

大学生创新能力提升与OBE课程测量工具的Rasch分析

蒋诗 石建飞 文春玲

黑龙江八一农垦大学

摘要: 大学生创新能力培养是高等教育的主要目标, 将促进大学生创新能力提升纳入OBE课程教学的预期成果并以宏观经济学课程为试点进行OBE课程改革初探。设定5点正填充评定量表, 运用Rasch模型测量OBE课程的教学效果与大学生创新能力培养目标在测量工具中的有效性。Rasch分析表明, 测量工具总体上是可靠并有效的, 项目测量的内容属于同一潜在维度, 且选项类别设置合理。

关键词: 成果导向教育(OBE); 创新创业教育; Rasch模型

【DOI】 10.12252/j.issn.2096-6288.2022.08.007

成果导向的OBE(Outcome-based Education)理念自20世纪90年代由Spady在其著作《以成果为本的教育: 争议和答案》中提出后被广泛接受和推广, 并成为美国、英国等发达国家高等教育改革的主流思想(刘衍聪, 李军, 2018)^[1], 它是一套系统的、以学生为中心的教学范式和方法, 强调教学设计和教学实施的目标要指向学生通过教育过程最后所取得的学习成果(Spady, 1994)^[2]。显然, OBE理念能够帮助高等学校提高教学质量、完成预期目标, 而就OBE课程实施的效果则可以通过正确测量学生对OBE课程的反应和认知来获得。Rasch模型一种功能强大的测量模型, 能为等级量表提供一个有效的框架。以黑龙江八一农垦大学宏观经济学OBE课程为背景, 在课程预期学习成果与评估任务中体现“创新”教育目标。在Christopher(2013)^[3]研究基础上, 增加测量大学生对创新兴趣与能力培养基本条件的认知, 设定包含14个项目的李克特量表, 运用Rasch模型与Winsteps 5.5.0.0软件考察量表的有效性与可靠性

一、OBE课程预期学习成果与评估任务转换

OBE课程教学与传统课程教学存在的主要差别是要将其课程目标(course objectives)转换为可预期的课程学习成果(Course Intended Learning Outcomes, CILOs), 其中一个显著的变化是从使能的行为转化成对学生示范的期望, 它是一种从关注教师到关注学生、从关注输入到关注输出的转变。选择黑龙江八一农垦大学两个班的宏观经济学进行OBE课程改革初探, 同时致

力于大学生创新能力提升为目的确定课程目标与评估任务/活动, 并将其与传统课程进行对比(表1)。其中“*”号标注的是对创新能力培养起到积极作用的预期成果, 而I-CILOs则是学生在具体学习生活中形成创新能力应具备的基本条件特征。

表2列出了传统课程目标与OBE课程目标的评估任务/活动, 依据Wang(2010)^[4]的四个原则来制定评估标准。从两个表中可以看出, 课程任务与评估成果有关的具体目标是存在密切联系的, 相对于传统课程目标的评估来看, OBE课程目标效果的评估更加多样化, 而且注重学生实际能力的提升和研究兴趣的培养, 同时也允许学生有不同的机会去展示他们的学习成果。

二、评定量表设定

学生数量共计75人, 调查问卷结合Christopher(2013)^[3], 共创建14个调查项目(items), 其中与创新能力培养相关的有7个项目, 括号中表明与哪个预期成果相关联, 这7个项目分别是: 教学活动的准备安排效果好, 教师在教学活动中提供的学习阅读材料有帮助(S-CIL01); 相对于其他课程, 该课程更有难度*(I-CIL02, 3); 相对于其他课程, 该课程课业负担更繁重*(I-CIL02, 3); 相比其他课程, 对该课程更有兴趣*(I-CIL01, 2, 3); 用在该课程上的课余学习时间多*(I-CIL03); 该课程对于现实问题分析具有指导作用*(I-CIL01, 2); 期待该课程对于创新竞赛有帮助*(I-CIL01, 2, 3)。项目的建立依据Joint Committee on Standards for Educational

Evaluation (1988) 的要求, 围绕着OBE课程教学改革是否适当、可行、有用且准确: 对原始数据设定5点正填充 (positive pack) 评定量表, 并采用Christopher (2013) [3] 的计分方式 (1/强烈不同意; 2/不同意;

