

# 基于数据挖掘的教师教育质量评价指标体系的构建

孙昊天, 程肖冰\*, 杨世元, 鲍嘉泽, 杨 森

北京联合大学城市轨道交通与物流学院, 北京

收稿日期: 2024年9月28日; 录用日期: 2024年10月21日; 发布日期: 2024年10月28日

## 摘 要

目前的教师教育质量评价一般通过专家评判和学生打分两部分构成, 存在评价不够客观、可信度低等问题。为解决这些问题, 本研究通过问卷调查获取数据, 并对数据进行信效度检验, 利用层次分析法和优序图法, 构建了全面、客观的教师教育质量评价指标体系, 实现对教师教学过程的科学评估。结果表明, 基于数据挖掘的教育质量评价体系能够显著提高教师评价的准确性与公正性, 为优化教学管理、促进教师专业发展提供了有力支持。

## 关键词

教育质量评价, 层次分析法, 优序图法

# Construction of a Teacher Education Quality Evaluation Index System Based on Data Mining

Haotian Sun, Xiaobing Cheng\*, Shiyuan Yang, Jiaze Bao, Sen Yang

College of Urban Rail Transit and Logistics, Beijing Union University, Beijing

Received: Sep. 28<sup>th</sup>, 2024; accepted: Oct. 21<sup>st</sup>, 2024; published: Oct. 28<sup>th</sup>, 2024

## Abstract

The current evaluation of teacher education quality generally consists of two parts: expert judgment and student scoring, which suffers from a lack of objectivity and low credibility. To address

\*通讯作者。

文章引用: 孙昊天, 程肖冰, 杨世元, 鲍嘉泽, 杨森. 基于数据挖掘的教师教育质量评价指标体系的构建[J]. 数据挖掘, 2024, 14(4): 230-242. DOI: 10.12677/hjdm.2024.144021

these issues, this study collects data through questionnaires and conducts reliability and validity tests on the data. By employing the Analytic Hierarchy Process (AHP) and the prioritization matrix method, a comprehensive and objective evaluation index system for teacher education quality has been constructed, enabling a scientific assessment of the teaching process. The results indicate that the education quality evaluation system based on data mining can significantly improve the accuracy and fairness of teacher evaluation, providing strong support for optimizing teaching management and promoting the professional development of teachers.

## Keywords

Educational Quality Evaluation, Analytic Hierarchy Process (AHP), Priority Matrix Method

Copyright © 2024 by author(s) and Hans Publishers 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

## 1. 引言

国外针对教学质量的评估较早,美国匹兹堡大学开发研制的教学质量评估系统(IQA)主要从任务认知需求、解释性话语和教师期望三个评估维度对教师进行评估[1]。英国学科教学质量评估,分别从教学质量、学习环境和学生获得评价教学质量出发,“实用主义”倾向明显,更为看重课程教学与行业社会的紧密结合[2]。澳大利亚形成了两个有代表性的标准形态,即 ALTC 标准和 AUTCAS 标准,旨在推动教师回归教学本位,促进教学与科研之间的良好互动[3] [4]。

在国内,教育质量评价的研究已取得显著进展,主要聚焦于理论研究、基于数据挖掘的实证研究以及基于调查问卷的评估分析三大方向。1) 理论研究。胡瑞、李彩云以新公共服务理论为依据,从治理目标、治理主体和治理内容入手,构建了高校教师评价理论分析框架[5]。张丽丽、章政根据“密尔道德评价理论”“压力认知评价理论”“第四代评价理论”“普遍模仿验证理论”等经典评价理论对我国教师师德的评价原则、评价体系构成要件进行重点分析[6]。2) 基于数据挖掘的研究。程小恩、温川飙利用真实的数据,刻画教师教学评价的特征属性,进行以教师为主体的数据挖掘,对教师个体的教学、科研等状况进行准确的、实时的数据分析[7]。王丽丽以高校教师科研绩效为研究对象,综合运用科学计量学和智能信息处理技术,尝试量化方法与技术化路径,试探索更加科学、合理、有效的高校教师科研绩效评价模式[8]。3) 基于调查问卷的研究。廖敏(2016)利用问卷的信度、效度建立评价指标体系[9]。何克抗教授(2006)指出问卷调查要与课程深层次整合[10]。

尽管这些研究为教育实践提供了丰富的理论指导与实践参考,但仍存在若干不足:1) 理论研究。理论研究虽深入,但在具体操作层面的指导略显薄弱;2) 基于数据挖掘的研究。基于数据挖掘的研究虽能精准捕捉学习行为数据,但在评价标准的明确性和解释性上有所欠缺;3) 基于调查问卷的研究。基于调查问卷的研究,尽管能系统构建评价指标,但易受主观因素影响,导致结果的可靠性面临挑战,特别是“邓宁-克鲁格效应”的潜在影响不容忽视。

鉴于此,本研究旨在弥补上述不足,创新性地提出将问卷调查与数据挖掘技术相结合的研究方法。问卷调查能够系统地设计评价指标,确保评价维度的全面性和明确性,同时通过严谨的问卷设计和数据分析流程,减少主观因素的干扰,提升评价结果的可靠性。而数据挖掘技术则以其强大的数据处理能力,深入挖掘学习行为数据背后的规律和模式,为教育质量评价提供更为客观、精确的数据支持。二者的有

机结合,不仅能够增强评价体系的科学性和有效性,还能促进教育质量评价在教育实践中的广泛应用与深入发展。

## 2. 教师教育质量评价指标体系的构建

### 2.1. 理论基础

#### 2.1.1. 李秉德教育理论

李秉德教学七要素是研究者们不断探讨和提高的重要内容。李秉德教学理念注重求真、尊重个人,以及教师应当重申“教得恰到好处”。李秉德提出的七要素包括:一是把学生作为主体放在教学者的位置;二是教之要有法,而其法即是按照自然规律进行;三是合理组织内容,引导学生理解;四是真切陈述观念;五是采用灵活多样的教学方式,适应学生的不同发展水平;六是在受教者中树立优越感;七是学生的学习兴趣,自主性和愉快感要被有效激发。

