

基于学习分析的网络课程评价指标体系研究

梁 凯

(洛阳理工学院 信息化技术中心,河南 洛阳 471023)

摘 要:目前网络开放课程的评价与管理是远程教育研究者关注的重点。基于对学习分析在网络开放课程评价中的主要作用的阐述,提出了由四个维度组成的学习分析参考模型,并制订了完整的评价指标体系,针对评价指标体系进行了信度与效度分析,最终结果显示评价指标体系的信度和效度较好。

关键词:网络课程;学习分析;评价指标;信度;效度

中图分类号:G434 **文献标识码:**A **文章编号:**1672-7800(2017)07-0041-03

DOI:10.16735/j.cnki.jet.2017.07.018

随着网络与通讯技术的发展,网络开放课程(简称网络课程)已经成为广大学习者认可的一种学习方式,因此对于网络课程的评价是教学管理的一项重要工作,一直以来也是教育信息化研究者关注的重点。评价指标是针对网络课程的量的规定,评价体系是相对完整的、具体的可测量的、行为化的一套评价准则,是评价对象主体各种要素及其互相关联的多个量规的有机整体^[1]。

1 目前网络课程评价标准存在的问题

1.1 忽视学习者参与因素

学习者作为网络课程的使用主体,也是课程评价中的主要因素之一,其参与课程学习的记录和数据也是课程评价的重要参考依据。目前美国的“在线学习的认证标准”(E—Learning Certification Standards)和国内的 CELTS—22.1《网络课程评价规范》等标准大多从教学内容、教学管理、技术应用等层面展开评价,没有明确的显示基于学习者参与度的评价标准。

1.2 缺少数理实证研究分析

通过对已有文献研究发现,对网络课程的评价指标体系的研究大致按照逻辑演绎和归纳总结两种思路。一种是从上到下、基于理论的、注重逻辑演绎的研究思路,与欧美的自下而上、基于数据驱动的、注重统计分析的研究思路完全相反,而且评价标准体系大都来源于“瀑布式”的理论推演,没有信度和效度分析。另一种思路是从研究者经验归纳得出的评价指标,在特定范围内具有指导实践工作的意义,但是指标的科学性也无法考证。当同一门课程采用不同的评价标准或方案进行考核时,得到的评价结果也没有

可比性,因此评价的可信度受到人们的质疑。

1.3 评价对象认可度不高

网络课程的考核和评价对网络学习活动的良性开展具有重要的意义,评价指标即评价量规,它对网络课程的建设发展具有重要的指导作用。调查发现,在网络课程的教学评价工作中,经常出现评价内容或标准的设定偏离评价主体的情况,因此造成评价不被教师认可,评价结果可信度低,导致人财物严重浪费。事实上只有评价主体接受评价结果,才能改进网络教学,促进学生的发展。

2 学习分析理论

学习分析技术最早出现于美国新媒体联盟(NMC)2010和2011的年度出版的“地平线报告”,报告中提到学习分析技术将成为未来的几年内的主流^[2]。目前教育技术界对其尚无统一定义。地平线报告将学习分析定义为:通过使用各种数据挖掘工具和可视化数据分析技术对学生的行为数据进行分析,研究其投入情况,并将其分析结果应用于个性化教学以及课程修订和教学评价中。美国高等教育信息化推进组织 EDUCAUSE 认为:学习分析技术即利用数据和模型来预测学习者的学习进度和投入情况,发现潜在的问题并及时给与干预^[3]。

我国部分学者认为,学习分析是围绕与学习者学习信息相关的数据,运用不同的分析方法和数据模型来解释这些数据,根据解释的结果来探究学习者的学习过程和情景,发现学习规律;或者根据数据阐释学习者的学习表现,为其提供相应的反馈,从而促进更加有效的学习^[4]。我们认为,学习分析的核心要素包括“测量学习效果、收集学习记录、

收稿日期:2017-05-29

基金项目:河南省教育科学“十二五”规划课题(2015—JKGHZD—0024)

作者简介:梁凯(1976—),男,河南洛阳人,洛阳理工学院信息化技术中心硕士研究生,研究方向为教育信息化。

生成学习记录报告,反馈优化学习”。

学习分析技术采用大数据的处理方法,收集并聚类分析学生的学习记录。其原本的作用在于将数据反馈给教师或学习者以便改进学习。从课程管理的角度来看,我们完全可以使用学习分析技术作为学生参与学习的一个量的标准,从侧面来反映出一个课程的优劣。

3 基于学习分析的评价模型

网络开放课程质量评价的指标体系应该是促进学习者发展所应达到的一系列相关目标组成的整体。我们针对学习分析的核心要素提出了一个三个维度组成的评价参考模型。该模型将数据与环境、技术、利益相关者、目标四个维度分别对应于网络课程评价的三种参与者:专家、同行、学生。依照国际上通用的三层结构模型,将上述的评价模型分为三个级别的指标组成,多个评价视角的相结合的评价指标体系。

