

基于学生综合能力提升的数智化课程评价实践研究

郭世琚¹, 邹韵²

(1.广州职业技术大学智慧汽车学院,广东广州 511483;

2.广州职业技术大学财经学院,广东广州 511483)

[摘要]课程评价改革具备较强的操作性特征,通过实验法进行应用和效果评估,以验证创新评价模型的有效性,弥补课程评价实践类研究的不足。实验针对不同类型的课程设计评价内容,构造基于学生综合能力提升的课程评价指标体系,采用层次分析法对指标进行赋权,并利用多种数智化教学工具记录过程数据,辅助评价分析,获得更为全面、精准和个性化的评价结论。根据实验结果的对比分析,改良后的评价方式在等级分布上更为均衡,成绩评定更加合理,能有效反映学生的学习质量和优缺点。在此基础上,进一步提出 AHP 赋权调整、优化人机协同机制、建立评价循环改进制度等对策建议,体现以课程评价促进学生全面发展的育人理念。

[关键词]课程评价;评价指标体系;综合能力提升;智能化评价;层次分析法

[作者简介]郭世琚(1988—),女,安徽淮南人,广州职业技术大学智慧汽车学院助理研究员,硕士,研究方向:学生教学管理、高等职业教育管理。邹韵(1988—),女,湖北武汉人,广州职业技术大学财经学院讲师,硕士,研究方向:金融、保险、高等职业教育学。

[基金项目]本文系 2023 年广东省高等职业教育教学改革研究与实践项目“基于学生综合素质评价改革的财经类课程评价体系构建与应用研究”(项目编号:2023JG606)。

[DOI] <https://doi.org/10.62662/kjxk0103007>

[中图分类号] G434

[本刊网址] www.oacj.net

[投稿邮箱] kjxk999@163.com

随着教育评价改革的深入,关于高职课程评价的研究也日渐丰富。纵观目前的研究成果,理论性探讨占据比重较大,而应用型研究相对缺乏。课程评价是人才培养过程中的重要环节,如何将理论方案进行落实,往往需要多轮实践验证和循环调整。评价改革不是纸上谈兵,而是要体现在日常的教育活动中,发挥其应用价值。因此,本文将创新课程评价的设计与实施,并基于研究团队的评价实验,展开案例分析研究,以期为高职课程评价改革提供实战经验借鉴。

一、课程评价的方案设计

(一)课程评价的对象与内容构成

本次评价实验我们选择了本校财经专业大二两个班级的学生共 70 人,他们已经拥有一定的专业基础知识储备,对学校的教学管理制度也较为熟悉,经过一年多的学习,对于课程评价形成了初步的认知和思考,能够提出自己的看法和建议。在课程的选取上,为了尽量兼顾全面性和可行性,选择了 4 门课程,其中 2 门是理实一体化课程,即需要同时进行理论学习和实操演练,但演练的方式是根据课程内容和相关岗位业务设计的模拟任务,2 门课程的形式类似,只是总体难度系数有所不同。另外

2 门属于实践类课程,由学校老师和企业导师共同教授,课程目标是让学生完成真实的企业工作任务,一门是营销类项目,另一门是金融科技应用项目。在课程前期由学校老师讲解任务中需要用到的专业知识和技术,课程中后期由企业导师布置具体任务,联合校内老师一起进行过程辅导,直至完成项目交付。这类课程更注重项目成果达标率,理论授课时间较少,考核占比较小。

(二)课程评价的指标:基于学生综合能力提升

不同于以往单纯地通过课程活动进行最终评分的方式,我们通过设计能力评价指标,将活动评分与指标进行分解对应,以各项指标权重作为数理计算根据,最后得出基于学生能力提升的课程评价结论。这样的方法可以有效避免终结性分数(如理论题测试)无法良好反映学生的能动性和态度变化,而主观形式的评分活动(如成果展示),其结论往往过度依赖评价者的主观意识和心理状态。通过能力型评价指标体系的测算,可以在较大程度上改善传统评价的固化思维,使得评价结论更为全面和科学,也更符合促进学生全面发展的教育理念。

根据团队前期的研究成果,设计出财经类课程五个维度的评价指标体系(详见参考文献 2),此处

以学生能力提升为核心,选取对应的评价指标。接下来我们利用层次分析法(AHP)为指标进行赋权,AHP通过将复杂的决策问题分解为不同的层次结构,然后进行两两比较以确定各个因素的相对重要性,最终计算出各备选方案的权重。