根据李秉德教育理论,本文以学生为中心,通过教师实现教学目的,以专家视角和学生视角对教师教学评价指标从以下七个方面进行设计。

1) 教师:教师的思想及业务水平、个性修养、教学态度、教学方法及能力等构成了教学活动中起主导作用的因素,否则学习活动就变成了自学。

2) 教学目的:教学目的是教学活动必不可少的要素之一,不同层次、不同性质或不同方面的目的形成的完整的体系或结构,将落实在学生身上。

3) 教学内容:教学内容或课程在教学活动中最具有实质性的因素,指的是一定的知识、能力、思想与情感等方面的内容组成的结构,具体表现为教学计划或课程方案、各科教学大纲或学科课程大纲、文字的或音像的各科教材。

4) 教学方法:教学方法是教师课内外所使用的各种教学方法、教学艺术、教学手段和教学组织形式,有显性和隐性之分。

5) 教学反馈:教学反馈是教学在师生之间信息传递的交互性决定的。

6) 教学环境:教学环境包括有形和无形两种。

7) 学生:学生是学习的主体,是教学活动的根本要素,指学生的身心发展水平、已有的知能结构、个性特点能力倾向和学习前的准备情况等。

#### 2.1.2. 建构主义理论

建构主义教育理论自20世纪后叶盛行至今,其思想最早是由瑞士心理学家皮亚杰通过研究儿童的认知规律提出来的,后来又汲取了维果斯基“最近发展区”理论和奥苏贝尔有意义接受学习理论的精髓,因此,较好地揭示了人类学习过程的认知规律(熊英,2010)[11]。建构主义理论是认知心理学的一个重要分支,近年来逐渐在西方盛行。建构主义认为世界是客观存在的,但对世界的理解和赋予意义却由每个人自己决定。建构主义教育理论虽然流派纷呈、内容丰富,但概括起来,其核心内涵是:以学生为中心,强调学生对知识的主动探索,个人在与周围环境的相互作用过程中逐渐构建起关于外部世界的知识,发现和对所学知识意义的主动建构过程。建构主义者认为,人们是在原有知识和经验的基础上以自己的经验为基础来构建或解释现实。学习者在接触任务时,有不同的先前知识,这些知识由其背景经验、先前学习及世界观、信念构成。学习者必须激活原有的经验,用自己的方式确保新的知识与已有知识之间形成意义的联结。进一步讲,知识是学习者在一定的情境下,借助他人(包括教师和学习伙伴)的帮助,运用必要的学习材料,通过意义建构的方式而获得的。

建构主义所蕴含的教学思想主要反映在知识观、学习观、学生观、师生角色的定位及作用等方面。

张建伟、陈琦将认知建构主义运用到教学,提出认知灵活性理论及其随机通达教学、自上而下的教学设计及知识结构的网络概念、情境性教学、支架式教学和教学中的社会性相互作用等理论[12]。

### 2.1.3. 多元智能

多元智能理论(The Theory of Multiple Intelligences)是美国哈佛大学教授、发展心理学家霍华德·加德纳(Howard Gardner)在对多元智能的长期研究后提出的观点,他认为人类的智力至少含有七项基本内容:语言智力(Linguistic Intelligence)、逻辑-数理智力(Logical-Mathematical Intelligence)、空间智力(Spatial Intelligence)、肢体-运动智力(Bodily-Kinesthetic Intelligence)、音乐智力(Musical Intelligence)、人际智力(Interpersonal Intelligence)以及内省智力(Introspective Intelligence)。这一理论认为学生拥有多种智力,强调人各有所长,重视学生的多元智慧,以多元弹性的评价方式发挥学生的不同智慧并且以此进行评价进而发展有效教学。加德纳分别于1997年和1999年对其理论进行了补充与完善,使原先确认的7种智能类别拓展到现在的9种,增加了自然观察(Naturalist)和存在理念(Existential)智能。

多元智能理论认为学习者个体的学习智能是有差异的。每个人都有其强项智能和弱项智能之分,以多元智能理论为指导的教育评价打破传统的智能测验的片面性与局限性,力图真实准确地反映学生智能。基于多元智能理论可以得出英语能力评价应该具有以下特点:学生的智能是多元化的;评价也应该是多元的;不同学生不同阶段表现出的智能特点不一样;评价与教学过程融为一体;评价要关注学生不同阶段的成长;学生的自我评价在整个评价中非常重要;传统的标准化测验无法衡量出每位学生的多元的潜能;重视非正式的评价。由于不同学生学习经历、家庭、社会、经济、文化等因素的不同影响,构成他们不同的禀赋、性向、才情、智能、喜好也有所不同,导致他们的学习兴趣、学习方式、学习能力各有差异[13]。

## 2.2. 教师教育评价指标的构建

以“翻转课堂/颠倒课堂/反转课堂/翻转教室/颠倒教室/反转教室”“翻转教学/颠倒教学/反转教学”“学业水平/学习成绩/学业成就/学业表现”为关键词检索中外文文献6568篇,去重后获得文献6290篇,筛选题目、摘要后共计2402篇文献,结合北京联合大学教师教学创新大赛评分标准,参考混合式教学有效性评价研究与实践、研究教育质量评价定义、目标,遵守问卷设计的原则,结合李秉德教学七要素对教师指标进行维度划分,初步设计教育质量评价指标体系:一级指标有5个,二级指标22个。

