

网络课程评价工具设计*

李兴敏

(深圳大学, 广东 深圳 518060)

摘要 本文首先论述了网络课程评价采用实证研究方法的客观依据,然后基于 2007 年网络精品课程的评价指标体系,结合模糊综合评价法对网络课程定量评价的判断能力,在多途径保证算法更为合理的基础上,采用计算机编程实现在线评价模型,为网络课程评价提供一种快捷、方便的新工具。

关键词 网络课程;实证研究;评价;指标体系

中图分类号 G434 **文献标识码**:A

一、研究背景

网络课程的评价通常包括两层含义:一是形成性评价,即在网络课程开发过程中实施的评价,评价的目的是改进和提高正在开发的网络课程质量;二是总结性评价,即在网络课程开发结束后进行的评价,评价的目的是对网络课程的性能和效果做出定量和定性的描述。本文所涉及的网络课程评价都围绕总结性评价展开。

笔者通过文献追溯发现,多数网络课程评价模型是将模糊数学建模方法运用于网络课程评价,即将影响网络课程教学质量与效益主要因素的定性评价转换为有限定量评价指标,并且设计出多种评价系统和评价模型,最后产生数字化的评价结果,这种实证的评价有其存在的客观背景。

1. 网络课程评价的主动权在于各级教育管理机构^[1],何时开展评价、如何进行评价由行政机关单方决定,而网络课程的开发者(教师)和使用者(校方)几乎没有发言权,只能被动接受评价,因为网络课程的评价结果经常是课程推广和财政支持力度的依据,这种单方面决定、行政色彩浓厚的评价,更注重的是外部显现的结果,经常忽视对网络课程整体水平的关注和内在素质的提高,客观上就是为排名而评。

2. 从教育研究方法论的文化背景分析,建国以来我们长期偏重于教学经验的简单总结,很少进行量化的评价分析。但 20 世纪 80 年代中期开始,教育统计学、测量学的研究进展到一定阶段,实证量化的方法在教育研究中广泛应用,而国内网络课程产生于 2000 年互联网普及前后,其发展历程刚好处于倾向使用实证方法体系评价课程的时期。

3. 从社会文化背景分析,唯实证化评价是人们长期对自然科学盲目崇拜的结果^[2]。来自西方世界的自然科学的巨大成就,使人们陶醉于其方法的优越性之中,从而在社会科学领域试图将教育实践中一切有序和无序的、清晰的和模糊的、可测的和不可

可测的东西都纳入科学实证的框架加以整合。

4. 电教从业人员知识结构背景近年已经发生了巨大变化,由于计算机类相关专业的连年扩招,大量的计算机类专业毕业生进入教育技术行业,他们依赖于自身专业背景和良好的数学功底,评价网络课程时更习惯于用实证方法研究问题。

5. 从物质基础上观察,目前的教育技术已经彻底摆脱了过去以录音、录像和广播电视为主要应用工具的原始阶段,计算机辅助教育得到了大规模的应用,每个学校的教育技术信息中心几乎都装备有本校计算机与网络的尖端设备,从而为教育技术工作者开展实证量化计算奠定了充分的物质基础。

6. 计算机辅助教育,尤其是最具划时代意义的远程教育近年得到迅猛发展。在这一历史阶段中,教育技术工作者已经逐渐掌握了计算机和网络的技能,这样他们用实证化方法来评价网络课程比以往更为方便和快捷。

然而综观大量的网络课程评价实证类研究文献,由于急于求得或过于关注数字化的评价结果,多数作者在研究过程中恰恰忽略了模型运用的合理性,其数字化的结果自然缺乏可信度^[3]。本文作者曾在完成广东省高校人文社科研究项目“IT 设备评标应用系统”的过程中,运用模糊优选理论,设计了 IT 设备采购综合评标模型和应用演算范例,验证了利用集合论对设备采购进行评比的可行性和合理性。所以本文旨在借鉴模糊优选理论设计评标模型的经验,将其应用于网络课程评价,同时借助业内普遍认可的评价指标体系,为网络课程评价提供进一步合理化的新工具。

二、多途径保证网络课程评价的合理性

1. 防止评价内容的技术取向和片面化

网络课程评价包含哪些内容是一个有争议的问题,而且受到评价者研究取向的影响,国内具体评价指标体系发展也较快,刘成新提出了网络教学设计的评价项目表,李葆萍等也指出网络课程评价的 9