首先,我们确定决策目标,即构建基于一级指标A学生综合能力提升的课程评价指标体系的合理权重,准则层为二级指标A1~A3,方案层为所有的三级指标。然后使用1~9标度法来表示相对重要性,数字越大代表重要性越高。通过逻辑判断,我们设定准则层的判断矩阵,如表1所示:

表1 准则层的判断矩阵

二级指标	A1	A2	A3
A1	1	2.5	2.5
A2	0.4	1	0.9
A3	0.4	1.11	1

将矩阵带入MATLAB软件进行一致性检验,计算判断矩阵的最大特征值 λ_{\max} 约为3.001,一致性指标 $CI \approx 0.0005$,随机一致性指标 $RI = 0.58$, $CR \approx 0.0008 < 0.1$,通过检验。随后我们用相同的方法分别对三个方案层进行处理,初始权数矩阵为:

$$A1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/2 & 1 & 2 & 3 & 4 \\ 1/3 & 1/2 & 1 & 2 & 3 \\ 1/4 & 1/3 & 1/2 & 1 & 2 \\ 1/5 & 1/4 & 1/3 & 1/2 & 1 \end{bmatrix},$$

$$A2 = \begin{bmatrix} 1 & 2 & 1/3 & 1/2 & 3 \\ 1/2 & 1 & 1/4 & 1/3 & 2 \\ 3 & 4 & 1 & 2 & 5 \\ 2 & 3 & 1/2 & 1 & 4 \\ 1/3 & 1/2 & 1/5 & 1/4 & 1 \end{bmatrix},$$

$$A3 = \begin{bmatrix} 1 & 2 & 3 & 4 & 1/2 & 3 \\ 1/2 & 1 & 2 & 3 & 1/3 & 2 \\ 1/3 & 1/2 & 1 & 2 & 1/4 & 1 \\ 1/4 & 1/3 & 1/2 & 1 & 1/5 & 1/2 \\ 2 & 3 & 4 & 5 & 1 & 4 \\ 1/3 & 1/2 & 1 & 2 & 1/4 & 1 \end{bmatrix}$$

经过软件分析,二级指标的判断矩阵A1、A2和A3的 λ_{\max} 分别为5.092、5.008和6.21,且CR值均小于0.1,全部符合一致性检验。由此,可以得到各

指标的权重如表2所示:

表2 基于学生综合能力提升的课程评价指标体系

二级指标 (权重)	三级指标	同级权重	全局权重
A1 专业能力 (0.557)	A11 专业基础知识和基本原理	0.418	0.233
	A12 专业实操技能及应用	0.264	0.147
	A13 专业岗位职责和规范	0.16	0.089
	A14 专业新技术应用	0.097	0.054
	A15 相关学科知识技能储备	0.061	0.034
A2 通用能力 (0.213)	A21 沟通与人际交往能力	0.16	0.034
	A22 语言表达能力	0.097	0.021
	A23 自主学习能力	0.418	0.089
	A24 实践与创新能力	0.264	0.056
	A25 数字化技术应用能力	0.061	0.013
A3 综合素质 (0.230)	A31 思想道德修养	0.265	0.061
	A32 职业道德与职业责任感	0.13	0.030
	A33 爱岗敬业与职业认同	0.079	0.018
	A34 健康的身心素质	0.052	0.012
	A35 团队合作意识	0.395	0.091
	A36 抗压和自我调节能力	0.079	0.018

二、课程评价的实施:数智化赋能

(一)课程评价实施方案

为了实现新的课程评价计划,得到更为准确的

实验数据,需要改变以往传统的人工评价方式,更多地运用智能化评价工具,在各项评价环节中获取动态、多源、精准的结果。对于理实一体化课程,评价需要兼顾理论和实操两个方面,以其中一门课程“保险实务”为例,评价活动的构成包括日常的过程性评价和期末的终结性评价。具体而言,过程性评价活动有到课率、个人平时作业(涵盖保险理论与业务实操)、小组作业与展示(考虑成员贡献度)、课堂问答、线上课程资源学习情况,终结性评价活动包括理论测试与实操项目考核。