### 2.2.1. 初步的教师教育质量评价指标的构建

组织第一次问卷发放、回收后,汇总分析问卷意见,增加教师基础评价指标:教师间相互帮助,相互指导纠正;教师不迟到、早退,教学富有热情。形成新的评价指标咨询问卷,通过对每一项评价指标及其评价标准的专家意见进行量化处理,对“合适”、“修改后合适”和“不合适”分别赋值3、2、1。利用Excel计算出各项指标的平均值、标准差等数据(例:平均值:“=AVERAGE(G2:G19)”标准差:“=STDEV(G2:G19)”),重复上述方法,至问卷咨询意见达到统一且指标意见评价总体满意。在第二次问卷回收并修改后,达到上述要求,得到初步的教师教育质量评价指标模型。

### 2.2.2. 教师教育质量评价指标体系的有效性检验

2023年7月20日到2023年7月30日,通过“问卷星”的形式,进行了第三轮专家问卷发放和回收,专家问卷共收回问卷207份,其中有效问卷207份。

为了便于对问卷结果进行数据计算分析,问卷主体结构分为一级指标、二级指标、意见选项三个部分。其中,意见选项设置为“非常合适”、“合适”、“一般”、“不合适”、“非常不合适”五个,量化处理分别赋值5、4、3、2、1。问卷形式见图1。

\* 1. 从“教学目标”方面评价

	非常合适	合适	一般	不合适	非常不合适
1.1 制定大多数学习者能够达到预定的教育目标	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.2 课程设计符合大多数学习者的需求和背景 (文化背景、学习风格和能力水平)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.3 设定了“以学生发展为中心”的课程目标, 培养学生主动学习、批判思维、创新创造和动手实践等能力	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.4 设立了明确合理的单元目标, 培养学生知识理解和技能应用的能力	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1. Questionnaire format  
图 1. 问卷形式

本文采用 SPSSAU 进行信度分析、效度分析、验证性因子分析和相关分析对问卷结果进行分析计算。  
数据分析步骤如下:

- Step 1: 信度分析, 确定问卷指标的可信度;
- Step 2: 效度分析, 确定问卷能否准确反映问题;
- Step 3: (验证性)因子分析, 说明各个指标间的关系;
- Step 4: 相关度分析, 验证各个二级指标间是否冲突。

1) 信度分析

信度即可靠性、可信度, 用于检测数据是否真实, 问卷填写者是否认真回答问题、回答是否可信等。一般使用克隆巴赫信度系数进行测量, 计算模型如式 1 所示, 如果  $\alpha$  系数大于 0.6 则表示通过信度检验。

$$\alpha = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum_{i=1}^k S_i^2}{S_x^2} \right) \quad (1)$$

式中,  $k$  为用于测量的题目数,  $S_i^2$  为第  $i$  个题目得分(测量值)的方差,  $S_x^2$  为量表总得分的方差。

计算过程如下:

$$\text{Step 1: } S_i^2 = \frac{\sum_{j=1}^{207} x_j^2 - 207 \bar{x}^2}{207} \quad (x_j \text{ 是第 } j \text{ 名专家或教师对该题的打分})$$

$$\text{Step 2: } S_x^2 = \frac{\sum_{i=1}^{24} \sum_{j=1}^{207} x_{ij}^2 - 24 \times 207 \bar{x}^2}{24 \times 207} \quad (x_{ij} \text{ 表示第 } j \text{ 名专家或教师对第 } i \text{ 题的打分})$$

$$\text{Step 3: } \alpha = \left( \frac{24}{24-1} \right) \left( 1 - \frac{\sum_{i=1}^{24} S_i^2}{S_x^2} \right)$$

SPSSAU 具体操作过程:

SPSSAU 菜单栏里的“问卷研究”中选择“信度”，在筛选样本中点击有效样本，选择方法有 Cronbach’s  $\alpha$  法、折半系数法、McDonald Omega 法和 theta 系数法四种方法，本次研究方法为 Cronbach’s  $\alpha$  法即内部一致性来检测该问卷信度。点击“开始分析”。SPSSAU 系统会根据上述过程自动计算出标准化 Cronbach’s  $\alpha$  系数并列出 Cronbach 信度分析表格，表格中也包含每一指标的校正项总计相关性(CITC)、项已删除的  $\alpha$  系数。

综上：得该问卷的信度系数值  $\alpha$  为 0.892，大于 0.8，因而说明研究数据信度质量高。

2) 效度分析

效度研究用于分析研究项是否合理有意义，使用因子分析进行数据分析。分别通过 KMO 值、共同度、方差解释率值、因子载荷系数值等指标进行综合分析，以验证出数据的效度水平情况。

KMO 值用于判断信息提取的适合程度，共同度值用于排除不合理研究项，方差解释率值用于说明信息提取水平，因子载荷系数用于衡量因子(维度)和题项对应关系。计算模型如式 2 所示，KMO 值大于 0.8，则表示研究数据非常适合提取信息(从侧面反应出效度很好)。

$$KMO = \frac{\sum \sum_{i \neq j} r_{ij}^2}{\sum \sum_{i \neq j} r_{ij}^2 + \sum \sum_{i \neq j} \alpha_{ij}^2} \quad (2)$$

式中  $r_{ij}$  表示简单相关系数， $\alpha_{ij,1,2,3,\dots,k}^2$  表示偏相关系数。同样可根据 SPSSAU 自动得出其他对应相关值。

SPSSAU 具体操作过程：

SPSSAU 菜单栏里的“问卷研究”中选择“效度”，在筛选样本中点击有效样本，维度个数设置为 5 个，点击“开始分析”。SPSSAU 系统会根据式 2 自动计算出标准化 KMO 值并列出 KMO 和 Bartlett 的检验表，见表 1。

