

以证据为中心的课学习形成性评价设计

段 斌, 杨宇佳, 印 峰, 苏永新

(湘潭大学 信息工程学院, 湖南 湘潭 411105)

摘要: 本文借鉴美国教育评价领域的主流研究方法, 运用“以证据为中心”的教育评价设计模式, 构建了通信工程专业核心课程学习过程的形成性评价框架, 设计了针对特定工程的学习产出评价量规, 有效支撑课程目标和毕业要求指标点, 聚焦学生学习效果。 “信息安全认知调查分析探究性实践任务”案例阐述了以证据为中心的学习成果评价设计方法, 评价产生的任务级反馈和总结性反馈可以判断学生各项能力的长处和短板, 从而有针对性地改进教学。

关键词: 能力; 以证据为中心设计; 形成性评价

中图分类号: G426

文献标识码: A

文章编号: 1008-0686(2019)01-0079-05

Evidence-Centered Design for formative Assessment of Curriculum Learning

DUAN Bin, YANG Yu-jia, YIN Feng, SU Yong-xin

(Xiangtan University, Xiangtan 411105, China)

Abstract: This paper draws lessons from the mainstream research methods in the field of education assessment in the United States, and constructs a formative assessment framework for the learning process of the core curriculum of communication engineering by using the "evidence-centered" educational assessment design model. The assessment gauge of learning output for specific projects is designed, which effectively supports the target points of curriculum and graduation requirements, and focuses on the learning effect of students. The case of "information security cognitive investigation and analysis inquiry practice task" expounds the method of assessment and design of learning achievement based on evidence. Task-level feedback and summative feedback generated by assessment can judge the strengths and weaknesses of students' abilities so that it improves the teaching.

Keywords: competency; evidence-centered design; formative assessment

0 引言

自上世纪末, 欧美各国工程教育认证组织先后改革了工程教育认证标准, 强调了产出导向 OBE (Outcomes-Based Education) 理念, 主要落在学生对 12 条毕业要求的达成^[1~2]。近年来, 我国高等教育亦逐步引入并发展成为具有中国特色的教学模式。仅靠教师讲课的教学模式和仅依据期末考试成

绩考核的评价机制已经无法满足人才培养需求。工程教育以往问责式的总结性评价(即标准化考试)的主角地位正逐渐被能及时调整教学的形成性评价所替代^[3]。有证据表明, 在形成性评价的课堂上, 学生的学习进度明显加快, 学生自主效能的显著提高, 学生学习的效果得到改善^[4~5]。

量规是基于绩效的评价, 与课程或学习标准紧密结合, 能充分运用特定的标准形成多主体、多维度

收稿日期: 2018-02-03; 修回日期: 2018-05-11

第一作者: 段斌(1966-), 男, 博士, 教授, 主要从事通信工程专业信息安全本科课程教学。E-mail: db61850@163.com

评价,适用于多样化学习活动效果的评价^[6]。

以证据为中心 ECD(Evidence-Centered Design) 的设计,是一种集设计、实施、评价和交互于一体的教育评价的整体方法。ECD 认为评价是一种证据推理过程,需根据学习者提供的有限证据对学习进行论证^[7]。

面向产出是 OBE 教育模式中重要的环节,但学习产出是比较抽象的、游离于具体的教学内容之外的概念,由此导致了评价是否聚焦课程目标的达成以及对相应毕业要求指标点的支撑,成为评价学习过程的难点。

为建立课程质量评价的制度,保证专业核心课程开展形成性评价工作的效果,本文在我校开设的“信息安全”课程教学中,对于以复杂工程问题解决为内容的学习,基于以证据为中心的模型设计和开发了“信息安全”课程教学形成性评价系统。

1 形成性评价框架

1.1 能力—证据—任务模型

形成性评价框架主要由三个部分组成:学习者能力模型、认知证据模型和复杂工程问题任务模型,简称为:能力—证据—任务模型,也称为以证据为中心的设计,如图 1 所示。

学习者能力模型(CM 或 SM):代表了学生的知识、技能和能力。以设计的 rubrics 为基础,以工程教育认证标准中的 12 条毕业要求为核心设计学生模型(实例化的能力模型)。

认知证据模型(EM):对学生完成任务情况的观察将为目标能力提供可衡量的证据,它描述了特

定任务表现的特征。

复杂工程问题任务模型(TM):提供一个框架,用于描述和构建学生将与之互动的情境,从而产生证据来推断学生的能力水平。以信息安全课程探究性学习任务为基础,构建任务模型。

