

## 网络课程评价标准的研制

滕霞<sup>1</sup> 段崇江<sup>1,2</sup> 张建伟<sup>1</sup> 王学优<sup>1</sup>

1、清华大学电教中心、现代教育技术研究所 北京 100084

2、云南师范大学计算机科学与技术学院 昆明 650092

【摘要】该研究按照标准研制过程在开发阶段和建立共识阶段上的要求，建立了初步的网络课程评价标准指标体系，随后采用德尔菲法进行了两轮专家评议，初步形成了具有共识性的网络课程评价标准指标体系及其评价方案。

【关键词】CELTS-22，网络课程，评价，标准，专家评议

【中图分类号】G434

【文献标识码】A

【论文编号】1009—8097（2003）01—0005—07

### 一、背景

现代远程教育以计算机网络（以及卫星数字通讯）技术为支撑，具有时空自由、资源共享、系统开放、便于协作等优点。世界各国在发展现代远程教育时深刻认识到，学习资源的可共享性和系统的互操作对于网络远程教育的实用性和经济性具有决定性意义。目前，虽然网络技术已为教育资源在低水平上的自治与共享（例如通过 HTTP 和 HTML）学习活动的合作（例如通过各种通信工具）提供了基本技术条件，但是允许教学资源在课程知识和教学管理水平进行交换的标准却没有很好地认定，因此妨碍了教学资源的大范围共享与交流。有鉴于此，国际上已有不少国家和组织致力于远程教育技术标准的研究。我国的远程教育技术已经开始进入以网络为基础的新阶段。然而，由于我国目前还没有制定关于网络远程教育的技术标准，各网络教育系统的资源自成体系，无法实现有效交流和共享，造成大量低水平的重复性开发工作，不但带来人力物力的浪费，而且将无法与国际网上的教育体系相沟通。根据以上对国内外现代远程教育形势的分析，教育部科技司于 2000 年启动了现代远程教育标准研究项目，简称 DLTS（Distance Learning Technology Standards）项目，以国际国内现代远程教育的大发展与大竞争为背景，以促进和保护我国现代远程教育的发展为出发点，以实现资源共享、支持系统互操作、保障远程教育服务质量为目标，通过跟踪国际标准研究工作和引进相关国际标准，根据我国教育实际情况修订与创建各项标准，最终形成有中国特色的现代远程教育标准体系。为了适应我国教育信息化的长远发展，2001 年 12 月，教育部现代教

育技术标准化委员会更名为教育部教育信息化技术标准委员会，英文简称为 CELTSC（Chinese E-Learning Technology Standardization Committee），相应地，原发布之远程教育技术标准（DLTS）将更名为网络教育技术标准，简称 CELTS（Chinese E-Learning Technology Standardization）系列标准。2002 年底，经国家标准化管理委员会办公室批准成立了“全国信息标准化技术委员会教育技术分技术委员会”。

网络课程的质量是制约整个网络教育质量的一个关键环节，网络课程的评价标准问题是当前网络教育发展中必须尽快解决的一个重要问题。在现代远程教育标准研究项目中，我们承担了 CELTS-22 “网络课程质量评价规范”的研制工作，本文就来介绍 CELTS-22 标准研制的方法过程和基本结果。

### 二、研究思路与方法

#### （一）研究目标

CELTS-22 标准研制的目标，是针对高等教育、职业培训以及基础教育领域中的网络课程，提出网络课程质量评价的指标体系和评价指南。本标准所针对的评价对象是作为最终产品或阶段性产品的网络课程，以网络课程产品用于教学或培训的潜在价值为着眼点，评价网络课程资源的内容和功能特性。但本标准不具体规定网络课程产品进入到实际教学和培训活动之后的应用方式或效果。一门合格的网络课程的实际教学效果还取决于对它的具体应用方式以及相关制约环节（如教学安排、学习支持服务、辅助资源等）。

本标准所针对的网络课程是基于 Web 的、以超媒体形式表现的、以异步学习为主的课程，它是基

于 Web 传输的为完成某学科或领域的教学目标而设计、组织起来的相对完整的学习经验体系,它包括围绕特定学习目标而设计组织的学习内容、为实施课程而设计的交互性学习活动、以及为评价课程效果而进行的测评等,是学习内容与学习活动的复合体。本标准针对的是网络课程,它也可以部分地适用于网络课件的评价。网络课件通常是结合某课程中的具体知识点或学习单元而设计的用于训练或教学的网络软件。本标准中的大部分内容适用于网络课件的评价,但单个网络课件一般无法涵盖本标准所规定的网络课程的全部质量特性。