**Table 1.** Preliminary KMO and Bartlett's test table  
**表 1.** 初步 KMO 和 Bartlett 检验表

KMO 值		0.885
Bartlett 球形度检验	近似卡方	2263.780
	df	276
	p 值	0.000

综上：由表得 KMO 值为 0.885，KMO 值大于 0.8，研究数据非常适合提取信息。

3) 验证性因子分析

通常使用标准载荷系数值表示因子与分析项(测量项)间的相关关系；如果某项呈现出显著性，且标准载荷系数值大于 0.7，则说明有着较强的相关关系；如果某项没有呈现出显著性，或者标准载荷系数值较低(低于 0.4)，则说明该项与因子间关系弱，可考虑移除该项。

SPSSAU 具体操作过程：

SPSSAU 菜单栏里的“问卷研究”中选择“验证性因子分析”，在筛选样本中点击有效样本，对五个因子(Factor)分别命名为：教学目的、教学内容、教学方法、教学反馈、教师环境，点击“开始分析”。SPSSAU 系统自动整理出 Component Matrix (初始因子载荷矩阵)与区分效度表，见表 2，从而可得出标准化载荷系数。

综上：本次问卷的标准化载荷系数绝对值均大于 0.6 且呈现出显著性，意味着有着较好的测量关系，说明这些二级指标可以很好地测量对应的一级指标。

**Table 2.** Differentiate validity table  
**表 2.** 区分效度表

	教学目标	教学内容	教学方法	教学反馈	教学环境
教学目标	0.689				
教学内容	0.436	0.757			
教学方法	0.411	0.223	0.684		
教学反馈	0.480	0.213	0.281	0.718	
教学环境	0.471	0.210	0.296	0.289	0.794

注：对角线数字为 AVE 平方根值。

AVE 计算模型如式 3 所示。

$$AVE = \frac{\sum \text{因子负荷量}^2}{(\sum \text{因素负荷量})^2 + (\sum \text{各测测量变项的测量误差})} \quad (3)$$

分析表 2 可以得出：

- ① 针对教学目标，其 AVE 平方根值为 0.689，大于因子间相关系数绝对值的最大值 0.480，意味着其具有良好的区分效度；
- ② 针对教学内容，其 AVE 平方根值为 0.757，大于因子间相关系数绝对值的最大值 0.436，意味着其具有良好的区分效度；
- ③ 针对教学方法，其 AVE 平方根值为 0.684，大于因子间相关系数绝对值的最大值 0.411，意味着其具有良好的区分效度；
- ④ 针对教学反馈，其 AVE 平方根值为 0.718，大于因子间相关系数绝对值的最大值 0.480，意味着其具有良好的区分效度；
- ⑤ 针对教学环境，其 AVE 平方根值为 0.794，大于因子间相关系数绝对值的最大值 0.471，意味着其具有良好的区分效度。

综上：这些一级指标间有很好的区分度，互不冲突。

4) 相关分析

Pearson 相关系数计算模型如式 4 所示。

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

SPSSAU 具体操作过程：

SPSSAU 菜单栏里的“通用方法”中选择“相关”，在筛选样本中点击有效样本，分析方法有 Pearson 相关系数法、Spearman 相关系数法、Kendall 相关系数法三种，本次研究选取 Pearson 相关系数法比较各个指标简单相关关系。点击“开始分析”，SPSSAU 系统自动整理出 Pearson 相关 - 标准格式并给出 Pearson 相关可视图，见图 2。

由图 2 可得：同级别一级指标对应二级指标形成的三角形区域红色深度较深，再一次说明各一级指标间有很好的区分度，互不冲突。对同级别一级指标对应的二级指标之间的相关系数进行分析，相关系数较大的两项合并。

综上：根据数据分析结果，相应增加、修改、删除评价指标，得到完善后的教师教育质量评价指标体系。

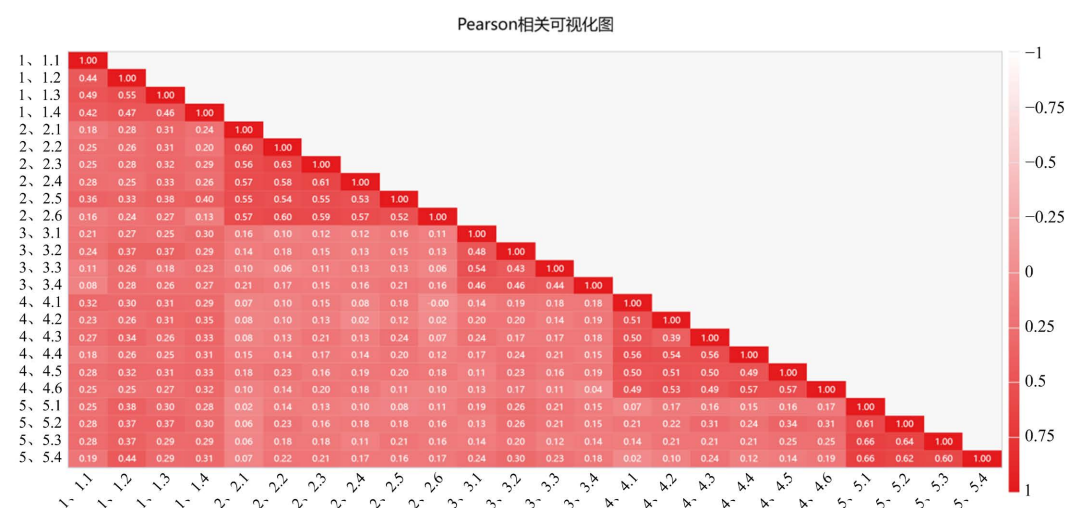


Figure 2. Pearson related viewable  
图 2. Pearson 相关可视图

3. 教师教育评价指标体系权重的确定

3.1. 数据来源

根据第三次问卷发放的结论及专家、教师和同学们提出的合理建议，第四次问卷设置进行相应调整。因为本次问卷结果要进行指标体系权重分析，问卷分为两类，一级指标权重问卷主体结构分为一级指标、权重分配(将 100 分配给以下五个一级指标)两个部分；二级指标权重问卷主体结构分为二级指标、意见选项两个部分，其中，意见分为“非常合适”“合适”“一般”“不合适”“非常不合适”五个选项。