个维度,徐福荫教授和李运林教授将其分为“教学性”“科学性”“技术性”“艺术性”和“使用性”共 5 个方面^[4],何克抗教授等为网络课程和各类电子教材制定了包括 3 个一级指标、10 个二级指标和 26 个三级指标的评价指标体系。2002 年 6 月 7 日,教育部教育信息化技术标准委员会发布了《网络课程评价规范 CELTS-22(标准草案征求意见稿)》,其采用的指标体系一度成为网络课程评价和建设的重要标准。该规范指标体系包括 4 个一级指标和 26 个二级指标。

评价指标体系不是一成不变的,并且始终处于一个不断接近如何更利于学生学习这一宗旨的过程^[5],目前评价指标设计逐渐成为网络课程建设的标准和依据,网络教育工作者更加理解网络课程的建设意义,在系统初期设计、中期开发和后期建设中有的放矢,有效地促进了网络课程的发展。

2007 年网络精品课程的评审和指标出台,进一步完善和细化了网络课程的评价体系,是一套适应性的评价指标体系。它包括教师队伍、教学内容与学习资源、远程导学与学习活动、学习支持服务、教学效果、特色与政策支持共 6 个一级指标,又进一步可划分为 18 个二级指标和 34 个三级指标^[6]。该指标体系投入使用,曾对 59 所现代远程教育试点高校申报的 189 门网络教育课程评审,通过资格审查、网上评审和会议终审三个环节,包括 168 名专家的专家组最终选出 50 门精品课程。

网络课程评价的指标体系需要进行一定程度的集中,虽然这种硬性的集中未见得其采用的指标体系是最优的,却是必须的,否则,网络课程的质量比较就无从谈起。2007 年出台的网络精品课程评价指标体系,理论上较为容易在当前条件下被业界普遍接受^[7],也是目前全国网络课程较为统一的评价标准。

2. 评价方法尽量合理与指标权重的慎重分配

2007 年因网络精品课程评审而出台的评价指标体系,并没有规定如何分配各个指标的权重,权重的变化就会引起评价结果的变化。所以对网络课程进行整体的质量评价,必须审定各个指标的权重。确定权重的方法有专家会议法、特尔斐法、层次分析法、逆向法。有关文献中比较认同集中专家力量利用层次分析法审定指标权重^[8],同时还要注意专家的专业组成应包括网络课程所在学科、电脑设计、教育心理学等领域,并配合使用座谈、访谈、观察等辅助手段保证层次分析法的正确运用。

3. 评价等级的设定和赋值

网络课程评价必须设定评价等级,评价等级划分得越细,计算出的隶属度矩阵就越准确,结论就越接近于科学^[9],但在线应用要考虑参评者对评价等级

的理解能力,等级含义要清晰明确,便于把握。笔者认为根据人们的生活习惯划分为 5 个等级(优秀、良好、一般、合格、不合格),更为容易被在线评价者理解,同时,每个评价等级的赋值也可以集中专家力量采用层次分析方法获得。

4. 防止研究对象选择的单一,分类在线评价者作为研究对象

绝大多数研究都将焦点集中在专家或教师上,从教师的关注、感受和教学行为、教学理念等不同的角度来研究课程的实施程度^[10],同时,校长也得到一些关注,而学生、行政人员和家长则少有研究。笔者认为对网络课程质量有评价资格的人员包括专家型教师、一般教师、在学学生及其家长以及网络课程教师所在单位的校长或者其委托的代表。由于参评人员的身份和资历不同、对网络课程了解程度和经验也不同,个人的表决权重也应有所不同,这点在参评者在线打分前的注册环节,就先由专家小组事先设计好个人的表决权重,并在本人不知情的情况下自动赋给(下面算法为简便起见,没有涉及)。

三、在线评价过程的解析

在线评价是网络课程评价发展的必然归宿,网络课程是在互联网平台上开发和应用于教学的,网络课程评价决不能失去这一平台,在线的评价可以包容各类参评人员,评价的过程趋于透明,评价结果学员随时可见并供选择课程时参考。