本标准针对的是基于 Web 的课程 (Web-Based Courses),是以 Web 作为主要传输途径而实现的网络课程,而不是以课堂教学或其它途径为主、以 Web 作为辅助工具实现的网络增强性课程 (Web-Enhanced Courses)。对于后者也可以部分地参考本标准进行评价,但一般无法涵盖本标准所规定的网络课程的全部质量特性。

本标准可以作为外部评价使用,即由课程开发者和使用者以外的第三方向开发完成的具体的网络课程作出评价,为选择课程的决策者、网络教育的管理者及课程的使用者提供被评价的网络课程的有效性和质量的依据。本研究制定的标准还可以作为内部评价使用,即由网络课程开发者进行自评,包括在开发过程中针对开发中的网络课程进行形成性评价和针对开发完成的网络课程进行总结性评价。

## (二) 研究方法

### 1、标准化及标准研制过程。

标准化是一个过程,主要是制定标准、实施标准、并使客观事物达到有序化的过程。标准是标准化的产物,是建立在科学技术和实践经验的综合成果的基础上、由一个获得权威机构认证 (Accreditation) 的标准开发组织提出的文件,是规范人类实践活动的有效工具。标准在其生存期 (Life Cycle) 中可以划分为四个阶段。(1) 开发 (Development): 通过文献和采用其他组织的标准建立初步的标准。(2) 建立共识 (Consensus-Building): 标准开发组织与其他组织形成协作和一致,尽快征集有关人员的意见,根据这些意见对初步的标准进行修改,逐步形成工作组草案、委员会草案、标准草案,最终建立通过认证的标准。(3) 维护 (Maintenance): 标准开发组接受询

问,对标准加以解释,在标准试用中搜集错误报告 (Defect Reports),保留反应记录 (Record of Responses),进行校正和技术勘误。(4) 修订 (Review): 根据发展增加新的标准条目,维持没有变化的标准,取消已经不再适用的条目。本研究遵循标准研制的一般程序,目前已经进行了标准的开发和建立共识两个环节的工作。

### 2、评价方法的选择。

教育软件的评价方法主要有四种:分析评价法、指标体系评价法、观察评价法和实验评价法,这些评价方法各有优缺点<sup>[1][2][3]</sup>。网络课程作为特殊的教育软件,这些评价方法也同样适用。经过对各种方法的比较,参照有关研究文献,我们在本标准中采用了指标体系评价法,因为分析评价法受参评人员的主观影响较大,对参评人员的素质要求较高,观察评价法和实验评价法需要投入很多的物力和人力,比较耗费时间,不便于对大量网络课程进行评价。指标体系法是目前国内外各种网络课程(课件)评价标准采用最多的一种评价方法,如美国培训与发展协会 (American Society of Training and Development, ASTD) 的 E-learning 课件认证标准 (The ASTD Institute E-learning Courseware Certification (ECC) Standards)、高级分布式学习 (Advanced Distributed Learning, ADL) 实验室的网络教学的设计与评价指南 (Guidelines for Design and Evaluation of Web-Based Instruction)、网络教学资源的应用评价<sup>[4]</sup>和网络课程绩效评价指标体系<sup>[5]</sup>等,这些标准及相关研究都使用了指

### 3、标准评议方法的选择。

我们采用德尔菲法对所制定的网络课程评价标准进行专家评议,建立共识,以保证标准的合理性、代表性、权威性。德尔菲法的一般流程为:(1) 确定预测问题,选择专家组;(2) 制定与发送第一轮问卷;(3) 回收第一轮问卷,整理结果;(4) 把第一轮问卷的统计结果附在第二轮问卷上发送给第一轮问卷的专家组;(5) 回收第二轮问卷并整理结果;(6) 此时若专家意见差异较大,则进行第三轮、第四轮……问卷调查,在不断的反馈与修改中使专家意见趋于一致而获得最终的预测结果;(7) 调查结果处理。