2023 年 8 月 8 日到 2023 年 8 月 17 日，笔者做了第四轮教师和学生咨询，通过“问卷星”的形式，将两份问卷通过微信分别相应发放给专家教师和学生。专家教师一级指标问卷共收回问卷 50 份，其中有效问卷 50 份，二级指标问卷共收回问卷 50 份，其中有效问卷 50 份；学生一级指标问卷共收回问卷 98 份，其中有效问卷 89 份，二级指标问卷共收回问卷 101 份，其中有效问卷 100 份。

3.2. 权重计算

3.2.1. 基于 AHP 层次分析法的一级指标权重计算

B65												
$=(B59*B64+C59*C64+D59*D64+E59*E64+F59*F64)/B64+(B60*B64+C60*C64+D60*D64+E60*E64+F60*F64)/C64+(B61*B64+C61*C64+D61*D64+E61*E64+F61*F64)/D64+(B62*B64+C62*C64+D62*D64+E62*E64+F62*F64)/E64+(B63*B64+C63*C64+D63*D64+E63*E64+F63*F64)/F64$												
	A	B	C	D	E	F	G	H	I	J	K	L
1	序号	教学目的	教学内容	教学方法	教学反馈	教师环境						
47	46	23	30	16	11	20						
48	47	24	10	41	13	12						
49	48	11	10	59	10	10						
50	49	15	12	46	11	16						
51	50	12	44	21	11	12						
52	平均值	17.32	23.64	34.52	12	12.52						
53		1	0.732657	0.501738	1.443333	1.383387						
54		1.364896	1	0.68482	1.97	1.888179						
55	构建判断矩阵	1.993072	1.460237	1	2.876667	2.757188						
56		0.692841	0.507614	0.347625	1	0.958466						
57		0.722864	0.529611	0.362688	1.043333	1						
58	整列求和	5.773672	4.230118	2.896871	8.333333	7.98722						
59		0.1732	0.1732	0.1732	0.1732	0.1732						
60		0.2364	0.2364	0.2364	0.2364	0.2364						
61	归一化处理	0.3452	0.3452	0.3452	0.3452	0.3452						
62		0.12	0.12	0.12	0.12	0.12						
63		0.1252	0.1252	0.1252	0.1252	0.1252						
64	权重	0.1732	0.2364	0.3452	0.12	0.1252						
65	最大特征值	5										
66	CI	0		RI	1.12							
67	CR	0										

Figure 3. Steps to calculate the weights of the first-level indicators  
图 3. 一级指标权重计算步骤

一级指标问卷问题少,目标明确,得出问卷填写结果威信性较高,适用于 AHP 层次分析法。

教育质量一级指标层次分析法的计算步骤见图 3。

从上图 3 可知,针对教学目的、教学内容、教学方法、教学反馈、教师环境总共 5 项,构建 5 阶判断矩阵进行 AHP 层次法研究(计算方法为:和积法),分析得到特征向量为(0.866, 1.182, 1.726, 0.600, 0.626),并且总共 5 项对应的权重值分别是: 17.320%、23.640%、34.520%、12.000%、12.520%。

本次针对 5 阶判断矩阵计算得到 CI 值为 0.000, 针对 RI 值查表为 1.120, 因此计算得到 CR 值为  $0.000 < 0.1$ , 意味着本次研究判断矩阵满足一致性检验, 计算所得权重具有一致性。

进一步利用 Matlab 进行检验, 见图 4。

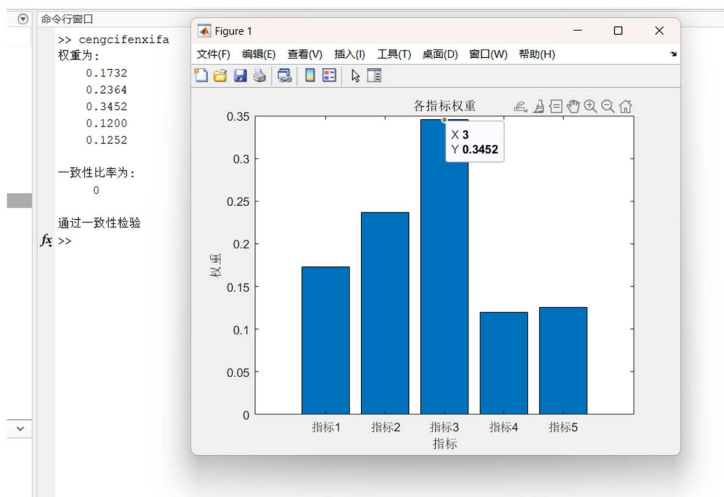


Figure 4. First-level index calculation test  
图 4. 一级指标计算检验

3.2.2. 基于优序图法的二级指标权重计算

对于优序图调查表, 被调查对象只需要考虑三个数字, 更加便于其在多因素多指标调查中理解相关目标, 做出合理判断, 无需像成分分析那样考虑每个对比目标的重要程度。这些对于被调查者的要求大大降低, 也提高了调查问卷的返回率和准确率。因此二级指标权重分析采用优序图法。