实践类课程以项目完成情况为考核重点,因此过程性评价的活动相对较少,例如金融科技应用实践课程,除了初期的大数据技术讲解与分析、项目方案展示等内容需要在教室完成,其余的活动场所并不固定,可能在校内实训室,也可能在企业,或者在图书馆,甚至是通过网络进行,因此类似到课率和课堂问答这类评价活动并不适用。另一方面,实践课程的最终项目成果可以从一定程度上反映学生平时的学习情况,因为只有把握好项目的每个步骤,才能展现更优质的成果。但我们也不能忽略项目过程的跟踪,这样才能及时发现问题并进行纠正,排除一些客观的不利因素和意外事件的影响,否则单纯的终结评价难免会出现不公平、不合理的情形。

(二)数智化评价工具赋能

理实一体化课程实施过程中,教师将学生的日常学习活动集中到智能化教学系统中,通过持续追踪学生在系统中的学习数据,实现过程分析与形成性评价。教学系统能够完成签到、在线互动、线上资源学习、布置平时作业、头脑风暴与小组讨论等多项活动的记录,快速评分并形成数据分析结果。同时,利用实操平台进行实训,平台可以记录学生的操作步骤、问题解决策略、结果达成度等情况,生成个性化的学习报告。终期测试也可以通过教学系统和实操平台进行,快速得到考核分数。智能系统不仅仅拥有布置任务和自动评分功能,还可以通过智能大模型算法,分析学生的活动数据,例如练习次数和时长、参与互动的频率、头脑风暴的质量、平时测试的得分增长率等等,通过横向和纵向比较,得到学生“活跃度”“成长值”“熟练度”“严谨度”等多维度的评判,提供动态、深入的学习者画像,将主观评价指标通过客观数据进行展示,让老师更为直观、全面地了解学生的学习习惯和质量。

实践课程的评价实施具有自身的特点,虽然对学生的知识层面考察相对较少,但能通过对学生完成工作任务的过程和结果评估,弥补普通课程在职业能力和通用能力评价方面的不足。另外,实践课程在评价主体上也更为多样化,除了校内教师以

外,企业导师的评价比重更大,对于团队项目还会组织学生进行组内评比和组间互评。通常在工作任务开展之前,会要求学生掌握一定的理论或项目常识,这部分的评价方式和理实一体化课程类似,可以通过智能化教学系统来进行。随后,根据任务要求,学生们需要制定项目方案,以往的评价方式几乎全部依赖老师的主观判断,在本次实验中,我们将学生提交的方案导入 AI 模型进行自动化评价。模型根据我们设置的评分标准,检查规范性、关键要素完整性、创新性、可行性等,作为评分的重要参考。同样,对于项目成果也可以参考 AI 模型的评价结论。此外,针对团队合作项目,我们利用在线文档、工作群等协作平台,追踪小组成员的分工、贡献度、互动质量、任务完成情况等,评估个人在团队中的表现。同时,利用教学平台的匿名投票功能进行学生互评,综合评判个人贡献,得到更为个性化的评价结果。如此一来,即使团队成绩不尽如人意,某些表现突出的学生仍然可以得到不错的分数。

(三)学生综合能力评价指标应用

按照以往的课程评价方式,教师将各项成绩进行统计,然后按照预先设定的比例计算出课程的最终成绩,这样的方式无法有效反映学生综合能力的提升状况。本次实验将课堂活动与学生能力提升的各项评价指标相匹配,得到每个指标的评分,若某一项指标对应多个课堂活动,则采用平均算法获得分数,最后用各项指标的分数与权重的乘积总和计算总得分。

以“保险实务”课程为例,到课率与 A34、A36 相关度较高;个人平时作业、课堂问答、线上课程资源学习情况、期末考核等与 A11~A14、A23~A25、A32~A33 等指标相对应;小组作业与展示还能对标 A21~A22、A35 等。设定好匹配规则后,利用智能分析模型,可以得到每门课程的学生各能力维度的评价情况,以及整体的综合能力评分。

三、课程评价实践结果分析与建议

(一)评价实践结果分析

我们按照传统的成绩计算方式和学生能力评价指标体系分别核算课程成绩并进行统计,通过对比分析可以看到,采用实验方法得到的四门课程的总体平均分从 82.5 提升到 86.3,且同一类型的两门课程,在平均分和成绩段分布上更为类似,说明对学生能力提升的评价具备更好的稳定性。

以营销实践课程为例,图 1 反映了两种不同方式计算的课程评价结果,相较于传统方法,实验结果显示不及格的学生数量有所减少,而优秀的数量增加了 10 个百分点。原本中等分数的学生占据绝对比重,明显高于其他类别,而实验方式得出的级别分布更均衡,中等与良好占比相差不大,总体而

