and good validity (0.7–0.8). However, the failure rate among students reaches 21.3%, reflecting learning difficulties. There is no significant difference in scores between students from different majors, attributed to the same teaching conditions and equal emphasis on the course by both majors. This study provides data support for the teaching of educational psychology courses, assisting teachers in improving their teaching based on the results, meeting the demands of teaching reform for precise teaching evaluation and quality enhancement, and promoting the development of educational psychology course teaching.

Keywords: Educational psychology course; Teaching reform; Examination papers analysis

Online publication: March 28, 2025

1. Introduction

Educational psychology course (EPC), as a core teacher education course for undergraduate students in educational majors, plays a foundational role in cultivating students' teaching and educational capabilities^[1]. In the current wave of teaching reform, the transformation of the educational evaluation system has become one of the core drivers for improving teaching quality^[2]. The traditional single-exam evaluation model can no longer meet the comprehensive assessment needs of modern education for students' overall quality and capability

development. Strengthening examination management, enhancing process assessment, and reasonably adjusting its proportion in the total course score have become important directions for teaching reform. However, in the current implementation process, accurately and objectively evaluating students' daily learning status still faces many challenges. It is urgent to build a scientific, comprehensive, and reform-oriented evaluation system that meets the needs of teaching reform. Given the high consistency in the teaching of EPC at Yunnan Technology and Business University (YTBV) for the 2022 Physical Education Major (PEM) and 2024 Early Childhood Education Major (ECEM), including the same teaching plan, content, class hours, textbooks, instructors, and exam papers, this study takes this opportunity to conduct an in-depth analysis of their final exam papers and scores. It aims to explore the commonalities and differences in the learning of EPC among students from different educational majors, providing strong data support and practical references for precisely optimizing the teaching of EPC under the background of teaching reform.

2. Materials and methods

2.1. General materials

This study selected the final examination papers in EPC from 94 undergraduate students of the 2022 PEM and 2024 ECEM at YTBV as the research subjects. During the grading process, a flow grading method was employed to ensure that the scoring criteria for each question were unified and objective. The total score for the exam paper was 100 points, with question types divided into two major categories: objective and subjective questions. The objective questions included single-choice questions (10 questions, totaling 15 points), multiple-choice questions (5 questions, totaling 15 points), and true-or-false questions (10 questions, totaling 10 points). The subjective questions consisted of term explanations (5 questions, totaling 20 points), short-answer questions (4 questions, totaling 20 points), and analysis questions (1 question, totaling 20 points). Both objective and subjective questions were designed to assess students' memory, understanding, and comprehensive application of knowledge, covering the core knowledge points and competency requirements of the EPC.

2.2. Test paper analysis methods

This study analyzed the final examination papers in EPC, scientifically evaluating aspects such as difficulty, reliability, and validity.

- (1) Difficulty: Refers to the level of difficulty of the test questions, which is a very important indicator for evaluating the examination papers^[3]. In this study, the classic difficulty classification standard was followed: 0.75–1 indicates an easy test paper, 0.6–0.74 is defined as a test paper of medium difficulty, and below 0.59 is judged as a difficult test paper. By analyzing the scoring situation of each question type, the overall difficulty coefficient of the test paper and the difficulty coefficient of different question types were calculated. This assessment judged the appropriateness of the test paper in terms of students' knowledge mastery level and provided a basis for subsequent teaching adjustments from the perspective of difficulty.
- (2) Reliability: Reliability was quantitatively measured using Cronbach's Alpha coefficient, which reflects the consistency of the scores for the test questions and is an important statistical representation of the reliability of the examination. Generally, a Cronbach's Alpha value greater than 0.9 indicates excellent consistency; 0.8–0.9 represents good consistency; 0.7–0.8 is within the acceptable range; and 0.6–0.7 means that the test paper requires significant revision^[4]. During the analysis, the Cronbach's Alpha value was further calculated after removing a specific question type (Question N). If the reliability of

the entire test paper significantly improved after removing the question, it indicated that the question might have issues affecting the overall quality of the test paper, such as ambiguous wording, and repetitive or off-topic knowledge points. This analysis helped identify areas for test paper optimization.