3/中立, 4/同意, 5/强烈同意)。需要说明的是由于对OBE课程建设属初步尝试, 试点学生人数较少, 但满足Rasch分析的条件 (其中为受试者人数, 为量表项目数)。

表1 传统课程目标与OBE课程预期学习成果比较

传统课程提纲 (course objectives)	OBE课程提纲 (CIL0s)
目的是使学生能够: (1) 理解经济学的基本内涵, 了解它与其他相关学科的关系以及应用领域 (2) 掌握 (微观) 经济学的主要内容 (3) 分析、解释 (微观) 经济现象与经济问题	课程完成后, 学生能够: CIL01* 清晰说明 (微观) 经济学的本质、结构与功能 CIL02* 掌握经济学的研究方法, 构建经济学逻辑思维 CIL03 清楚理解经济学重要理论和应用的作用与价值 CIL04* 具备一定的 (微观) 经济学理论素养, 学会运用经济学定理、规则描述、解释经济现象, 发现并解决经济问题 CIL05* 产生一定的研究兴趣 (了解经济分析的延伸学科, 如统计学、计量经济学等) 课程完成后, 在创新能力 (Innovation) 体现上, 学生期望能够: I-CIL01 乐于思考, 善于发现问题 I-CIL02 努力“创新”, 积极参加各类创新竞赛活动 I-CIL03 不惧挫折, 逐步培养锲而不舍、百折不挠的精神

表2 传统课程与OBE课程目标评估任务/活动 (assessment tasks) 比较

传统课程目标评估任务 (course objectives)	OBE课程目标评估任务
(1) 平时成绩占比30%, 包括出勤、课堂表现, 作业及小测验 (课堂或课外) 等 (2) 期末成绩占比70%, 闭卷考试, 题目类型: 选择、辨析、问答、作图、计算、分析等。	平时成绩占比50%: (1) 将学生分组, 每周有一组介绍一个经济现象, 并提出相应的个经济问题→CIL02, 3, 4 (20%) (2) 分组撰写一篇专业论文, 组内学生均要作出相应贡献, 根据贡献获得加权分数→CIL04, 5 (30%) (3) 就文章框架进行15分钟的小组展示, 组内学生获得该组分数→ CIL01, 2, 3, 4&5 (20%) (4) 10次线上周测验→CIL02, 3, 4 (30%) 期末成绩占比50%

三、Rasch测量的基本检验

1. 拟合度检验。Rasch测量的拟合度指标有: (1) 残差均方 (加权的Infit与未加权的Outfit), 二者分别对目标上 (位于参与者能力度量附近的项目的不稳定影响) 和目标外 (远离参与者或项目度量的异常) 的反应模式敏感; (2) 均方指数MnSq (Mean Square), 其期望值为1, 用来表示测量中异常的大小, 它为最主要参考的拟合度指标; (3) 显著性t检验的标准化ZSTD统计, ZSTD指标则提供了范围在-1.96与+1.96之间的t检验结果。取MnSq取值范围为0.5~1.5之间, Rasch结果显示, 项目14的Infit与Outfit值 (1.75和1.92) 大于1.5, 说明项目14与其余13个项目相关的潜在单一维度不匹配, 这个项目测量的潜在特质可能属于一个不同的维度。从项目的相关性上看, 相比所有项目代表的单Rasch维度, 项目14的相关度也不高, 只有0.47, 这也

进一步说明直接将“创新”的目标 (或预期学习成果) 纳入评估OBE课程目标还需审慎考虑, 而从带有“*”号与“创新”能力培养相关的项目上看, 间接考察促进创新能力提升目标的基本条件在评估OBE课程中更为稳妥。项目14的MnSq值与相关度指标虽然未在可接受的范围, 但是与可接受范围差距并非很大, 鉴于样本数据有限与调查的偏差, 是否删除项目14还要比对Rasch模型其他测量结果后再确定。

