

基于多参数项目反应理论的大学生物理教学效果实证研究

贾 飞 张 睿 赵 晖 时 钟

(同济大学物理科学与工程学院, 上海 200092)

摘 要 随着教育数据挖掘技术的迅速发展,教育工作者可以借助先进的数据分析方法深入了解和优化教学过程。本研究结合多维项目反应理论和教育数据挖掘技术,采用 R 软件中的 MIRT 程序包对 2022—2023 学年下学期至 2023—2024 学年下学期的三个学期的大学物理期末考试数据进行分析。研究重点关注学生在记忆、概念理解和综合应用三个维度上的学业表现,并通过对学生多维能力的估计,评估教学内容和方法的有效性。研究结果表明,课后作业和课堂测试与期末考试各维度成绩之间存在显著的正相关性,尤其是在综合应用能力方面,课后作业的设计对学生的综合能力提升起到了重要作用。此外,研究还通过终结性评价数据的异常分析,识别了在特定维度上表现不佳的班级,提出了相应的改进措施。基于聚类分析的结果,教学班级被划分为表现优秀、良好和待改进三类,为教学改进提供了数据支持。最后,本研究为教学管理者提供了基于数据的决策支持,能够帮助其识别教学中的强项与弱项,优化资源配置,推动教学质量的提升。

关键词 多维项目反应理论;教育数据挖掘;教学效果

EMPIRICAL EVALUATION OF COLLEGE PHYSICS INSTRUCTION BASED ON A MULTIDIMENSIONAL ITEM RESPONSE THEORY FRAMEWORK

JIA Fei ZHANG Rui ZHAO Hui SHI Zhong

(College of Physics and Engineering, Tongji University, Shanghai 200092)

Abstract As educational data mining (EDM) technology advances rapidly, educators can gain in-depth insights into and optimize the teaching process with advanced data analysis methods. This study combines Multidimensional Item Response Theory (MIRT) with educational data mining. It uses the MIRT package in R software to analyze college physics final exam data from three semesters, from the 2022—2023 spring semester to the 2023—2024 spring semester. The focus is on student academic performance in memory, conceptual understanding, and comprehensive application. By estimating students' multidimensional abilities, it evaluates the teaching content and methods' effectiveness. Results show a significant positive correlation between after class homework, classroom test scores and final exam performance across dimensions, particularly in comprehensive application ability, where homework design is key for improvement. Moreover, the study identifies underperforming classes in

收稿日期: 2025-05-20

基金项目: 大学物理自适应智慧课程研究, 2025 同济大学教改项目; 以暑期物理提高班为载体的大学物理课程基础能力强化教学改革研究, 同济大学本科教育研究与改革项目(2025), 项目号: 20250504。

通信作者: 张睿, 98733@tongji.edu.cn。

引文格式: 贾飞, 张睿, 赵晖, 等. 基于多参数项目反应理论的大学生物理教学效果实证研究[J]. 物理与工程, 2025, 35(5): 46-54.

Cite this article: JIA F, ZHANG R, ZHAO H, et al. Empirical evaluation of college physics instruction based on a multidimensional item response theory framework[J]. Physics and Engineering, 2025, 35(5): 46-54. (in Chinese)

specific dimensions through analyzing anomalies in summative assessment data and suggests improvements. Based on cluster analysis, teaching classes are categorized into excellent, good, and needing-improvement groups, offering data-backed support for teaching improvement. Lastly, this study provides data-driven decision making support for teaching managers. It helps them identify teaching strengths and weaknesses, allocate resources better, and enhance teaching quality.

Key words multidimensional item response theory; educational data mining; teaching effectiveness

在当今的教育领域,如何有效提升教学效果已成为教育工作者和研究者关注的焦点。随着教育数据挖掘(Education Data Mining, EDM)技术的兴起,教育者得以利用先进的数据分析方法深入探讨教学过程、学生行为与学习结果之间的关系。EDM过程包括收集数据、应用模型进行分析,以及通过模式识别和挖掘与教育机构、学生、教学方法等相关的有用信息^[1]。这些信息不仅能够为高等教育主管部门提供决策支持,帮助改进政策,还为教育机构提供了监督和优化教师与学生的互动的依据,学生则可以利用这些信息识别学习中的薄弱环节,以便更有针对性地提升个人学业表现^[2]。中共中央、国务院印发《教育强国建设规划纲要(2024—2035)》中指出:“建立基于大数据和人工智能支持的教育评价和科学决策制度”^[3],这为教育数据挖掘技术在教育领域的应用和发展提供了政策指引和有力支持。