体系的一级指标分为专家、同行、学习分析三个维度,二、三级指标分析研究了美国教学设计与适用专业委员会建立的“在线学习的认证标准”(E-Learning Certification Standards)、美国国家教育协会、美国高等教育政策研究所和 Black Board 公司联合发布的远程学习评价标准(Quality On The Line)^[5]和 CELTS-22.1《网络课程评价规范》中的相关指标。如表 1 所示。

表 1 评价指标体系

维度	一级指标	二级指标综合
专家维度	课程理念	广泛性、社会性、主题性、开放性
	课程目标	清晰性、指导性、全局性、结构性、可操作性
	课程内容	科学性、实用性、兴趣性、规范性、系统性、先进性、拓展性
	课程评价	及时性、激励性、导向性、客观性、有效性、多元化性
同行维度	教学内容	教学资源建设、课程内容完整性、课程内容设计、拓展资源
	教学设计	教学应用、教学管理、教学档案、学生档案
	界面设计	基本技术规范、页面效果
	技术运行	技术要求规范、整体效果
学习分析维度	学生参与者学习状况	录频度、时长,学习资源数量、学习资源访问量频度、时长,作业提交率,论坛话题数量、话题平均活跃度,单元测验平均成绩
	教师角色	平均在线时长,教学资源知识产权状况(独创率),作业批改率,答疑次数和频度
	辅导状况	

4 评价指标体系的信度与效度分析

信度(Reliability)即可靠性,它是指采用同样的方法对同一对象重复测量时所得结果的一致性程度。效度(Validity)即有效性,它是指测量工具或手段能够准确测出所需测量的事物的程度^[6]。效度测量结果与要考察的内容拟合度越高,则效度越高;反之,则效度越低。信度和效度是评价指标体系进行实证研究的重要部分,只有做了信度和效度研究才能确定评价指标体系的可靠性,才具有公信力。我们在河南省某高校推行网络辅助教学,使用本评价指标体

系进行考评,考评采用问卷法和基于学习分析的数据挖掘相结合的方法,并进行对比研究,分析评价指标体系的信度和效度。

4.1 对象样本

调查取样是在洛阳理工学院网络课程遴选专家中展开,研究采取网络投票问卷调查的方式,采集网络课程评价的各种统计样本数据。

第一次评价问卷为 2011 年 10 月,参与评价的课程为 35 门,参与网络评价问卷的人数共 65 位,其中专家共 20 位,同行 45 位。共收到问卷 2275 份,每门课平均调查时间为 28min。第二次评价问卷时间为 2012 年 11 月,参与评审问卷的人员与第一次相同,共评审 48 门课程,收回问卷 3120 份。样本数分布如表 2 所示。

表 2 调查样本分布表

院系名称	评价专家	评价同行	比率
材料科学与工程学院	3	5	0.12
计算机与信息工程学院	4	5	0.14
电气与自动化学院	2	5	0.11
机械工程学院	3	5	0.12
经济与工商管理学院	1	5	0.09
人文与社科学院	2	5	0.11
外国语学院	1	5	0.09
教育科学与音乐学院	2	5	0.11
环境与化学工程学院	2	5	0.11
合计	20	45	1

4.1 分析方法

自制调查量表。专家维度和同行维度的在线评价问卷依据本评价指标体系制定,其内容主要包括 3 个评价维度,9 个一级评价指标,35 个二级评价指标,如课程目标、课程内容(专家)、课程实施、课程评价、课程理念、教学内容(同行)、教学设计、界面设计、技术运行等 9 个一级指标。回收的数据采用 SPSS15.0 软件进行统计分析,验证评价体系的信度与效度。

4.2 结果分析

信度分析。信度分析主要检查内部一致性,可用它表示评教指标内部之间的关系,评测评价指标中各个条目是否测量了相同的内容或特质。我们将被试内容多个部分的同种属性,使用相同的测量标准检验,反映评价指标之间存在着高的相关性。

从表 3 数据中可以看出,A 评价中各指标的 α (Cronbach's Alpha)等于 0.923,大于 0.9,在 B 测试中各指标的 α (Cronbach's Alpha)等于 0.889,可以认为基本接近 0.9。综合分析显示,两次测试中 α 值都满足量表的信度系数要求,说明评价指标之间内在信度较好。

表 3 两次测试的 Cronbach's 系数

测试	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
第一次试用评价(A)	0.923	0.919	9
第二次试用评价(B)	0.889	0.893	9

表 4 剔除某一级指标后信度分析

A 第二次评价					
类别	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
课程目标	62.4488	35.750	0.792	0.517	0.892
课程内容	62.8744	33.423	0.873	0.640	0.882
课程实施	62.7680	36.583	0.792	0.501	0.892
课程评价	62.6403	34.623	0.898	0.658	0.879
课程理念	62.5978	39.904	0.624	0.349	0.909
教学内容	62.1714	34.7340	0.8580	0.2870	0.907
教学设计	62.2991	37.3370	0.7980	0.6540	0.895