## 三、研究过程

### (一) 初步的网络课程评价标准

在标准研制过程中,我们对国内外有关网络课

程评价标准的研究做了综合分析<sup>[6]</sup>。基于对国内外大量网络课程的分析,参照ASTD的E-learning课件认证标准、ADL的网络教学设计与评价指南以及其他相关标准和研究文献,我们提出了网络课程评价标准的框架和初步指标体系,包括四个维度:1、课程内容(8条标准):指课程内容本身的学术质量和组织结构,这是决定网络课程质量的核心要素;2、教学设计(17条标准):指对课程的教学目标、教学过程及教学测评方法的合理设计,这是决定网络课程质量的关键,是网络课程区别于一般网络软件的特殊维度;3、界面设计(10条标准):指对学习者与网络课程系统之间的信息交流方式的设计,简便易用的人性化的界面设计是决定网络课程质量的另一个重要因素;4、技术(5条标准):指技术的可靠性、适当性和兼容性,这是网络课程质量的前提和例基本保障。各条标准的描述方式如表1所示。

表1 初步的网络课程评价标准指标体系的描述方式示

标准	描述	属性
1.5 内容分块	按主题把内容逐级划分为合适的学习单元或模块,每个页面集中表达一个主题,每个段落集中表达一个意思。	必需

## (二) 评议专家组的建立

我们邀请了17位专家组成了本标准的评议专家组,这些专家包括:1、国内著名教育技术学专家或曾经在网络课程评价、CAI课件评价上有专门研究的教育技术专业研究人员;2、网络课程开发人员;3、网络学院技术负责人员。评议专家覆盖了北京、上海、山东、重庆、湖南、香港等地区。在评议过程中,专家组成员具有很大的稳定性,为我们获得可靠的结果提供了基本保证。

## (三) 第一轮专家评议

第一轮评议的内容主要包括五个侧面:1、评议各个指标的描述的明确性。即评议各标准描述得是否清楚明确,采用李克特四点量表进行判断(1非常含糊-4非常明确)。2、评议各标准的必要性,即评议各标准是否能体现相应维度的质量特性,是否作为独立的指标纳入网络课程评价标准指标体系。如果这条标准对于可靠、有效的网络课程具有必不可少、不可替代的作用,就选“1,作为必需的标准”;如果这条标准不是必不可少的,但是对可靠、有效的网络课程具有重要的意义,就选“2,作为建议的标准”;如果本标准有一定意义,但不足以作为独立标准,就选“3,不作为独立标准,合并到其它

条目中”;如果本标准意义很小或完全没有必要,就选“4,没有必要,删掉”。3、请专家针对各条标准提出修改意见,填写在“修改建议”一栏中。4、请专家补充我们在各个维度上没有考虑到的指标。5、采用开放式问题,请专家提出针对整个标准体系的意见。

## (四) 第一轮评议结果及标准修订

专家组对第一轮评议问卷中的五个侧面进行了认真评议,提出了自己的意见和修改建议。从总体上说,专家组对整个指标体系的框架没有任何异议,部分专家指出要区分网络课程与网络教学平台之间的联系和差异,考虑哪些功能和特性应由网络教学平台提供,不宜在网络课程中进行评价。

专家组对初步的网络课程评价标准的明确性和必要性的评议结果如图1、图2所示。由图1可见,所有标准的描述都比较明确——描述的明确性等级在比较明确(3)和非常明确(4)之间。