随着教育数据量的激增,如何有效提取有价值的信息、分析学生的学习过程和预测学业成绩,成为教育研究中亟待解决的问题。教育数据中包含了大量关于学生学习行为、兴趣、成绩、参与度等方面的信息,这些数据为教育者提供了改进教学策略、提升教学质量和学生学习效果的宝贵依据^[4]。因此,教育数据挖掘的应用领域越来越广泛,尤其在高等教育领域,为教学决策和资源配置提供了更加科学的数据支持^[5]。

在传统的教学评估中,单一的评估方法往往局限于学生某一方面的学习成果,而忽略了学生综合能力的培养。为了更加全面地评估学生的学业表现,本研究引入了多维项目反应理论(Multi-dimensional Item Response Theory, MIRT)^[6]。MIRT通过为测验中所涉及的每个维度引入能力和项目区分度参数,模拟测验题目和学生之间的交互作用^[7]。作为一种非线性模型,MIRT可以表征题目参数、学生的多维潜在能力以及其正确作答概率之间的关系,通过此模型,可以对学生在

多个维度上的能力进行独立推断^[8]。本研究使用的大学物理期末考试数据涉及的试题综合性较强,一道题目往往包含多种能力要求。因此,考虑到学生在作答过程中可能的猜测因素,本研究采用了目前广泛应用的三参数多维 Logistic 模型(M3PL)^[6],用于试题参数估计和学生多维能力估计。在 MIRT 分析中,能力维度的确定是研究的关键。首先需要考虑试卷是否涉及了多个维度的能力考察,其次则要依据数据与 MIRT 模型的拟合度进行选择。研究表明,如果数据与 MIRT 模型拟合度高,就应该采用 MIRT 分析^[9]。三维模型的拟合指标都是最低,说明三维模型与数据的拟合程度最好^[10]。

本研究采用 MIRT 模型对 2022—2023 学年下学期、2023—2024 学年上学期、2023—2024 学年下学期的多个学期期末考试数据进行分析,探讨学生在记忆、概念理解和综合应用等多个能力维度上的表现。具体而言,本研究主要包括以下几个方面:首先,对试卷质量进行分析;其次,探讨过程性评价与终结性评价之间的关系,评估其对教学效果的促进作用;再次,分析学生在记忆、概念理解和综合应用三个维度上的能力表现,通过数据可视化和聚类分析,识别教学过程中的异常情况并提出针对性的改进措施;最后,基于分析结果,为教育管理者提供数据支持,推动教学质量的持续改进。

通过多维度数据的综合分析,本研究旨在为教育实践提供可操作的建议,帮助教师深入理解学生的学习需求,优化教学策略,从而促进学生的全面发展,并为教育管理者提供决策支持,提升整体教学效果。

1 数据来源、数据处理及试卷质量分析

1.1 数据来源

本研究对学生在“大学物理”期末考试中产生的数据进行收集和挖掘。数据包括 2022—2023

学年下学期、2023—2024 学年上学期、2023—2024 学年下学期三个学期的工科物理期末考试答题记录以及与之相关的其他数据信息,相关数据信息见表 1。在数据收集过程中可能会出现缺失数据或异常数据,需要对数据进行预处理。在预处理数据过程中,清除不符合格式的异常数据或将异常数据当作缺失数据。表中为清除缺失数据后的人数。

表 1 试卷数据集

时间	2022— 2023(下)	2023— 2024(上)	2023— 2024(下)
样本数	2314	2244	2241
题目标记	Q1~Q23	Q1~Q24	Q1~Q24

1.2 数据的信度和效度分析

根据 SPSS 统计软件分析,三个学期试卷的信度分析见表 2,效度分析见表 3。三次考试数据的整体信度 Alpha 系数>0.8,结构效度 KMO>0.9,巴特利特球形度检验的显著性<0.05,表示期末考试结果具有较高的信度和准确性。

表 2 试卷信度分析

时间	可靠性统计		项数
	克隆巴赫 Alpha	基于标准化项的克隆巴赫 Alpha	
2022—2023 (下)	0.835	0.839	23
2023—2024 (上)	0.852	0.855	24
2023—2024 (下)	0.84	0.844	24

1.3 试卷质量分析

Winsteps^[11]是一款用于项目反映理论(Item Response Theory, IRT)的分析软件,广泛应用于

心理测量和教育评估领域。该软件能够生成 Wright Map(怀特图,亦称项目特征曲线图),以直观展示 IRT 分析结果。本研究采用 Winsteps 对收集的数据进行预处理,并据此绘制怀特图(见图 1)。怀特图的纵轴表示学生的能力水平,横轴表示题目的难度参数。综合分析怀特图,可以发现测试题目能够覆盖不同能力水平的学生群体,且题目的难度分布合理,这表明所设计的测试具有较高的整体质量。