- (3) Validity: Validity was primarily determined by examining the correlation coefficients between the scores of each type of question and the total score of the test paper. Since the total score can comprehensively reflect students' overall mastery of the course, ideally, the scores of each question type should show a significant positive linear correlation with the total score. If a question type had a very low or negative correlation with the total score, it might indicate serious problems in the difficulty setting of the question, such as being too difficult and causing widespread student failure, or too easy to effectively distinguish student levels, or even having incorrect answers. This study categorized validity into three levels based on the validity coefficient: validity greater than 0.7 is considered good, 0.4–0.7 is fair, and less than 0.4 is considered poor. This classification comprehensively evaluated the effectiveness of the test paper.

2.3. Statistical analysis of scores for different majors

A comprehensive and in-depth analysis and comparison of the EPC exam scores between students of the PEM and ECEM were conducted, covering key indicators such as failure rate, pass rate, excellence rate, average score, subjective question scores, and objective question scores. The analysis not only intuitively presented the differences and commonalities in the overall score distribution between the two majors but also provided insights into the professional characteristics of students' knowledge mastery and ability application by examining the scores of different question types. This analysis offered data-driven insights for formulating cross-major teaching strategies.

2.4. Statistical methods

IBM's Statistical Package for the Social Sciences (SPSS) V.23 software (<https://www.ibm.com/analytics/spss-statistics-software>) was used to conduct statistical analyses of the relevant data. For quantitative data, the mean \pm standard deviation (SD) was used to represent the results, and *t*-tests were employed to assess the significance of differences between groups. For count data, rates (%) were used, and chi-square (χ^2) tests were conducted to determine statistical differences in score distribution between the two majors. A significance level of $P < 0.05$ was set to ensure the reliability and scientific nature of the research findings.

3. Results

3.1. Analysis of the EPC exam papers

The overall average difficulty of the exam paper was 0.7, falling within the range of 0.6–0.74, indicating that the overall difficulty was moderate. Further analysis of the different question types revealed that the average difficulty of objective questions was 0.72, while that of subjective questions was 0.68. Among the subjective questions, short-answer questions and analysis questions posed a greater challenge to students, resulting in a relatively higher overall difficulty for subjective questions. This may be due to these types of questions requiring students to have a strong ability to integrate knowledge and express it logically. Students often struggle to accurately extract key points of knowledge and present them in an organized manner, reflecting their weaknesses in knowledge application and thinking expansion. This finding provides a clear direction for adjusting future teaching priorities.

The reliability of the exam paper was 0.78, which falls within the range of 0.7–0.9, indicating an acceptable level of reliability (as shown in **Table 1**). After removing different question types, the reliability values of the exam paper fluctuated slightly, suggesting that the various question types in the exam paper have relatively stable internal consistency and the overall structure is reasonable. However, compared to the high-reliability standard (e.g., above 0.9), there is still room for improvement. Future efforts could focus on optimizing question-wording and enhancing the relevance of knowledge points to further refine the exam paper and improve reliability, ensuring that the exam results more accurately reflect students’ true abilities.

Table 1. Reliability analysis of the EPC exam paper

Item	Cronbach’s Alpha value
Entire exam paper	0.78
After removing [Single-choice questions]	0.72
After removing [Multiple-choice questions]	0.71
After removing [True-or-false questions]	0.76
After removing [Term explanation questions]	0.77
After removing [Short-answer questions]	0.76
After removing [Analysis questions]	0.71

The validity of each question type relative to the total score of the exam paper ranged from 0.7 to 0.8, indicating good validity (as shown in **Table 2**). This means that there is a strong positive correlation between students’ scores on each question type and their total scores, suggesting that the exam paper effectively measures students’ overall mastery of EPC knowledge.

Table 2. Validity analysis of the EPC exam paper

Question type	Validity
Single-choice questions	0.86
Multiple-choice questions	0.86
True-or-false questions	0.82
Term explanation questions	0.72
Short-answer questions	0.71
Analysis questions	0.85
Total score	1.00

3.2. Statistical analysis of exam scores for different majors

Overall score distribution: In this exam, a score of 90 and above was defined as excellent, with an overall excellence rate of only 5.4% (5 out of 94 students). Scores between 60 and 90 were considered passing, with a total pass rate of 73.3% (69 out of 94 students). Scores below 60 were considered failing, with a total failure rate of 21.3% (20 out of 94 students). The overall score distribution indicates a low excellence rate and a relatively high failure rate, reflecting that students’ performance in the EPC needs improvement. There are many deficiencies in knowledge mastery and application, and the teaching process requires deeper explanations and reinforcement training for key and difficult knowledge points to enhance students’ learning outcomes.