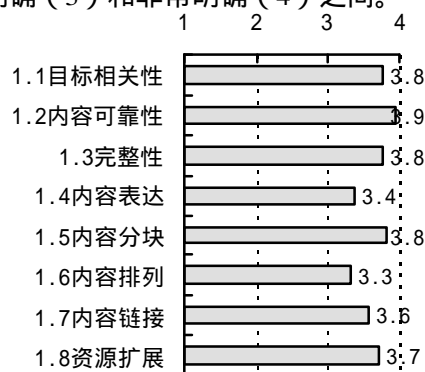


图1a 课程内容维度

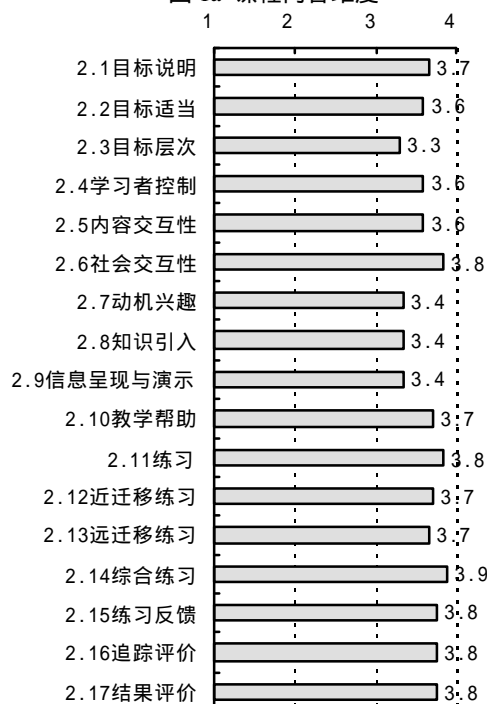


图1b 教学设计维度



图 1c 界面设计维度

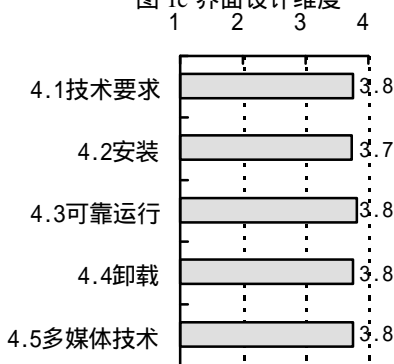


图 1d 技术维度

图 1 第一轮评议：明确性评议结果

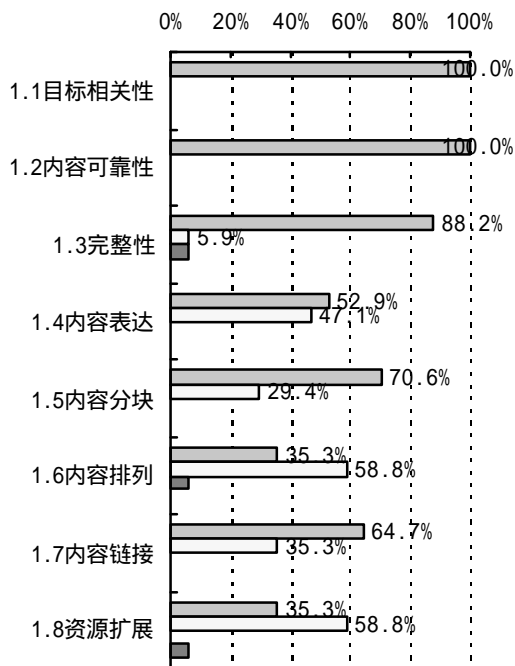


图 2a 课程内容维度

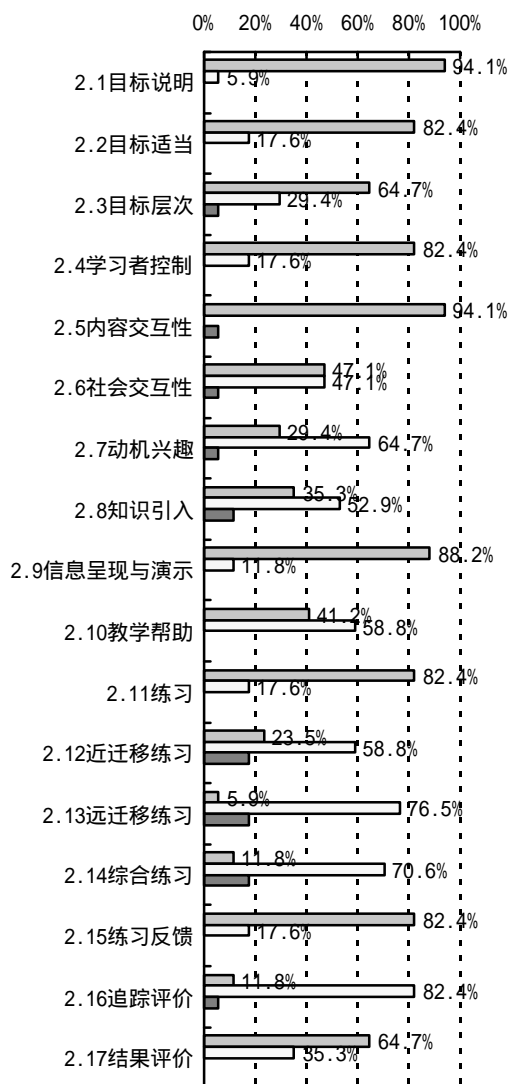


图 2b 教学设计维度

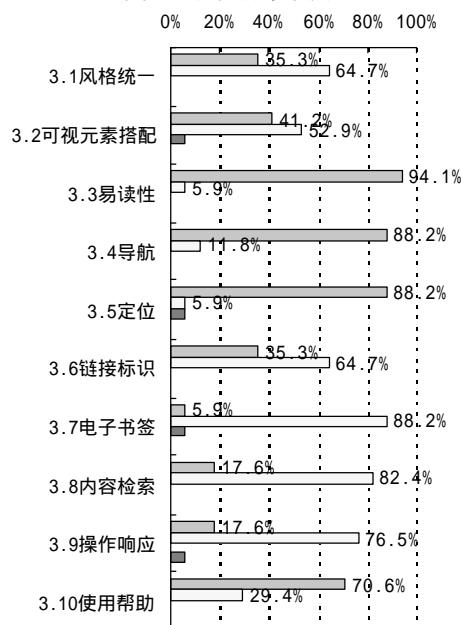


图 2c 界面设计维度

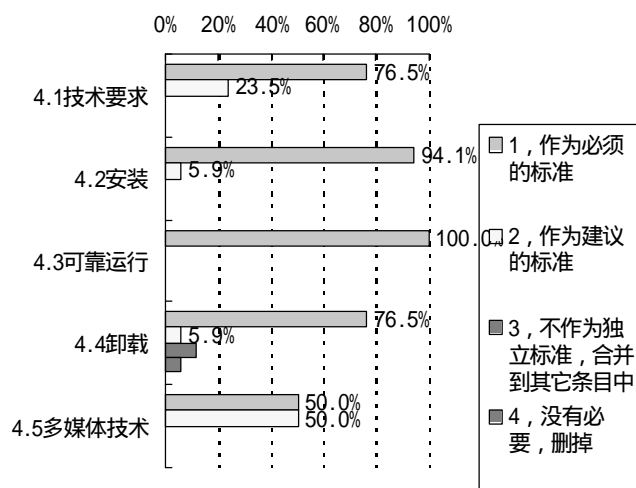


图 2d 技术维度

图 2 第一轮评议：必要性评议结果