1.4 MIRT 模型数据处理

本研究采用了 MIRT 模型来分析期末考试试卷。试卷题型包括选择题、填空题和计算题。为了适应 MIRT 模型的分析要求,我们对学生的答题数据使用二级计分制进行处理,即将每个项目上的作答记录简化为“对”与“错”两种结果,分别用数字“1”和“0”表示。对于选择题,其答题记录直接用“1”和“0”来表示答题的“正确”和“错误”。对于填空题,根据题目分值进行了如下编码:3 分的题目用“1”和“0”表示答题分数分别为“≥2 分”和“≤1 分”;4 分的题目用“1”和“0”表示答题分数分别为“≥3 分”和“≤2 分”;5 分的题目则用“1”和“0”表示答题分数分别为“≥3 分”和“≤2 分”。对于计算题,每题总分为 10 分,同样用“1”和“0”来表示答题分数分别为“≥5 分”和“≤4 分”。

通过这种数据处理方式,我们能够将复杂的得分数据转换为适合 MIRT 模型分析的二元数据,从而更准确地评估题目性能和学生能力。这种数据处理方法不仅提高了分析的准确性,而且确保了结果的可靠性和有效性。

2 研究内容

本文采用 MIRT 对 2022—2023 学年下学期至 2023—2024 学年上学期三次期末考试的成绩进行深入分析。通过这一分析,我们能够获取每

表 3 试卷效度分析

KMO 和巴特利特检验		时间		
		2022—2023(下)	2023—2024(上)	2023—2024(下)
KMO 取样适切性量数		0.943	0.949	0.934
近似卡方		9220.194	9881.094	9860.794
巴特利特球形度检验	自由度	253	276	276
	显著性	0.000	0.000	0.000