就二级指标“教学目的”为例, excel 的教育质量二级指标优序图法的计算步骤见图 5。

	A	B	C	D	E	F	G
1	序号	1.1 制定大多数学习	1.2 课程设计符合大多	1.3 设定了“以学生发展为中	1.4 设立了明确合理的单元目标, 培养学生知识理解		
42	41	2	2	1	2		
43	42	3	5	3	3		
44	43	5	5	5	5		
45	44	2	2	2	2		
46	45	5	5	1	3		
47	46	4	4	4	4		
48	47	4	5	4	4		
49	48	3	4	3	1		
50	49	5	4	5	5		
51	50	4	5	1	4		
52	平均值	3.6	3.62	3.38	3.2	求和(TTL)	权重
53		0.5	0	1	2.5	0.3125	
54	优序图权	1	0.5	1	1	3.5	0.4375
55	重计算表	0	0	0.5	1	1.5	0.1875
56		0	0	0	0.5	0.5	0.0625

Figure 5. The “Teaching Objectives” weight calculation steps  
图 5. “教学目标”权重计算步骤

进一步利用 Matlab 进行检验并完成其余四个一级指标的计算。见图 6、图 7、图 8、图 9 和图 10。

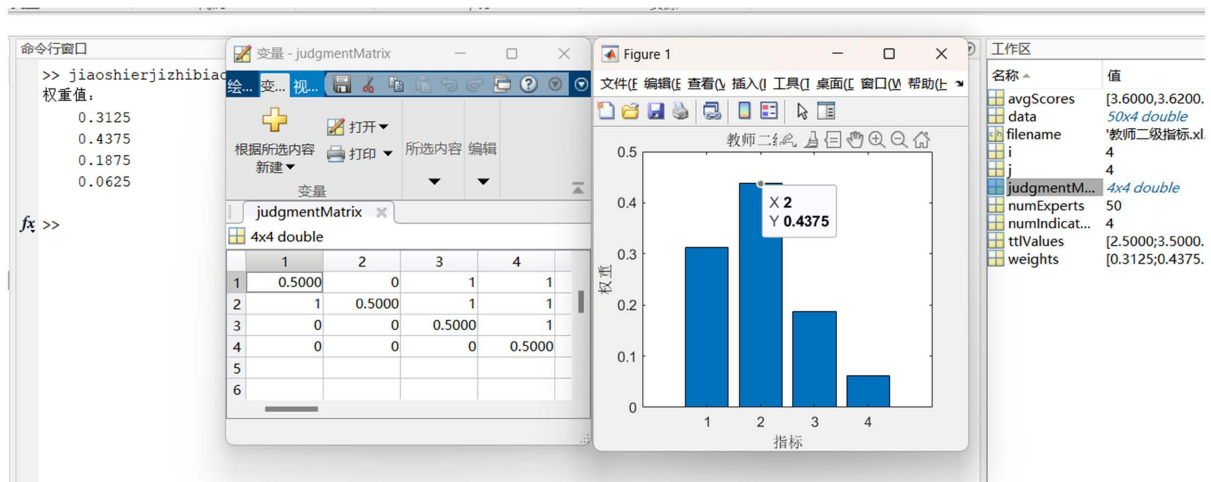


Figure 6. The “Teaching Objectives” weight calculation test

图 6. “教学目标” 权重计算检验

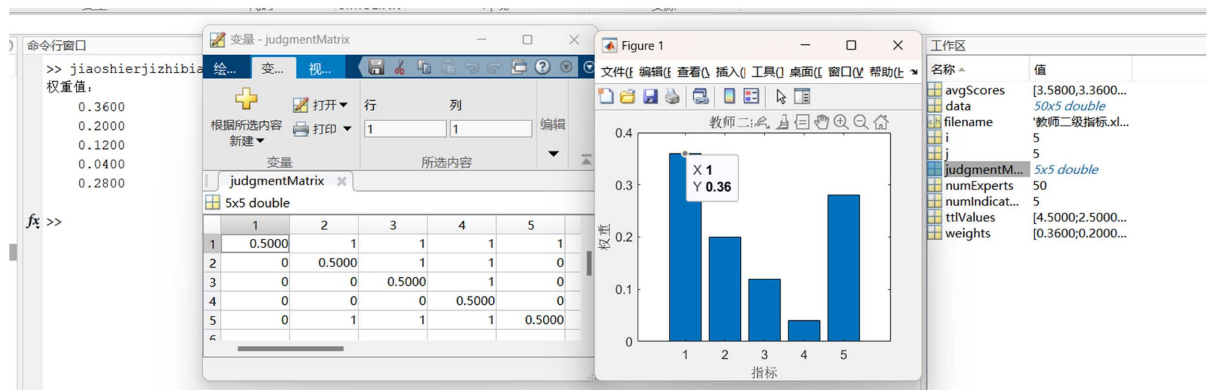


Figure 7. “Teaching Content” weight calculation test

图 7. “教学内容” 权重计算检验

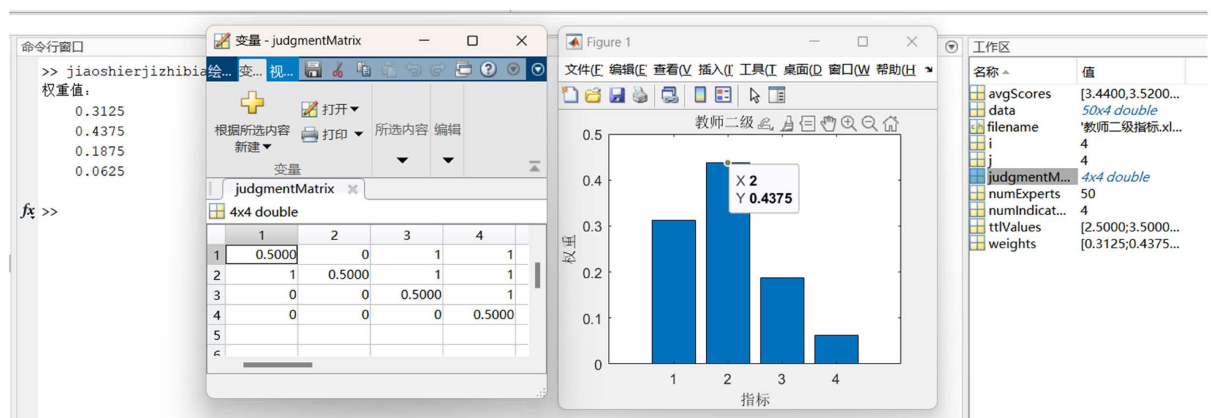


Figure 8. “Teaching Method” weight calculation test

图 8. “教学方法” 权重计算检验

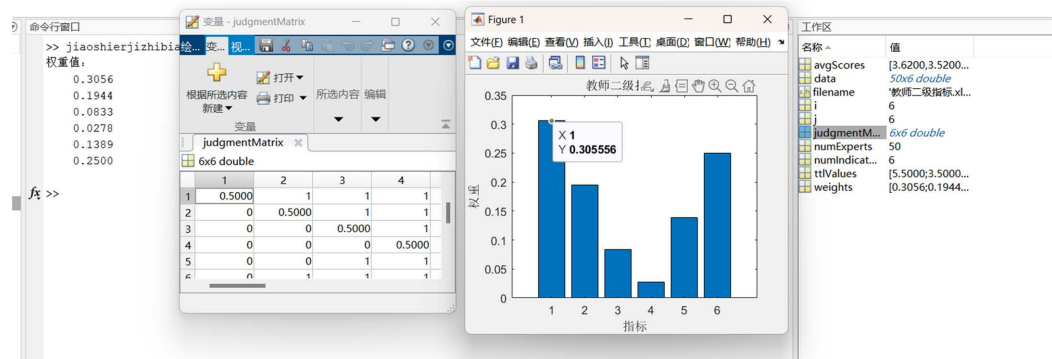


Figure 9. “Teaching Feedback” weight calculation test  
图 9. “教学反馈” 权重计算检验

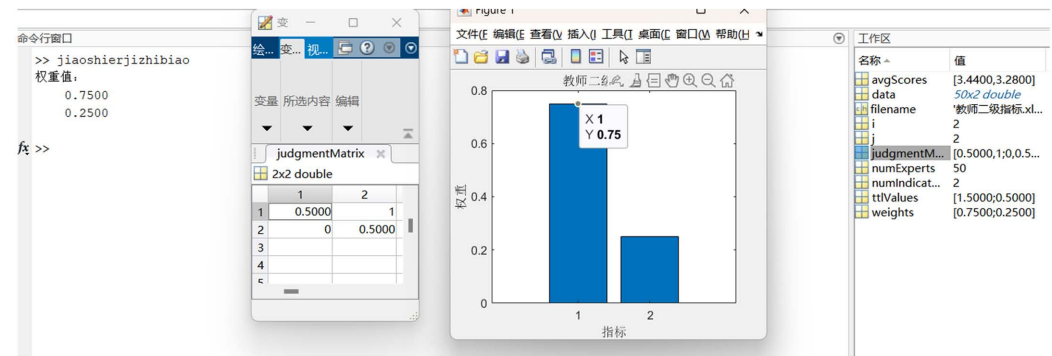


Figure 10. “Teacher Environment” weight calculation test  
图 10. “教师环境” 权重计算检验