在处理专家组对各条标准的必要性评价时，我们使用了“主流意见”的概念。主流意见是指在每条标准的必要性评议意见中选择比例最高的选项，如标准 1.3 的主流意见是“作为必需的标准”。由图 2 可见，除了标准 1.4、1.6、1.8、2.6、2.8、2.10、2.12、3.2 和 4.5，其他标准都具有 60% 以上的主流

意见。

专家组同时提出了在四个维度中需要补充的一些指标。如建议在课程内容维度增加“教学内容的开放性和可扩充性”等；建议在教学设计维度增加“媒体设计的有关内容”等；建议在界面设计维度增加“信息呈现的可控性”等；建议在技术维度增加“可移植性”、“兼容性”等。

根据第一轮评议结果，课题组对初步的网络课程评价标准进行了修订。修订的主要内容是：1、提高标准的明确性。一是增加了对某些标准的举例或具体界定范围；二是修改了部分标准的描述。2、增加评定网络课程等级的“评分标准”栏目。我们按网络课程满足相应指标的程度把每条标准划分为四个评分等级。3、根据关于必要性的专家评议结果，修订了第一轮评议中“暂定的必要性”，在网络课程评价标准中建立“属性”栏目。4、增加、删除、合并了一些标准。调整后标准条目总数为 35 条。经过上述修订的网络课程评价标准的描述格式如表 2 所示。