参与网络课程评价者在线填写评语等级记录表(即下页表,评价因素及其权重值也体现在该表中),每个评价者根据对网络课程的判断,对表中的每个评价因素判定其评价等级(优秀、良好、一般、合格、不合格),然后在表格中打“√”,形成评委对某门网络课程的隶属度矩阵,从而计算一段时期内全体评价者对该门课程的综合评价,直接在计算机上操作并自动生成结果。算法解析如下。

第一步,开列评价因素集 $A, A = \{a_1, a_2, \dots, a_{34}\}$, 2007 年网络精品课程评审时出台的评价指标体系中计 34 项评价因素列于表格。虽然国内已经建立了很多套关于网络课程评价的指标体系,但是着眼新近出台并能够被多数专家学者认可的指标体系则选择余地很小,下页表的指标体系具有一定的代表性,在这套指标体系的基础上做网络课程评价工具的设计工作,具有一定的应用价值。

第二步,专家组通过层次分析法获得因素权重 c_i , 其中 $i=1, 2, \dots, 34$, 组成权重向量 $C, C = (c_i)_{1 \times 34}$, 见上表,即对应表中 34 个评价因素,每一个评价因素的权重值,可以理解为每一个评价因素在总共 34 个

评语等级记录表

评价内容			评价等级				
一级指标	二级指标	三级指标及其权重(a _i , c _i) $\sum_{i=1}^{34} c_i = 1$	优秀 B ₁ =90	良好 B ₂ =80	一般 B ₃ =70	合格 B ₄ =60	不合格 B ₅ =50
教师队伍	课程负责人与主讲教师	课程负责人与主讲教师(a ₁ , c ₁ =0.02)	190	167	41	2	0
	队伍结构	队伍结构(a ₂ , c ₂ =0.02)	193	151	49	7	0
	教学研究与从业经验	教学研究与从业经验(a ₃ , c ₃ =0.02)	134	161	93	12	0
教学内容 与学习资源	教学计划	内容选取(a ₄ , c ₄ =0.03)	138	170	77	15	0
		载体形式(a ₅ , c ₅ =0.03)	159	140	83	18	0
		传递方式(a ₆ , c ₆ =0.02)	131	157	94	18	0
	内容组织	规范性(a ₇ , c ₇ =0.02)	184	141	66	9	0
		知识覆盖面(a ₈ , c ₈ =0.02)	176	151	61	12	0
		扩展性资源(a ₉ , c ₉ =0.02)	167	163	63	7	0
	人机交互	界面设计(a ₁₀ , c ₁₀ =0.02)	162	176	55	7	0
		文字符号(a ₁₁ , c ₁₁ =0.03)	128	213	49	10	0
		交互设计(a ₁₂ , c ₁₂ =0.03)	113	222	57	8	0
	技术特征	技术适当性(a ₁₃ , c ₁₃ =0.02)	182	175	36	7	0
遵循标准(a ₁₄ , c ₁₄ =0.02)		185	159	44	12	0	
远程导学 与学习活动	导学策略	多种策略合理运用(a ₁₅ , c ₁₅ =0.01)	126	169	88	17	0
		激发学习兴趣(a ₁₆ , c ₁₆ =0.04)	130	178	72	20	0
	活动组织	学习目标(a ₁₇ , c ₁₇ =0.03)	151	148	78	23	0
		活动指导(a ₁₈ , c ₁₈ =0.02)	123	165	89	23	0
		活动支持条件(a ₁₉ , c ₁₉ =0.02)	176	149	61	14	0
	学习评价	评价手段(a ₂₀ , c ₂₀ =0.02)	168	159	56	17	0
		反馈及时性(a ₂₁ , c ₂₁ =0.02)	159	171	58	12	0
		学习过程记录(a ₂₂ , c ₂₂ =0.04)	154	184	50	12	0
	实践教学	实践条件(a ₂₃ , c ₂₃ =0.04)	120	221	44	15	0
		实践活动组织(a ₂₄ , c ₂₄ =0.02)	105	230	52	13	0
学习支持服务	服务内容	学术性支持(a ₂₅ , c ₂₅ =0.03)	175	182	31	12	0
		非学术性支持(a ₂₆ , c ₂₆ =0.03)	178	166	39	17	0
	服务人员	人员配备(a ₂₇ , c ₂₇ =0.04)	119	176	83	22	0
	途径与保障	学习支持服务的途径(a ₂₈ , c ₂₈ =0.04)	123	185	67	25	0
		制度(a ₂₉ , c ₂₉ =0.04)	144	155	73	28	0
规范性(a ₃₀ , c ₃₀ =0.02)	116	172	84	28	0		
教学效果	学生反馈	学生反馈(a ₃₁ , c ₃₁ =0.02)	169	156	56	19	0
	作业评价	作业评价(a ₃₂ , c ₃₂ =0.02)	161	166	51	22	0
	同行及自我评价	同行及自我评价(a ₃₃ , c ₃₃ =0.09)	152	178	53	17	0
特色及政策支持	特色及政策支持	特色及政策支持(a ₃₄ , c ₃₄ =0.09)	147	191	45	17	0