3.3. 教师评价指标权重分配结果与建模

综上所述，教师教育质量评价指标体系见下表 3。

Table 3. Evaluation indicators system for teacher education quality  
表 3. 教师教育质量评价指标体系

一级指标	一级指标权重(%)	二级指标	二级指标权重(%)	最终权重(%)
教学目的	17.32	制定大多数学习者能够达到预定的教育目标	31.25	5.4125
		课程设计符合大多数学习者的需求和背景(文化背景、学习风格和能力水平)	43.75	7.5775
		设定了“以学生发展为中心”的课程目标，培养学生主动学习、批判思维、创新创造和动手实践等能力	18.75	3.2475
		设立了明确合理的单元目标，培养学生知识理解和技能应用的能力	6.25	1.0825
教学内容	23.64	教学内容与教学目标紧密结合起来	36.00	8.5104
		教学内容有深度(内容覆盖学科核心概念和原理)和广度(内容覆盖学科多个领域和相关知识点)	20.00	4.728
		使用质量较高的教学资源(权威教材和参考书籍、优质网络资源)	12.00	2.8368
		教学内容满足行业与社会需求，具有科学性	4.00	0.9456
		结合所授课程特点、思维方法和价值理念，深挖课程思政元素，有机融入课程教学	28.00	6.6192

续表

教学方法	34.52	教学方法多样, 如讲授、演示、讨论、案例分析等	31.25	10.7875
		教学方法能够根据学习者的学习成绩、知识掌握程度而差异化处理(对于低成就者, 有特殊的支持和关注; 对于高成就者, 鼓励其参加拓展竞赛等)	43.75	15.1025
		设计有趣且具有挑战性的教学活动(如: 以互动、游戏、形象举例方式)	18.75	6.4725
		尊重学习者的思想和感受, 注重情感教育	6.25	2.1575
教学反馈	12.00	评价过程应公开、透明, 评价者和被评价者都能够了解评价的流程和标准	30.56	3.6672
		评价过程充分考虑了不同学习者的差异性(记录学生成长档案, 横向与纵向比较)	19.44	2.3328
		教育方法可根据学习者的评价和反馈而灵活改变	8.33	0.9996
		能够全面评价学习者的知识、技能和能力的发展(学习成绩、作业表现、项目成果、口头表达能力、批判性思维能力、合作能力等方面)	2.78	0.3336
		及时向学生提供反馈信息, 以促进其学习	13.89	1.6668
		将评价标准与教学目标紧密结合起来	25.00	3
教师环境	12.52	获得奖项(比赛、项目)	75.00	9.39
		教师不迟到、早退, 教学富有热情	25.00	3.13