表 2 修订后网络课程评价标准的格式示例

标准	属性	描述	具体说明	评分标准	备注
1.7 资源扩展	建议	提供与课程内容相关的、有学习价值的外部资源链接。	<ul style="list-style-type: none"> <li>提供足够的相关外部资源链接和参考资料；</li> <li>所提供的资源出处准确无误；</li> <li>所提供的资源有较高的学习价值。</li> </ul>	0, 没有提供任何该课程以外的资源链接和参考资料。 1, 只提供了很少的该课程以外的相关资源链接和参考资料。 2, 提供了一定的外部资源链接和参考资料, 但数量不充分, 或者所提供的有些资源对学习价值不大。 3, 提供了足够的与该课程内容密切相关的外部资源链接和参考资料, 这些资源能够较好地促进和拓展学生的学习。	资源扩展包括在线资源和离线资源, 在线资源主要包括但不限于: <ul style="list-style-type: none"> <li>相关网络课程;</li> <li>专业学科网站;</li> <li>相关网上图书与杂志;</li> <li>著名研究者或研究机构的主页;</li> <li>相关的政府网站, 等等。</li> </ul>

#### (五) 第二轮专家评议结果及修订

在第一轮评议结果及对初步的网络课程评价标准修订的基础上，课题组进行了第二轮专家评议。第二轮评议的内容主要包括三个侧面：一是向各位专家反馈关于第一轮评议必要性的统计数据 and 专家本人在第一轮评议中的选择，请专家在参考其他专家的意见的基础上再次判断每条标准的必要性等级；二是用开放式问题请专家对新增加的各标准的具体评分方法进行评议；三是用开放式问题征求专

家组对标准的总体意见。

总体上专家组对指标体系的框架没有提出修改意见。专家组的意见包括：1、把更多的标准的属性确定为必需标准, 使本标准发挥更好的导向作用。2、建议在第二轮统计结果修订的基础上，组织力量进行一次“试评价实验”，进一步考察本标准的适用性，并作相应改进和完善。