Comparison of scores between majors: A detailed comparison of scores between students of the PEM and ECEM revealed no statistically significant differences in pass rate, excellence rate, total scores, objective question scores, or subjective question scores ($P > 0.05$). This suggests that under the same teaching conditions, students from both majors have a similar level of knowledge mastery and ability in EPC. Further analysis indicates that this similarity is due to the shared teaching resources and assessment standards, which effectively eliminate the interference of teaching-related factors on scores. Additionally, as students in educational majors, both groups face the requirement of Teacher Qualification Exams and have a similar level of emphasis on the EPC, resulting in comparable learning motivation and engagement, and thus no significant difference in scores between the two majors.

4. Discussion

4.1. Analysis of the EPC exam papers

In summary, the analysis of the difficulty, reliability, and validity of this exam paper indicates that its overall design is scientifically sound and rational. The moderate difficulty ensures effective differentiation of students' knowledge levels, the acceptable reliability guarantees the relative stability and credibility of the exam results, and the good validity confirms that the exam paper accurately reflects students' mastery of the course content. However, in the face of the higher demands of teaching reform for student capability development, the exam paper still has room for improvement in guiding students to enhance their higher-order thinking and innovative application abilities. Future efforts should focus on strengthening the assessment of students' analysis, synthesis, and evaluation skills in exam question design, thereby promoting deeper development in teaching and evaluation.

4.2. Statistical analysis of exam scores

Reflection on learning status: The exam scores and question type distribution reveal that students currently rely heavily on rote memorization to cope with exams in EPC, especially obtaining scores in single-choice questions through extensive practice. However, they perform poorly in short-answer and analysis questions, which require deeper understanding and comprehensive application of knowledge. This learning approach is significantly different from the goals of teaching reform, which advocate fostering students' ability for autonomous learning, deep thinking, and problem-solving in real-life situations. In future teaching, instructors should actively change their teaching strategies by introducing diverse teaching methods, such as case-based teaching and project-based learning, to stimulate students' interest and guide them to actively construct their knowledge systems and enhance their ability to apply knowledge.

Impact of major differences: Although there is no significant difference in scores between students from different majors, the potential impact of major background on the learning process should still be considered. Students majoring in PEM may tend to understand EPC from the perspective of sports training and physical education practice, while those in ECEM may focus more on the application of knowledge in early childhood education scenarios. In teaching, instructors should fully explore the integration points between major characteristics and educational psychology knowledge to conduct targeted teaching activities, thereby promoting the organic combination of theoretical knowledge and professional practice among students and improving teaching effectiveness.

5. Conclusion

Exam paper analysis, as an important means of evaluating teaching effectiveness, plays an irreplaceable role in the process of teaching reform. Through a systematic analysis of the EPC exam papers and statistical analysis of the scores, teachers can gain precise insights into the quality of the exam papers and a deep understanding of students' knowledge mastery. This process also helps identify weak links in the teaching process and pain points in students' learning. In future teaching, teachers should closely follow the concepts of teaching reform, center on the students, and develop personalized teaching plans based on their learning situations. They should strengthen formative evaluation and feedback, guide students to change their learning methods, and gradually improve the quality of EPC teaching. This approach will lay a solid foundation for students' professional growth and career development, and contribute to the achievement of educational major talent cultivation goals.

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Wang X, 2015, Educational Psychology, Peking University Press, Beijing.
- [2] Tan H, Wang Y, 2021, Key Issues in Implementing the Overall Plan of Deepening the Reform of Educational Evaluation in the New Era. *Journal of China Examinations*, (8): 14–20.
- [3] Gao W, Li H, Cheng W, et al., 2024, Data Analysis and of the Final Exam Papers of Physiology from Traditional Chinese Medicine Undergraduates of Different Professional Directions. *Chinese Medicine Modern Distance Education of China*, 22(12): 185–188.
- [4] George D, Mallery P, 2016, SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 update (14th ed.), Allyn and Bacon, Boston.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.