根据如上评价体系, 进行建模, 得到量化的打分, 见式 5。

$$\gamma = \sum \sum \alpha_{ij} \chi_{ij} \tag{5}$$

$\gamma$  为教师总分,  $i$  为第  $i$  个一级指标,  $i = 1, 2, 3, 4, 5$ ,  $j$  为第  $j$  个二级指标,  $\alpha_{ij}$  为第  $i$  个一级指标的第  $j$  个二级指标的权重,  $\chi_{ij}$  为教师在第  $i$  个一级指标的第  $j$  个二级指标的对应分数。

4. 结论

综上, 该评价体系全面覆盖教学目的、内容、方法及反馈等多个维度, 确保评价的系统性和科学性。具体而言, 教学方法的权重以 34.52% 位居首位, 这充分证明了教学方法的多样性和个性化对提升教学质量的关键作用。教学方法中, 差异化处理学习者需求(43.75%)和设计有趣且具有挑战性的教学活动(18.75%)尤为关键, 体现了对学生个体差异的尊重和激发学习兴趣的重要性。紧随其后的教学内容权重为 23.64%, 凸显了教学内容与教学目标的紧密结合、内容的深度和广度、教学资源的质量以及满足行业与社会需求的重要性。特别是, 深挖课程思政元素并有机融入课程教学的策略(28%), 展示了教育不仅是知识的传授, 更是价值观塑造的过程。教学目的(17.32%)、教学反馈(12%)和教师环境(12.52%)虽在总权重中占比较低, 但各自的作用亦不容忽视。教学目的强调了学生中心的教学理念, 包括设定合理的课程目标、培养多元能力等; 教学反馈则关注教学过程的信息传递与调整, 促进教与学的良性互动; 而教学环境则包括物理与心理环境, 对学生的学习体验和学习效果产生重要影响。这一体系为优化教师教育提供了实践指导和启示。

在督导组的带领下, 为了响应本校教学比赛的号召, 将该教师教育质量评价指标体系实践在我校 21 级中俄合作教学的管理基础课程中, 通过低代码构建平台亦搭在钉钉上搭建相关教学平台, 实时记录实证班级教育质量评价的具体实施准备与实施过程, 结合实证班级教育质量评价成效检测, 重点思考运用指标时遇到的问题, 与 20 级学生进行纵向对比, 试点班级的教学质量平均提升了 16.81%, 其中, 学生学业成绩优秀率提高了 12.62%, 学习兴趣和满意度分别提升了 36.71% 和 18.95%。同时, 教师的教学热情

和创新能力也得到了显著提升，具体表现为教学方法的多样化、教学反馈的及时性和教学环境的优化。这些具体的量化数据不仅验证了本研究的科学性和先进性，更为教育管理者提供了宝贵的实践经验和决策依据，为全面提升教师教育质量奠定了坚实基础。

## 基金项目

北京联合大学 2023 年教育教学研究与改革项目(JJ2023Y033)；北京市教委科技一般资助项目(KM202011417002)。

## 参考文献

- [1] 周九诗, 鲍建生. 美国 IQA 教学质量评估系统的探索与反思[J]. 比较教育学报, 2021(5): 87-99.
- [2] 丁磊. 英国高等教育质量评估新进展[D]: [硕士学位论文]. 石家庄: 河北师范大学, 2021.
- [3] 李作章, 张雷生. 从“有效教学”走向“优质教学”: 澳大利亚大学教学质量标准的演进逻辑[J]. 大学教育科学, 2020(5): 112-119.
- [4] 魏振锋, 汪焰. 国家级职业教育教师教学创新团队教学质量评价的实证研究[J]. 现代职业教育, 2024(20): 1-4.
- [5] 胡瑞, 李彩云. 高校教师评价改革的“破”与“立”——基于新公共服务理论的分析框架[J]. 现代教育管理, 2024(8): 1-9.
- [6] 张丽丽, 章政. 新时代我国教师师德考核评价体系研究: 评价依据、借鉴与要件[J]. 教育科学研究, 2024(7): 77-82.
- [7] 程小恩, 温川飙. 基于线上线下混合教学的教师教学能力三维评价模式的研究[J]. 电脑知识与技术, 2023, 19(1): 143-145.
- [8] 王丽丽. 高校教师科研绩效量化评价研究[D]: [博士学位论文]. 哈尔滨: 哈尔滨师范大学, 2017.
- [9] 廖敏. 问卷的信度和效度以及 SPSS 软件运用[J]. 科技展望, 2016, 26(16): 290.
- [10] 张文兰, 何克抗. 网络环境下小学英语教学设计型研究及效果分析[J]. 教育信息化, 2006(12): 7-10.
- [11] 熊英. 建构主义大学英语教学设计[J]. 教育理论与实践, 2010, 30(12): 53-55.
- [12] 张建伟, 陈琦. 从认知主义到建构主义[J]. 北京师范大学学报(社会科学版), 1996(4): 75-82, 108.
- [13] 霍艳娟. 中国 EFL 学习者英语能力评价研究[D]: [博士学位论文]. 上海: 上海外国语大学, 2017.