2. Rasch模型的两个假设检验。(1) 单一维度假设测试项目仅包含一种潜在特质, 即一个测试是否测量了它想要测量的潜在特征而并非以外的结构 (Wright & Stone, 1999) [5]。若满足单一维度假, 三个重要参数需符合: 残差特征值小于3.0; ②首成分残差小于15%; 有超过40%的原始方差可以被量表所解释。Winsteps运行结果显示, 残差特征值为2.1849, 小于3, 说明仅有一

个维度；首成分残差为13.5732；而原始方差可以被量表所解释的部分为51%，均在合理取值范围内，满足单一维度假设。（2）局部独立假设认为在潜在特征条件下，一个测试项目不与其他项目共变。验证这一假设的参数为标准化残差的相关性，一般要小于0.3，如果大于这一数值预示着存在相当程度的局部依赖性。运行结果显示，除项目14，其余项目参数均小于0.3。

3. 可靠性、区分度指标与类别设置分析。一般而言，Rasch模型要求个体与选项的可靠性的可接受范围分别为大于0.8与0.9，而个体与选项的区分度应分别大于2和4，若区分度大于2则意味着该测试能够区分项目/人的难度/能力。Rasch测量分析结果显示，参与者与项目的可靠性分别为0.85与0.96，对应的区分度分别为3.25与5.58，均在可接受的范围内。类别设置分析是用来判断量表选项类别（category）设置是否有效，验证的统计指标主要包括类别平均值（observed average），Andrich临界值（Andrich threshold）与类别测量值（observed measure）。Rasch模型选项分类检验结果表明三个平均值均单调递增，总体上看量表选项设置是有效的。然而，类别差异大小并非等距，3/中立与4/同意之间区分度明显，这可能同样因为被试数量不足，或不可避免的存在Halo Effect（因为成见而不假思索地为所有选项选择同一个值）。因而在后续的研究中，一方面需要增加样本容量，同时尝试更加精细的分类调整，如6点或7点量表，另一方面正确引导学生对OBE课程的认知，着力提升学生问卷作答的质量。

综上，Rasch的区分度结果表明自评量表关于参与者与项目具有较好的区分能力，参数估计具有较好的稳定性，并且信度结果表明量表在潜在特质的不同水平之间也具备充分的区分能力。同时，量表选项类别设置是有效的，但选项类别间差距欠均衡。就Rasch模型的整体测量结果，建议先不删除项目14，虽然实际数据不能完全符合Rasch模型的规定，但从解决问题这一目的来看，只要理论结果与实际近似就可接受，也不需要过分精确（杨建原和曾薇，2012）^[6]。因而本研究设定的评

定量表可以被看成是一种有效并可靠的测量工具，可以作为用来测量OBE课程实施效果以及促进创新能力提升目标的参考。

参考文献

[1] 刘衍聪, 李军. 基于OBE理念的应用技术型人才培养方案的设计[J]. 中国职业技术教育. 2018(4).

[2] Spady, W. Outcome-based education: Critical issues and answers. Arlington[R], VA: American Association of School Administrators. 1994.

[3] Christopher D., Gavin, T.L., Trevor G.B. & Ronnie S. Understanding outcome-based education changes in teacher education: evaluation of a new instrument with preliminary findings[J]. Asia-Pacific Journal of Teacher Education. 2013, 41(4).

[4] Wang, L. Designing and implementing outcome-based learning in a linguistics course: A case study in Hong Kong[R]. Procedia-social and behavioural sciences. Kyrenia: Elsevier. 2010.

[5] Wright, B., & Stone, M. Validity. Measurement essentials (2nd ed) [R]. WIDE RANGE. 1999.

[6] 杨建原, 曾薇. Rasch模型在等级量表设计中的应用[J]. 中国考试. 2012(05).

作者简介: 蒋诗(1976.4-), 女, 汉, 黑龙江省北安市, 博士, 黑龙江八一农垦大学, 研究方向: 宏观经济统计计量。

石建飞, 男, 1975年12月-, 硕士, 副教授, 硕士研究生导师, 黑龙江八一农垦大学信息与电气工程学院。

文春玲, 女, 1977年11月-, 博士, 讲师, 硕士研究生导师, 研究方向为农业经济与政策。

课题来源: 黑龙江省高等教育教学改革项目“基于OBE理念的大学生创新创业教育融入高校人才培养方案的研究”(项目编号: SJGY20200498)