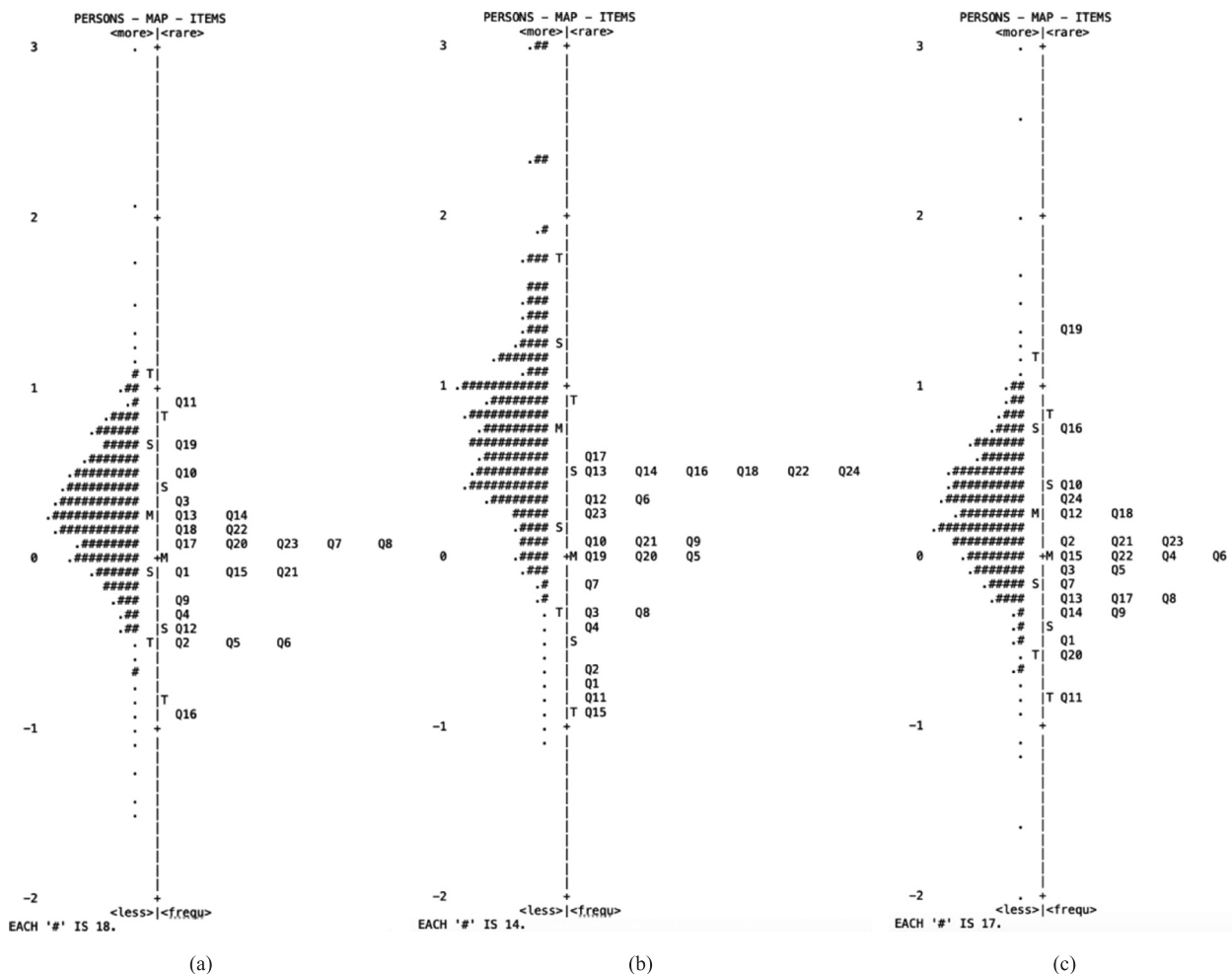


图1 试卷怀特图

(a) 2022—2023(下); (b) 2023—2024(上); (c) 2023—2024(下)

个学生在记忆、概念理解和综合应用三个维度上的能力值。进一步地,通过对每位老师所授教学班级学生的能力值求平均,能够计算出每个班级在这三个维度上的平均能力值。据此,我们可以从记忆、概念理解和综合应用这三个维度的平均值来评估每个班级的整体表现。

记忆维度旨在评估学生对知识点的基础记忆和重复能力。若学生在此维度得分较低,表明其基础知识掌握不牢固,教师可以考虑增加重复练习,以加强学生的记忆能力。概念理解维度用于考查学生对核心概念的理解及其在新情景中的应用能力。若得分较低,则意味着学生对核心概念的理解不足,教师需加强对概念的解释和讨论,以促进学生的深入理解。综合应用维度评估学生在复杂问题中应用知识解决问题的能力。若在此维度得分偏低,表明学生在知识的实际应用上存在

问题。教师可以通过更多情景问题和案例教学来帮助学生提升应用能力。

通过对学生在三个维度上的整体表现进行分析,教师可以评估教学内容在各个维度上的有效性,并据此调整教学内容。如果在记忆维度普遍得分较高,但在理解和应用上得分较低,这表明教学可能过于侧重于基础知识的记忆,而缺乏对概念的深入理解和应用能力的培养。在这种情况下,教师应调整教学策略,加强概念讲解和应用训练,以促进学生的全面发展。

2.1 过程性评价与终结性评价的相关性

本研究通过分析学生在2023—2024学年上学期的平时课堂测试、课后作业与期末考试三个维度(记忆、概念理解和综合应用)之间的相关性数据(见表4),我们可以深入理解不同评估方式与期末考试各维度成绩之间的关系。分析结果表

表 4 某教师 2023—2024 学年上学期两个教学班平时课堂测试、课后作业与期末考试三个维度的相关性

		平时数据		期末考试数据		
		课后作业	课堂测试	记忆	概念理解	综合应用
平时数据	课后作业	1				
	课堂测试	0.540**	1			
期末考试数据	记忆	0.381**	0.258**	1		
	概念理解	0.307**	0.273**	0.158**	1	
	综合应用	0.448**	0.306**	0.344**	0.258**	1

** . 在 0.01 级别(双尾),相关性显著。个案数=306

明,课后作业与课堂测试成绩之间、课后作业成绩与期末考试各维度成绩之间、课堂测试成绩与期末考试各维度成绩之间、期末考试各维度之间均存在显著的正相关性。特别是,课后作业与期末考试综合应用维度的相关性最高,而课堂测试成绩与期末考试各维度成绩之间的相关性略低于课后作业。

课后作业与课堂测试:两者之间的显著相关性反映了学生在课堂参与度和课后复习效果上的一致性。这表明,学生在课堂上的学习投入与课后复习和巩固练习在一定程度上是相辅相成的。

过程性评价与终结性评价:相关性结果表明,学生的学习是一个连续的过程,早期的学习活动(如课后作业和课堂测试)与后期的评估(如期末考试)之间存在密切的联系。这种连续性揭示了过程性评价在学生在学习过程中不可忽视的价值。

课堂测试、课后作业与期末考试:期末考试与课后作业的相关度明显高于与课堂测试的相关性。这种高相关性主要源于课后作业和期末考试在内容深度、知识覆盖范围和题目设计目标上的高度一致性。此外,学生在完成作业时表现出的自主学习行为和时间投入,也进一步增强了课后作业与期末考试之间的关联性。相比之下,课堂测试由于侧重即时反馈,题目范围较窄且时间有限,其与期末考试的相关性相对较低。

课后作业与考试的综合应用维度:课后作业与考试中综合应用维度的相关性最高。这表明课后作业在促进学生综合应用能力方面起到了关键作用。这一现象可能与课后作业的设计中融入了应用型问题和实践任务有关,这些设计不仅帮助学生巩固知识,还加强了知识的灵活应用能力。