评价因素中的重要性体现。评价因素权重是专家组意志和倾向性的集中体现,它们的设置是否合理,直接影响网络课程的评价结论,本文第二部分已经介绍了常用确定因素权重的四种方法,其中根据网络课程的性质恰当地选择专家组成员,配合使用会谈等保证措施,利用层次分析法审定获得因素权重被认为是比较有说服力的权重获取方法。

第三步,统计评价时间段内参加总人数 n 和判定每个评语等级的人数,建立评价矩阵 U, $U=(u_{ij}/n)_{34 \times 5}$, $i=1 \dots 34, j=1 \dots 5$, u_{ij} 表示因素 a_i 对 b_j 的隶属程度^[1],在线评价开始到统计这一时段参与网络课程

评价的群体,他们会就评价因素课程负责人与主讲教师等因素给出评语等级的判定,这一群体中分别判定某一行上评价因素对应 5 个等级的人数比例,就对应形成评价矩阵 U 的一行。开放式在线评价使网络课程评价的群体得以有效扩大——教师、学生、家长和管理人员都能参与,从根本上克服了以往参加评价者限于被邀请专家的片面性,同时专家可以远程登录参与,既保证了他们参与评价的机会又缩短了评价周期。

第四步,计算评价向量 D, $D=(d_j)_{1 \times 5}$

评价向量 D 等于单因素评价矩阵 U 与权重向

量 C 的积, $D=C \times U$, 评价向量的每一项 $d_j = \sum_{i=1}^{34} c_i u_{ij}$ ($j=1,2,3,4,5$), $D=C \times U = (c_i)_{1 \times 34} \times (u_{ij}/n)_{34 \times 5} = (d_j)_{1 \times 5} = (d_1, d_2, d_3, d_4, d_5)$, 本步计算得到的评价向量反映了评价群体对该门网络课程总的看法, 即在五个评语等级中, 人们的看法更集中于哪一个等级, 形成根本观点在还需第五步计算完成后。

第五步, 建立综合评价模型计算综合评价值 S , 根据第四步结果, 加上专家组通过层次分析方法获得的评语等级集 $B = \{b_1, b_2, b_3, b_4, b_5\}$ (也就是优秀、良好、一般、合格、不合格共 5 个等级在专家组心目中的量化百分综合估值), 综合评价值 $S = D \times B^T$, 为网络课程的最后评分。评语等级的设定不宜过多或者过少, 过多则评语指向性模糊, 参与评价者不易判定; 过少则不能充分体现网络课程各个评价指标的质量差别。评语等级的赋值也是评审工作重要的一环, 它在将量化的评价结果转化成定性的评价结论过程中起着重要的作用, 所以一定要保证其等级评语赋值与具有话语权的专家们的看法量度一致。本步骤再次采用第三步中取得权重的层次分析法, 来取得等级评语值。

假定评价因素集 A 含 34 项评价因素, 如上页表所列, 评语等级集 $B = (90, 80, 70, 60, 50)$, 在线参加评价总人数 400 人, 他们就 34 项评价因素给定的评语等级见上页表, 可以整理得到权重向量 $C_{1 \times 34}$;

$$C_{1 \times 34} = (0.02 \ 0.02 \ 0.02 \ 0.03 \ 0.03 \ 0.02 \ 0.02 \ 0.02 \ 0.02 \ 0.02 \ 0.03 \ 0.03 \ 0.02 \ 0.02 \ 0.01 \ 0.04 \ 0.03 \ 0.02 \ 0.02 \ 0.02 \ 0.02 \ 0.04 \ 0.04 \ 0.02 \ 0.03 \ 0.03 \ 0.04 \ 0.04 \ 0.04 \ 0.02 \ 0.02 \ 0.02 \ 0.09 \ 0.09)$$