专家组第二轮评议的必要性结果如图 3 所示。除了标准 1.6、2.2、3.2 和 3.5，其他标准都具有 60%

以上的主流意见。

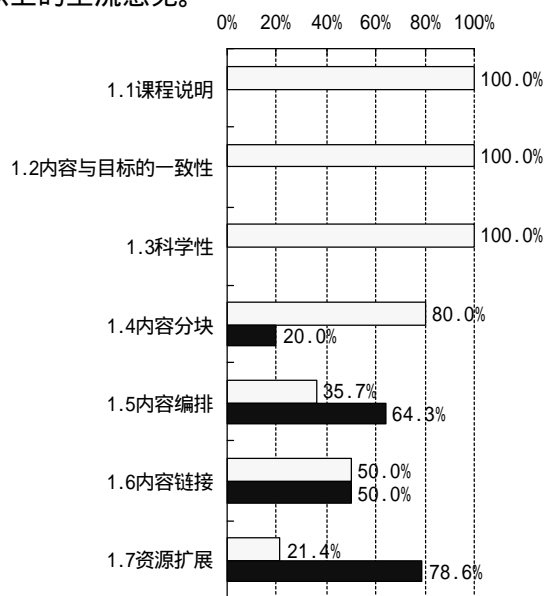


图 3a 第二轮评议课程内容维度必要性结果

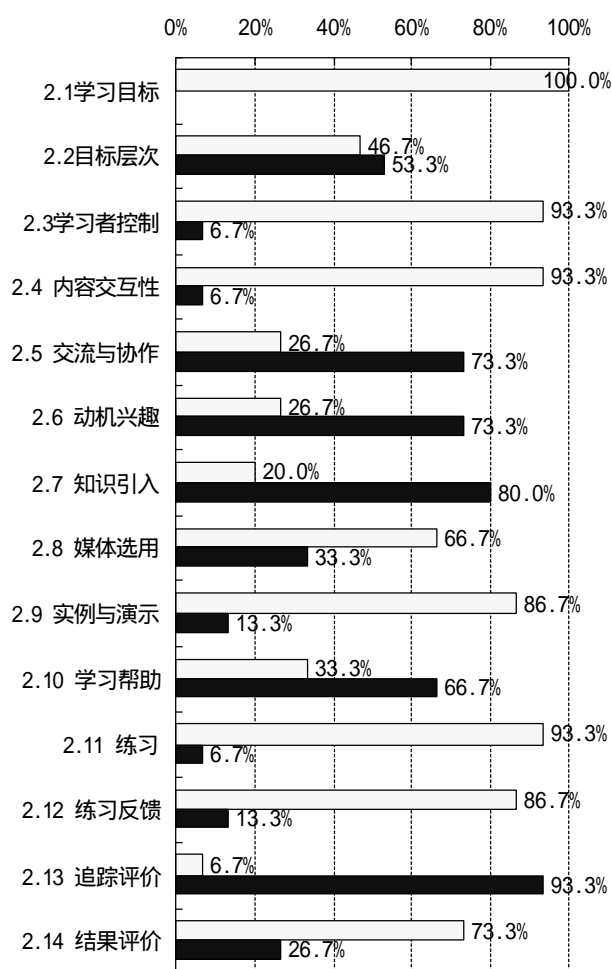


图 3b 第二轮评议教学设计维度必要性结果

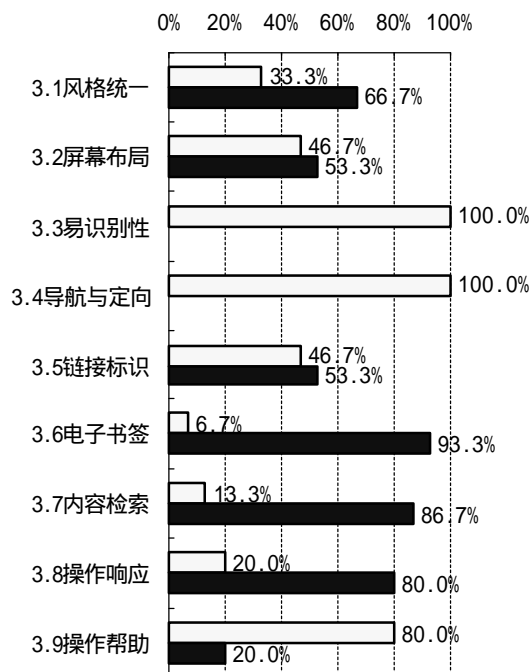


图 3c 第二轮评议界面设计维度必要性结果

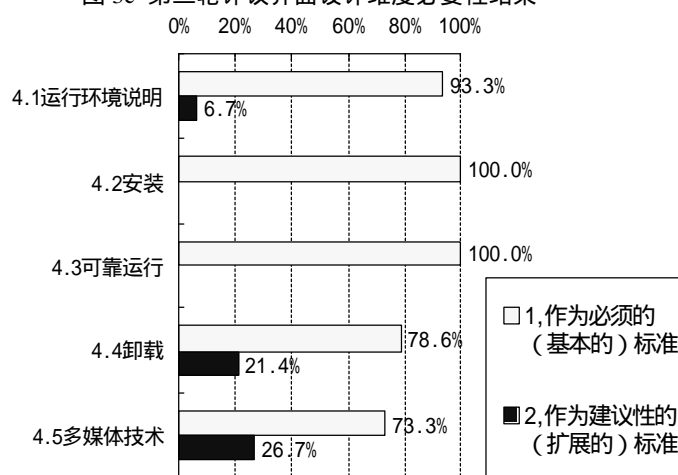


图 3d 第二轮评议技术维度必要性结果

图 3 第二轮评议必要性结果

我们对两轮评议中关于各标准的必要性的主流意见进行了对比，以非参数检验方法（<sup>2</sup>）对主流意见的比例做了差异检验，结果如表 3 所示。经过两轮评议，专家组对指标体系的意见更加集中，除了界面设计维度以外，其他维度在第二轮评议中的主流意见比例都显著高于第一轮评议中的主流意见比例（ $p<0.05$ ）。在总体上，经过第一轮评议后修订的标准在第二轮评议中的主流意见的比例（81.9%）显著高于第一轮评议中的主流意见的比例（73.6%），说明标准基本形成了共识性意见，达到了研究的预期目标。

	课程内容	教学设计	界面设计	技术	总体
第一轮评议	74.3%	73.4%	77.1%	69.4%	73.6%
第二轮评议	81.8%	81.0%	79.3%	89.0%	81.9%
差异检验	$\chi^2=4.34, df=1, p=0.037^*$	$\chi^2=4.16, df=1, p=0.041^*$	$\chi^2=0.24, df=1, p=0.623$	$\chi^2=40.86, df=1, p=0.000^{***}$	$\chi^2=4.34, df=1, p=0.037^*$