期末考试各维度之间的相关性:期末考试与

各能力维度之间的显著正相关性表明,这些维度在一定程度上存在相互关联性。这暗示了学生在某一维度上的能力可能对其他维度的表现产生积极影响,进一步强调了综合性教学的必要性。

这些发现为教师和教学管理者提供了重要的指导意义。首先,**评估工具的重要性:**课后作业和课堂测试在评估学生学习成果方面具有重要价值。教师应充分利用这些工具,不仅用于检测学生的学习进展,还能有效促进其学习效果。其次,**有助于优化教学策略:**结果表明,加强课后作业的设计和课堂互动可以显著提高学生在期末考试中的表现,尤其是在综合应用能力的提升方面。最后,**应注重综合应用能力的培养:**期末考试的各维度之间的正相关性表明,综合应用能力的培养对于学生全面发展具有关键作用,通过引入更多情景问题、案例教学以及综合训练,教师可以更有效地提升学生在多个维度上的能力表现。

通过实证分析揭示过程性评价与终结性评价之间的正向关联,为教学实践提供了基于数据的改进方向。这不仅为教师优化教学方法提供了科学依据,也为教育管理者制定教学改革策略和资源分配方案提供了数据支持,从而进一步推动教育质量的持续提升。

2.2 通过终结性评价,预警教学异常情况

图 2 展示了三个学期学生在记忆、概念理解和综合应用三个维度上的目标达成度箱线图。通过对比分析可以发现,大多数教学班级的表现处于正常范围,但少数班级在某些维度上表现出异常。

例如,在 2023—2024 学年上学期(见图 2(b)),教学班级④在记忆维度上的表现明显偏低,反映出学生对基础知识的掌握不够牢固。针对这一情况,教师可以适当增加基础知识的重复练习和巩

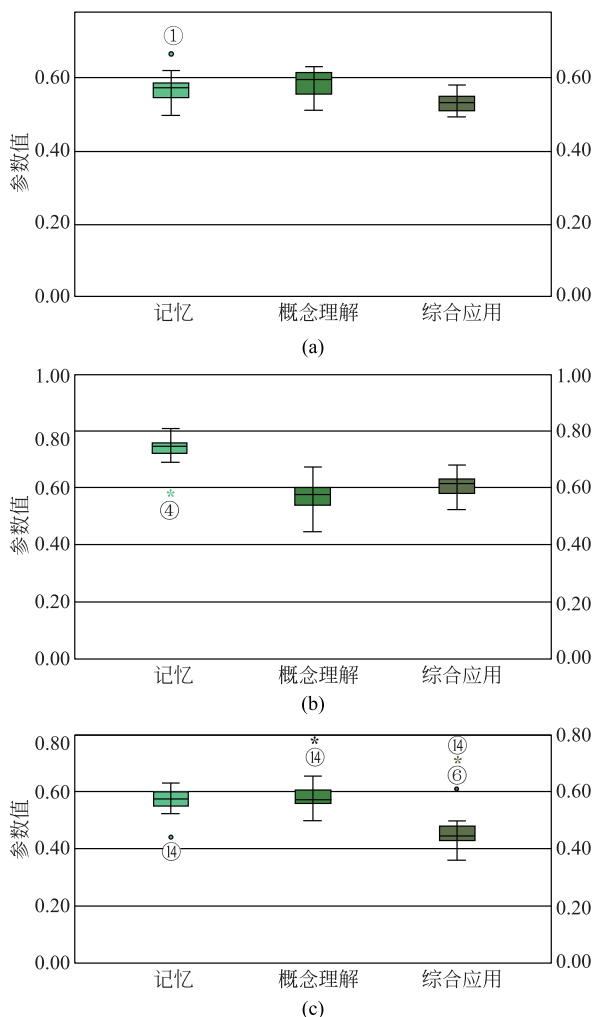


图2 目标达成度

(a) 2022—2023(下)目标达成度箱线图; (b) 2023—2024(上)目标达成度箱线图; (c) 2023—2024(下)目标达成度箱线图

固环节,以增强学生对知识点的记忆和理解。而在 2023—2024 学年下学期(见图 2(c)),教学班级 ⑭在记忆维度上出现异常,但其在概念理解和综合应用维度上表现突出。这表明该教师的教学方法倾向于强调深度学习和核心概念的掌握,并在案例教学和例题分析方面具有显著优势。然而这种教学方式可能忽略了对学生基础记忆能力的系

统性训练,导致学生在记忆维度上的表现较为薄弱。因此,该班级的教学改进方向应包括在现有教学方法的基础上,增加基础知识的记忆性训练,以达到更均衡的教学效果。

箱线图的可视化特性为教师提供了一种直观且高效的工具,用于识别教学过程中出现的异常情况,这种数据可视化不仅帮助教师了解自身教学的优势与不足。还能后续教学调整提供明确的方向。通过箱线图展示的数据,教师可以在未来的教学过程中有针对性地优化教学方法和内容,从而提升整体教学质量和学生的学习效果。此外,这些数据对教育管理者同样具有重要的参考价值。当发现某些教学班级在特定维度上的表现异常时,教育管理者可以加强对相关教师的指导与培训,帮助其调整教学策略以满足学生的学习需求。