言更符合正态分布。

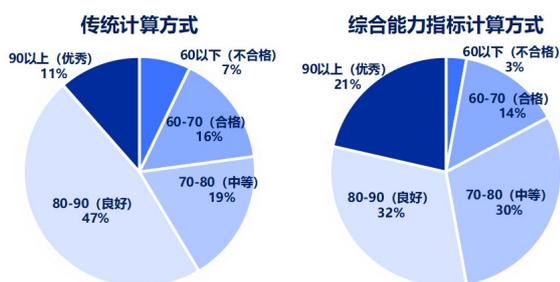


图1 两种计算方式下的课程成绩分布

出现上述变化的原因在于,采用传统计算方法,实践课程的项目成果评价占比非常高,一旦有学生的销售任务没有达标,就很可能出现低分或不及格的情况。但按照综合能力评分标准,A2和A3项下的指标占比近45%,学生在完成项目过程中体现出的个人素质能够拉高这部分的评分,使得综合成绩超过合格线。同样,经过调整评价方式,也涌现出更多优秀的学生,对班级形成正向激励。另一方面,对于小组项目,按照传统方式,组内成员的成绩差异不会太大。对于同一班级,除非有较为突出的表现或明显的不恰当行为,否则老师会进行整体把握,组间的分数差别一般也不会超过10~15分,这就是大部分学生集中在“良好”等级的原因。但是综合能力评价方法将原本集中的评分项目按能力类别进行了细分,淡化了统筹评分的效果,使得成绩分布更为均衡,反映出学生的个性化差异。

此外,我们还将每个学生的三维能力评价结果制作成雷达图,从而更直观地看到学生的优势和薄弱之处,再通过开展对应的教学活动和任务,适配个体差异,让学生扬长避短,进行针对性的训练。图2是某位同学在营销实践课程中A2的评价雷达图,根据测评结果,该学生在A2能力类别中平均得分78.6,其中最突出的是A23自主学习能力,表明学生在课外时间能发挥主观能动性,自觉学习相关知识,夯实基础。但在A21和A22方面评分较低,语言表达能力和沟通技能有待加强,需要在以后的学习中,更多分担小组任务,加强团队合作,通过发言、展演、陌拜、推介等方式克服恐惧心理,训练表达技巧。

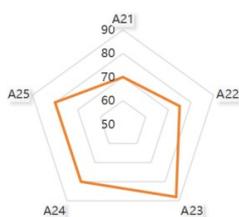


图2 学生A2能力评价雷达图(示例)

评价实验结束后,我们对参与评价的校内教师、企业导师和学生进行问卷调查。多数人认为按照综合能力评价指标进行课程评价的方式更为合理,更能真实反映学习质量、找准学生薄弱之处,对此表述赞同的比例分别为74%、68%和77%,评价改革结果的满意度较高。

(二) 对策建议与改进措施

通过评价改革实践,我们将课程评价从专注教学活动本身转变为以能力提升程度和学生个人价值塑造为出发点,使得评价过程和结果都更有说服力,也能为后续的个性化人才培养提供良好的参考。但在实验过程中我们也发现一些不足之处有待改进。

1. 评价指标体系 AHP 赋权调整

评价指标的权重设定是影响评价结果的关键因素之一,AHP的赋值方法存在主观性强、静态化的固有问题。在确定初始权数时,主要依赖专家或过往经验的判断,容易受到认知局限的影响。针对这一缺陷,可以通过结合主观判断与客观数据的混合赋权法来降低主观偏差。在确定初始权重后,引入客观数据,如学生就业数据、获奖情况等,对权重进行修正,适当提高相关性高的指标的权重,平衡主观逻辑与客观实际。另一方面,目前指标体系权重是固定不变的,但在不同的年级阶段,对于不同性质的课程,应进行适当的调整。例如低年级侧重理论知识和思想塑造,高年级更重视实操技能和岗位素养;理论课与实践课、必修课与选修课的教学目标和评价重点也会有所差异。后期可以通过建立权重动态调整模型,将授课阶段、课程类型、外部需求等因素作为影响因子,设定每个因子的影响系数,在这些因子发生变化时,通过系数来调整指标权重,实现动态匹配。