评价矩阵 $U_{34 \times 5}$ 是一个 34 行 5 列的大矩阵, 第一行数值为 $(0.475 \ 0.4175 \ 0.1025 \ 0.005 \ 0)$, 即是总人数 400 人中对课程负责人与主讲教师 a_1 指标评价为五个等级的比例, 同理得到剩下 33 行的数值, 共同构成该矩阵 $U_{34 \times 5}$;

评价向量 $D_{1 \times 5} = C_{1 \times 34} \times U_{34 \times 5} = (0.371625 \ 0.438675 \ 0.14965 \ 0.04005 \ 0)_{1 \times 5}$;

综合评价值 $S = D_{1 \times 5} \times B^T = (0.371625 \ 0.438675 \ 0.14965 \ 0.04005 \ 0)_{1 \times 5} \times (90 \ 80 \ 70 \ 60 \ 50)_{5 \times 1} = 81.41875$;

对照评语等级 (优秀、良好、一般、合格、不合格) 对应的评语等级集 $B = (90, 80, 70, 60, 50)$, 该门网络课程的综合评价应该为良好。开放式的在线评价不局限于某一门网络课程, 从理论上说, 每一门在线

的网络课程都是评价的对象, 参评人员也是开放管理的, 只需在参与评价前进行身份注册, 从而杜绝重复评价和划定个人权重, 评价的周期是相对的, 在线学习者随时可以参与评价, 随时可以浏览系统已经计算出的评价结果和暂时形成的评价结论。

四、结束语

这套网络课程在线评价系统基于 2007 年网络精品课程评审指标体系, 达到了由定性评价到定量评价再到定性评价的目的, 运用过程中综合各方面因素确定网络课程的评价等级, 体现了网络课程的整体质量, 利于选拔优秀网络课程, 可以实现对网络课程的有效监控和管理, 从而避免了一些在网络课程开发工作上的低水平重复。

当然, 客观地说绝没有一种最佳的实施网络课程评价的工具, 每一种工具都有一定的适用范围。但是由于实证研究背景的客观需要, 作为一种依据相对权威的指标体系并相对合理而便捷的评价工具, 必然会在当前的一定时期内, 发挥其应有的作用。另外上文对于网络课程评价方法的演算范例完全利用计算机编程, 借助互联网实施评价过程, 并且对在互联网上收集到的评价数据随时保存和有效管理, 根据网络课程评价的发展趋势, 适时地修改软件评价模型, 也就产生不同的评价结果, 该系统作为一种非单纯唯数据的评价系统, 对于长时间地研究网络课程的评价, 积累不同时期的评价结果而用于课程研究的分析预测, 也具有一定的历史意义。

参考文献:

- [1][2] 于洪卿. 论课程评价中的实证化倾向[J]. 江西教育科研, 2006, (3): 62-64.
- [3][5] 李兴敏. 基于模糊数学方法综合评价网络课程的新思考[J]. 现代远距离教育, 2008, (4): 29-31.
- [4] 南国农, 李运林. 信息化教育概论[M]. 北京: 高等教育出版社, 2004.
- [6] 曾海军等. 基于精品课程评审指标探索网络教育资源的建设与共享[J]. 中国远程教育, 2007, (10): 47-52.
- [7][9] 李兴敏. 结合层次分析法对网络课程进行模糊综合评价[J]. 电化教育研究, 2008, (8): 95-99.
- [8] 郭向勇等. 基于层次分析法的网络课程学习评价模型[J]. 电化教育研究, 2008, (3): 72-75.
- [10] 唐丽芳. 我国课程实施过程评价研究进展及走向[J]. 东北师范大学学报(哲学社会科学版), 2007, (2): 129-133.
- [11] 郭向勇, 傅国强. 以模糊数学方法构建多媒体教学质量综合评价体系[J]. 电化教育研究, 2007, (3): 76-80.

收稿日期 2008 年 12 月 29 日
责任编辑 张静然

* 本文是广东省高校人文社科研究项目“IT 设备评标应用系统”(项目编号: 02SJD63004) 成果之一。