这种基于数据驱动的分析方法为教师提供了更加精确的教学反馈,使其能够识别教学中的薄弱环节并采取针对性的改进措施。这不仅有助于教师的自我反思和专业提升,也为教育管理者提供了科学的决策依据,促进了教学质量和学生学习成效的整体提升。通过这一方法,可以进一步推动教育实践的系统性优化,为学生的全面发展奠定坚实基础。

2.3 通过聚类分析反馈教学质量,建立教学档案

本研究采用基于 WCSS(Within-Cluster Sum of Squares)的手肘法来确定 K 均值聚类中的最优簇数。WCSS 是指簇内各点到簇中心的平方距离之和,而手肘法的基本原理是通过观察 WCSS 随着簇数的增加和变化的趋势,找到“肘部”点,即 WCSS 减小速率显著下降的位置。该点对应的簇数被认为是最优簇数。本研究利用手肘法对每个班级的平均能力值进行聚类分析,旨在将教学效果类似的班级划分为不同群组,从而反映不同教师的教学质量。

图 3 展示了 2023—2024 学年上学期三个维

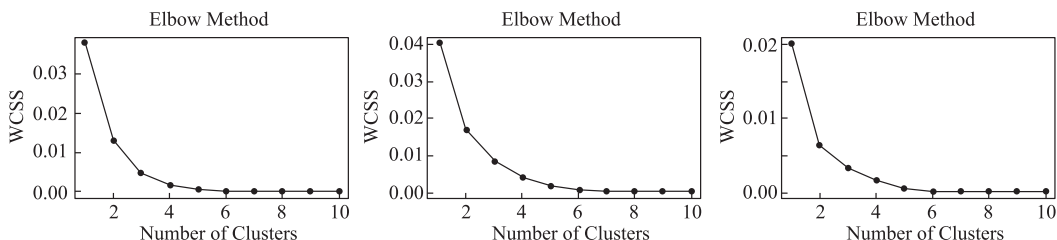


图3 2023—2024(上)三个维度手肘法结果图

度手肘法确定最佳聚类数的结果,分析表明,每个维度最佳聚类数为 3。对于 2022—2023 学年下学期和 2023—2024 学年下学期的数据分析结果,与 2023—2024 学年上学期的结果一致(图略)。

根据聚类分析结果,将平均能力值按照记忆、概念理解和综合应用三个维度划分为三类(见表 5):1 表示教学效果优秀、2 表示教学效果良好、3 表示教学需要改进。在此基础上,本研究对三个学期的教学班级进行教学质量分类,并建立了教师教学档案(见表 6)。通过对多个学期教学效果分析,班级可以进一步分为以下三类:

一是**表现优秀的班级**:三个维度均表现优秀的班级可以视为教学效果的标杆。通过记录教师的教学方法、课程设计、学生反馈以及成绩分布等,这些优秀实践可以作为成功案例,为其他教师提供教学优化的参考与借鉴。

二是**需要重点关注的班级**:在三个维度上均表现欠佳的班级需要特别督导。此类教学班级可

能面临多方面的问题,如学生基础薄弱、课堂互动不足或教学资源分配不合理。应重点关注教师的教学策略、课程内容以及学生的学习状态,提出针对性改进措施,以提升教学效果。

三是**教学设计某些方面存在问题的班级**:部分班级在某些维度上表现突出,但在另一些维度上表现较弱,这可能反映了教学设计中的结构性问题:如教学内容过浅,教学内容缺乏深度,难以满足学生的认知需求。建议增加教学内容的深度和广度,确保教学内容具有挑战性但仍在学生的理解范围内。通过引入实例、模拟和实验,帮助学生理解抽象概念;定期进行概念测试,评估学生的理解程度并相应调整教学计划。或者教学内容过深。过于复杂的教学内容可能超出学生的认知能力,影响其学习效果。建议重新审视教学大纲,去除或者简化过于复杂的部分,并通过螺旋式教学法逐步引导学生深入理解复杂概念。同时,选择更具代表性和相关性的案例,帮助学生掌握核心