2. 优化人机协同机制

采用数智化评价工具可以切实提高效率,支持个性化学习,但在实际运用过程中也常常出现一些矛盾,主要集中在两个方面,一是智能化“过多”的问题,二是智能化“过少”的问题。“过多”是指过度依赖技术而降低了评价主体的自主性,尤其是对于一些有较强人文属性的指标,如A31思想道德修养、A34健康的身心素质等,需要教师的主观洞察和长期观察来评判。另外,学生在掌握智能平台的评分标准后,可以进行“策略化”破解,如平台根据学生前后期测验分数的差值来判断成长值,某些学生则利用此规律故意拉大分差以获得进步奖励分。而“过少”是指目前的智能化教学系统尚未达到高级智能的标准,无法完全满足教学和评价活动的全部需求,有些操作比较繁琐,教师和学生也没有养成随时登录和记录的习惯,反而降低了效率,使得评

价结论失真。因此,我们在使用数智化评价工具时,不能忽视了人为的干预,需建立人机分工协同机制。技术主导可量化、高频次的指标,人工主导人文性、低频次的指标,并进行协同校准,对于技术输出的疑似异常结果,由教师复核并修正。教学平台开发方需不断完善系统,优化操作体验,贴合教学实际需求。针对学生的破解策略,采用更为先进的智能技术分析答题时间分布、修改痕迹等行为数据来加以识别。

3. 建立评价的循环改进制度

课程评价是为了衡量学生的学习质量,但更重要的是了解学生的学习状况以便促进后续的能力培养,因此评价不是终点,而是“支点”。采用综合能力评价指标来进行测评也正是基于这样的教育理念,在将来的教学实践中,可以将能力评价节点提前,即进行期中考试,然后根据评价结果及时做出调整,形成“评价—改进—再评价”的循环模式。对于评价指标权重和课堂评价活动,由课程负责人和授课老师牵头做出调整;对于异常评价数据,可以通过与学生单独谈话、小组访谈等方法,找出问题原因并进行修正。另外,结合学生的评价雷达图,运用智能化手段辅助分析,适当调整学习策略,

为后续能力提升找准方向。完成改进后,在后半学期通过新一轮评价验证效果,增强评价的有效性,提高师生对评价结果的信任度,让课程评价成为连接教学实施与学生成长的纽带,切实发挥促进学生综合能力全面提升的重要作用。

参考文献:

- [1] 杨传明.应用型高校经管类专业实训实践校企耦合评价体系研究[J].职教论坛,2015(3):85-88.
- [2] 邹韵.“两型五维”高职人才培养视角下财经类课程评价指标设计[J].教育科学文献,2025(2):65-71.
- [3] 牟蕾.高等院校创新实践教育质量关键影响因素研究[D].西安:西北工业大学,2015.
- [4] 苏哲斌.课堂教学质量的过程性综合评价方法[J].教育探索,2012(9):89-90.
- [5] 丁念金.学习过程评价的基本构架[J].教育测量与评价(理论版),2012(6):29-32,53.
- [6] 马文娟,马文艺.AIGC背景下高职院校电商专业数智化转型路径研究与实践[J].湖北开放职业学院学报,2025,38(7):156-158.
- [7] 王茜雯,吴全全,闫智勇.职业行动建构导向的学习质量评价模型的学理与探思[J].中国职业技术教育,2021(34):57-64.

Practical Research on the Digital and Intelligent Curriculum Evaluation Based on the Improvement of Students' Comprehensive Ability

GUO Shi-jun¹, ZOU Yun²

- (1. Smart Automotive School, Guangzhou Polytechnic University, Guangzhou Guangdong 511483;
2. Finance School, Guangzhou Polytechnic University, Guangzhou Guangdong 511483, China)

Abstract: Curriculum evaluation reform demonstrates strong operational characteristics. This paper evaluates the application and effect by experimental method, so as to verify the effectiveness of innovative evaluation model and make up for the deficiency of practical research on curriculum evaluation. The experiment designed evaluation content for different course types, establishing a comprehensive student competency enhancement evaluation system. The Analytic Hierarchy Process (AHP) was employed to assign weights to these indicators and various digital and intelligent teaching tools are used to record process data and assist evaluation analysis, so as to obtain more comprehensive, accurate and personalized evaluation conclusions. Through the comparative analysis of experimental results, the new evaluation method achieved more balanced score distributions, reasonable grade assessments, and effective reflection on students' learning quality and strengths or weaknesses. Building on these findings, some suggestions were put forward, including AHP weighting adjustment, optimization of man-machine cooperation mechanism and establishment of evaluation cycle improvement system. These measures reflected the educational philosophy of promoting students' comprehensive development through curriculum evaluation.

Key words: course evaluation; evaluation index system; comprehensive ability improvement; intelligent evaluation; Analytic Hierarchy Process