B 第一次评价					
类别	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
课程目标	63.3478	50.187	0.662	0.474	0.904
课程内容	64.1087	51.032	0.796	0.685	0.894
课程实施	59.087	37.992	0.843	0.768	0.896
课程评价	64.1087	49.388	0.762	0.665	0.895
课程理念	64.3043	48.972	0.782	0.682	0.892
教学内容	64.0435	53.909	0.604	0.425	0.910
教学设计	64.8696	48.471	0.824	0.730	0.888

在表 4 中,第 6 列表明删除特定指标条目后的 α (Cronbach's Alpha)系数。第 5 列表示该指标与其他指标的复相关系数,反映了当前指标与剩余指标的之间的总体相关度。从表中不难看出,A 试用评价中,去除任何一项指标都会使评价量表的信度系数降低。两次评价中“教学内容”指标与其他指标的相关系数分别为 0.287 和 0.329,当去除“教学内容”后,信度系数有所增加,分别为 0.907 和 0.910,可以说明当前指标与剩余指标之间相关不高。在“教学内容”的深度分析检查中,认为其二级指标“教学资源建设”与剩余一级指标的相关系数不高,与课程目标、课程理念、教学设计、课程实施的相关系数分别为 0.163、0.172、0.231、0.327*,说明仅仅与课程实施的相关系数具有实际意义,其他相关系数均不明显。

4.3 结构效度检验

(1)KMO 和 Bartlett's 检验。两次测试的检验结果显示,Bartlett 的球形度检验结果都具有判定意义,所以判定相关系数矩阵与单位矩阵有着显著差异;结果还显示 A 评价的 KMO 度量 0.810,B 评价的度量为 0.793,二者都大于 0.7,利用 Kaiser 标准度量,可以看出评价指标体系的效度较好,如表 5 所示。

表 5 两次评价 KMO 和 Bartlett's 检验

Kaiser Meyer Olkin Measure of Sampling Adequacy.	A 评价	B 评价
Approx. Chi Square	0.810	0.793
Bartlett's Test of Sphericity	453.376	349.357
df	113.000	89.000
Sig.	0.000	0.000

(2)结构方程模型适配。A 评测中数据模型的 Chi—

square 检验值为 2.456,P=0.003,p<0.01,具有显著性。B 评测中数据模型的 Chi—square 检验值为 8.234,p=0.004,P<0.01,具有显著性。所有检测数据都表明其协方差矩阵与评价数据的 NFI、RFI、CFI、IFI 模型适配度指数 S>0.9,表示模型匹配度良好。A 评测数据的 RMSEA=0.310,B 评测中数据 RMSEA=0.297,可以看出整体 RMSEA>0.1,从麦卡勒姆的 RMSEA 模型分割原理分析^[7],如果 RMSEA>0.10,可表明模型的适配度不高。

表 6 评价模型适配度

模型	χ^2/df	P	NFI	RFI	RMSEA	CFI	IFI
A 试用模型	2.456	0.0003	0.952	0.908	0.310	1.000	1.026
B 试用模型	8.234	0.0004	0.960	0.925	0.297	1.000	1.026

有学者研究提出^[8],当 RMSEA 值在样本较少的检验环境中可能存在虚高现象,导致研究模型的适配度出现不佳的现象。表 6 数据中整体模型适配度良好,仅存在 RMSEA>0.10,分析原因可能为样本数量偏小造成。A 测试中样本数为 2275,B 测试中样本数为 3120,在结构方程模型测试中都不太满足样本最小测试量。

5 结语

本文梳理了学习分析在网络开放课程评价中的主要作用,在四个维度组成的学习分析参考模型的基础上,研究制订了完整的评价指标体系。在网络课程建设的实践应用中,通过专家的评价数据分析评价指标体系的信度和效度,结果显示评价指标的效度较好。但由于本研究环境的局限,仍有个别指标不太理想,仍需未来进一步分析验证。

参考文献:

[1] 高丽. 高质量 OA 期刊文献的选取方法研究[D]. 西安:西安电子科技大学,2014.

[2] 魏顺平. 学习分析技术:挖掘大数据时代下教育数据的价值[J]. 现代教育技术,2013(2).

[3] 王苗. 学习分析技术在网络课程学习中的应用实践研究[D]. 长春:东北师范大学,2014.

[4] 顾小清,张进良,蔡慧英. 学习分析:正在浮现中的数据技术远程教育,2012(1).

[5] 孙传远,孙少华. 专家视野中的开放课程质量评价指标体系——基于德尔菲法的调查分析[J]. 职教论坛,2014(9).

[6] 牟宁. 面向制造业第三方物流企业服务能力评价研究[D]. 昆明:昆明理工大学,2010.

[7] MACCALLUM R C, WIDAMAN K F, ZHANG S & HONG S. Sample size in factor analysis[J]. Psychological Methods,1999(1).

[8] 吴明隆. 结构方程模型—AMOS 的操作与应用:第 2 版[M]. 重庆:重庆大学出版社,2010:41.

(编辑:邬江华)