表 5 三个不同学期在三个维度的聚类结果

教师代码	2022—2023(下)			2023—2024(上)			2023—2024(下)		
	记忆	概念理解	综合应用	记忆	概念理解	综合应用	记忆	概念理解	综合应用
①	1	2	1	2	2	2	1	3	3
②	3	2	3	\	\	\	2	3	3
③	\	\	\	1	2	2	1	3	3
④	\	\	\	3	3	3	2	3	3
⑤	\	\	\	1	1	3	1	2	3
⑥	2	1	2	1	1	1	1	2	1
⑦	\	\	\	\	\	\	2	3	3
⑧	2	1	1	1	1	2	2	3	2
⑨	2	1	3	2	2	2	1	2	3
⑩	3	2	3	1	2	2	2	3	2
⑪	\	\	\	\	\	\	2	3	2
⑫	2	2	3	1	2	2	2	2	3
⑬	2	2	2	1	2	2	1	2	3
⑭	2	3	3	1	1	3	3	1	1
⑮	3	1	2	2	2	3	1	3	2
⑯	\	\	\	1	2	2	\	\	\
⑰	2	1	1	\	\	\	\	\	\
⑱	1	3	2	\	\	\	\	\	\
⑲	3	1	3	1	1	2	\	\	\
⑳	2	1	2	\	\	\	\	\	\

表 6 教师教学档案

教师 代码	表现欠佳的次数			表现好的次数			总教学 次数
	记忆	理解	应用	记忆	理解	应用	
①		1	1	2		1	3
②	1	1	2				2
③		1	1	2			2
④	1	2	2				2
⑤			2	2	1		2
⑥				2	1	2	3
⑦		1	1				1
⑧		1		2	2	1	3
⑨			2	1	1		3
⑩	1	1	1	1			3
⑪		1					1
⑫			2	1			3
⑬			1		1		2
⑭			1	2			3
⑮	1	1	2	1	2	1	3
⑯	1	1	1	1	1		3
⑰				1			1
⑱					1	1	1
⑲		1		1			1
⑳			1	1	2		2
㉑					1		1

内容。针对这些情况,可以给学生提供个性化支持。针对学生的能力差异,提供不同难度的学习资源和任务,设计适合不同水平学生的个性化学习方案。

2.4 分析结果为教学管理者提供导向支持

通过对教学效果数据的分析,教学管理者能够制定更加科学、合理的决策,这对于推动教学改革和促进教育发展具有重要意义。

首先,教学效果分析可以帮助管理者识别教学中的薄弱环节,为制定针对性的改革策略提供依据。例如,管理者可以基于分析结果改进课程内容、优化教学方法或调整评估体系。同时,通过深入了解不同教学资源的有效性,管理者能够更科学地分配预算和教学资源,确保教育投入产生最大化的收益。

其次,分析结果可以为教师培训和专业发展提供有力支持。通过数据驱动的反馈机制,教师能够清晰地认识到自身在教学中的优势与不足,

从而在特定领域进行有针对性的提升。例如,针对数据中反映出的学生学习效果较弱的领域,教师可以通过培训掌握新的教学方法和工具,进而提升教学技能。

最后,定期开展教学效果分析有助于在教育机构内建立一种持续改进的文化。这种文化不仅激励教师不断优化教学质量,还能激发教育创新,推动教师的专业成长。在这种数据支持下,管理者可以制定长效机制,激励教师持续追求卓越,为学生提供更高质量的教育服务。

因此,教学效果分析结果为教学管理者提供了基于实证数据的宝贵导向支持,使其能够以科学决策为基础,有效推动教育实践的改进和发展。这一过程不仅有助于提升整体教学质量,还为教育创新和教师发展创造了更多可能性,从而实现教育体系的可持续提升和优化。

3 讨论

本研究基于多维项目反应理论(MIRT)和教育数据挖掘技术,对多个学期的期末考试数据进行了深入分析。研究的主要目标是评估学生在不同教学维度的表现,并通过数据分析揭示教学中的薄弱环节,从而为教师和教育管理者提供改进教学质量的依据。

教学效果评估 研究表明,过程性评价(如课堂测试、课后作业)与终结性评价(期末考试)之间存在限制的正相关性。这表明学生在课堂中的学习投入和课后复习巩固对于期末考试的成绩起到了重要作用,尤其是在综合应用能力的提升上。此外,通过对期末考试不同维度(记忆、概念理解、综合应用)之间的相关性分析,进一步证实了学生在某一维度的能力对其他维度的表现具有积极影响。教师可以通过强化课后作业和课堂测试的设计,有效促进学生在期末考试中的表现,尤其是在综合应用能力方面。

教学异常预警 通过对终结性评价数据的分析,研究能够识别在特定维度上表现异常的教学班级。箱线图的可视化展示使得教师能够直观地识别出教学中的不足,及时发现哪些班级在记忆、概念理解和综合应用等维度上存在问题。例如,某些班级在记忆维度上表现不佳,可能表明基础知识的掌握不足,教师可以在教学中增加更多的