注：\*  $p<0.05$ , \*\*  $p<0.01$ , \*\*\*  $p<0.001$

表 3 标准属性的两轮专家评议中持主流意见的平均比例的对比

在第二轮专家评议的基础上，课题组对部分标准的必要性属性进行了修订，形成了《网络课程质量评价规范》(标准草案征求意见稿)（可从 [www.celtsc.edu.cn](http://www.celtsc.edu.cn) 下载），课程内容维度有 7 条指标，教学设计维度有 11 条指标，界面设计维度有 9 条指标，技术维度有 5 条指标，整个规范共包括了 32 条评价指标。

#### 四、结论与展望

按照标准研制过程的要求及各阶段的任务，课题组在标准开发和建立共识阶段初步建立了包含 32 条标准的《网络课程质量评价规范》(标准草案征求意见稿)，具有初步的共识性。在此基础上，我们又对网络课程质量评价过程做了研究，并利用本规范对若干网络课程进行了试评，获得了更多的反馈信息。目前，我们正对本规范进行进一步的完善和后续研究，以便形成工作组标准和正式标准。标准研制是一项需要广泛参与和不断完善的工作，欢迎诸位研究者、课程开发者及课程使用者提出宝贵意见。

致谢：本研究得到了“现代远程教育关键技术与支撑服

务系统和天地网结合项目”和教育部人文社会科学“十五”规划项目的资助。本研究得到了陈丽、丁兴富、高利明、郭文革、黄宝财、刘成新、刘革平、倪友基、尚俊杰、师书恩、王吉庆、熊才平、余胜泉、袁建新、张小真、郑贤义和祝智庭等众多专家的鼎力支持，特此诚挚致谢。

#### 参考文献

- [1] Horton, W. Evaluating e-learning [M]. American Society of Training and Development, 2000
- [2] 杨真，陆达．积极开展教育软件的评价工作[J]．电化教育研究，2001，(1)：51-53
- [3] 师书恩．计算机辅助教学软件的评价[J]．现代教育技术，2001，11(1)：60-63
- [4] 刘成新．网络教学资源的设计、开发与评价[J]．电化教育研究，2000，(3)：27-30
- [5] 熊才平，吴瑞华．网络课程绩效评价指标体系的构建研究[J]．现代教育技术，2001，11(4)：13-17
- [6] 段崇江，张建伟，王学优，滕霞．网络课程评价标准研究述评[z]．昆明：中国电教协会 2002 年会，2002

### The Development of the Standards for Quality Evaluation of the Web-Based Courses

Teng Xia Duan Chongjiang Zhang Jianwei Wang Xueyou

**Abstract** : Following the process and method of standardization, a draft standard for quality evaluation of web-based courses was firstly developed. Delphi method for consensus building was then applied to conduct survey from experts at multiple times. The working draft was further modified through consensus building process. The evaluation strategy, criterias and guideline were established based on the draft standard.

**Keyword**: CELTS-22, Web-based course, evaluation, standard, expert review

责任编辑：务本

#### 域外采风

##### 新加坡进入电子书包时代，学生将扔掉传统书包

新加坡的中学生将不再需要背着沉重的书包，取而代之的是一个储存他们所有课本、笔记本和作业资料的“电子书包”，而它的重量轻于 1 公斤。

由新加坡教育部、国家电脑局和两家科技公司耗资 200 万新元（约合 119 万美元）合作发展的电子书包，实际上是一个手提式电脑，能够让学生储存和取出英文、数学、文学等科目的教材。这意味着中学生将来或许不需要背着沉重的书包上学。这个重量轻于 1 公斤的电子书包，将能储存他们现有的课本、笔记本和作业等资料。学生只须开启电脑，就能够掌握一切资讯。

当地一家中学已经从 2002 年 7 月起开始试用这种电子书包。这家学校校长严正发表示，学校中试用电子书包的 160 多位一年级中学生，都喜欢这种新的多媒体教学方式。教育部计划在 2002 年年底完成有关实施电子书包计划的审查工作，并考虑开始向全国其它学校推广电子书包。

资料来源：佛山教育信息港 <http://www.fsje.net>  
(编辑：李志河)