基础练习和复习。该预警系统不仅帮助教师进行自我调整,也为教育管理者提供了支持,确保及时干预并提供必要的指导。

教学质量反馈 采用聚类分析,将不同班级的教学效果分为三个类别:优秀、良好和待改进。通过将每个班级的平均能力值进行聚类,能够为教师和教育管理者提供更精准的教学档案,并为后续的教学改进措施提供科学依据。对于教学效果不佳的班级,可以加强对教师的指导和培训,调整教学方法,以提升整体教学质量。反之,对于表现优秀的班级,可以总结并推广其有效的教学实践。

教学管理者支持 分析结果为教学管理者提供了重要的决策支持,帮助其识别教学中的强项和弱项。基于数据的分析使得教学资源能够更加合理地分配,并优化教学策略。此外,定期的教学效果分析为管理者提供了持续改进教学质量的依据,促进了教育创新和教师专业成长。

4 结语

本研究通过对教学数据的深入分析,揭示了过程性评价与终结性评价之间的关系,识别了教学中的薄弱环节,并提供了改进措施。通过多维项目反应理论与教育数据挖掘的结合,教学管理者能够更有效地进行数据驱动的决策,优化教学方法和资源分配。最终,这些分析结果不仅能够帮助教师提升教学质量,还能为教育管理者提供制定更具针对性的改革策略的依据,从而促进教育质量的持续提升。

参 考 文 献

- [1] BATOOL S, RASHID J, NISAR M W, et al. Educational data mining to predict students' academic performance: A survey study[J]. *Education and Information Technologies*, 2022, 28: 905-971.
- [2] BURGOS C, CAMPANARIO M L, de la PEÑA D, et al. Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout[J]. *Computers & Electrical Engineering*, 2018, 66: 541-556.
- [3] 中共中央、国务院印发《教育强国建设规划纲要(2024—2035年)》(2025年1月19日), http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202501/t20250119_1176193.html.
- [4] KABAKCHIEVA D. Predicting student performance by using data mining methods for classification[J]. *Cybernetics and Information Technologies*, 2013, 13(1): 61-72.
- [5] FERNANDES E, HOLANDA M, VICTORINO M, et al. Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil[J]. *Journal of Business Research*, 2019, 94: 335-343.
- [6] RECKASE M D. *Multidimensional item response theory* [M]. Springer-Verlag New York, 2009.
- [7] RECKASE M D. The past and future of multidimensional item response theory[J]. *Applied Psychological Measurement*, 1997, 21: 25-36.
- [8] 康春花,辛涛. 测验理论的新发展:多维项目反应理论[J]. *心理科学进展*, 2010, 18(3): 530-536.
- [9] KAND C H, XIN T. New development in test theory: Multidimensional item response theory[J]. *Advances Psychological Science*, 2010, 18(3): 530-536. (in Chinese)
- [10] 许志勇,丁树良,钟君. 高考数学试卷多维项目反应理论的分析及应用[J]. *心理学探新*, 2013, 33(5): 438-443.
- [11] XU Z Y, DING S L, ZHONG J. The analysis and application of MIRT in mathematics paper in college entrance examination[J]. *Psychological Exploration*, 2013, 33(5): 438-443. (in Chinese)
- [12] 侯虎. 可解释人工智能在物理试卷分析中的应用[D]. 上海:同济大学,2023.
- [13] HOU H. Application of explainable artificial intelligence in the analysis of physics test papers[D]. Shanghai: Tongji University, 2023. (in Chinese)
- [14] LINACRE J M. WINSTEPS manual for rasch analysis [M]. winsteps.com, 1981.
- [15] 袁庆新,周永卫. 多狗追问题的力学解法[J]. *力学与实践*, 2010, 32(1): 81-82.
- [16] YUAN Q X, ZHOU Y W. A mechanics method of the dog-and-dog pursuit problem[J]. *Mechanics in Engineering*, 2010, 32(1): 81-82. (in Chinese)
- [17] 赵振国. 追击问题更直观的求解[J]. *力学与实践*, 2000, 22(5): 68.
- [18] ZHAO Z G. A more intuitive method to the pursuit problem[J]. *Mechanics in Engineering*, 2000, 22(5): 68. (in Chinese)
- [19] 马秀艳,于凤军. 多狗追问题的解析解[J]. *河南师范大学学报(自然科学版)*, 2001, 29(4): 109-111.
- [20] MA X Y, YU F J. An analytic solution of a kinematical game[J]. *Journal of Henan Normal University (Natural Science Edition)*, 2001, 29(4): 109-111. (in Chinese)
- [21] 李力. N狗追问题的复分析方法[J]. *大学物理*, 2015, 34(8): 19-19.
- [22] LI L. An ingenious complex analysis solution of N dogs running problem[J]. *College Physics*, 2015, 34(8): 19-19. (in Chinese)

(上接第45页)