2 学习产出评价量规

对于更复杂的、非结构化的问题以及未被广泛认可的答案的问题,使用量规对于评价学生的绩效更为重要^[8]。在证据层中依据能力和任务设计评价量规,一方面在学习后运用量规准确判定能力的达成,形成可衡量的有力证据;另一方面在学习前公布量规可以对学生学习起到导向作用,有效减少学习的盲目性。参考华盛顿协议的界定和标准中 12 条毕业要求,我们设计开发了一款针对“信息安全”课程的“大学生信息安全认知分析”的学习产出量规,有效支撑课程目标和毕业要求指标点,聚焦学生学习效果。

我们将开展“信息安全”课程教学的通信工程专业的大学生分成四个学习小组,向他们发放了大学生信息安全认知问卷,问卷由大学生密码设置倾向问卷和大学生信息安全意识与行为问卷组成^[9],包括 Acxivity(1-4) 4 个子活动,这里我们以 Activity2 中的 Rubric1 为例进行说明。

Rubric1: 专业知识和推理

收集了通信工程 2 组的 108 份大学生信息安全认知问卷,通过 SPSS 统计分析软件对其进行探索性因素分析。

由于进行探索性因素分析需掌握相关的数学和

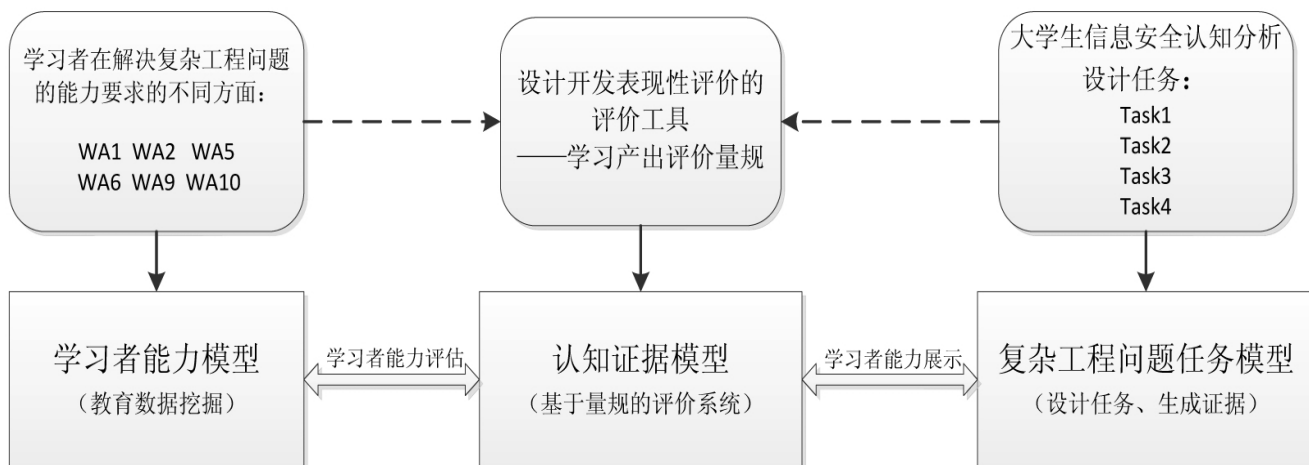


图 1 以证据为中心的形成性评价框架

专业知识: 例如 PCA、因子分析、极大似然法等。因此我们可确定 Rubric1 的评价元素为: 能力指标点 1-1。为了描述评价元素的不同水平, Rubric1 四个等级分别设置为:

A. 熟练运用数学与专业知识对大学生信息安全意识与行为问卷进行了探索性因素分析, 且能进行累积方差解释;

B. 运用数学与专业知识对大学生信息安全意识与行为问卷进行了探索性因素分析, 没有进行累积方差解释;

C. 部分运用数学与专业知识对大学生信息安全意识与行为问卷进行了探索性因素分析;

D. 没有运用数学与专业知识对大学生信息安全意识与行为问卷进行探索性因素分析。

3 “信息安全”课程分析实例

在“信息安全”课程教学过程中, 我们通过基于以证据为中心的课学习形成性评价系统观察和评价学生的学习状态, 发现问题, 及时纠正或帮扶, 并客观判定与毕业要求指标点相关的课程目标的达成情况。

3.1 能力—证据—任务模型及实例化

1) 学习者能力模型

图2代表2015级通信工程专业“信息安全”课程之“大学生信息安全认知分析”的知识、技能和能力方面的变量组成的图谱。表1则是对学习者在解决复杂工程问题的能力要求的不同方面所进行的说明。

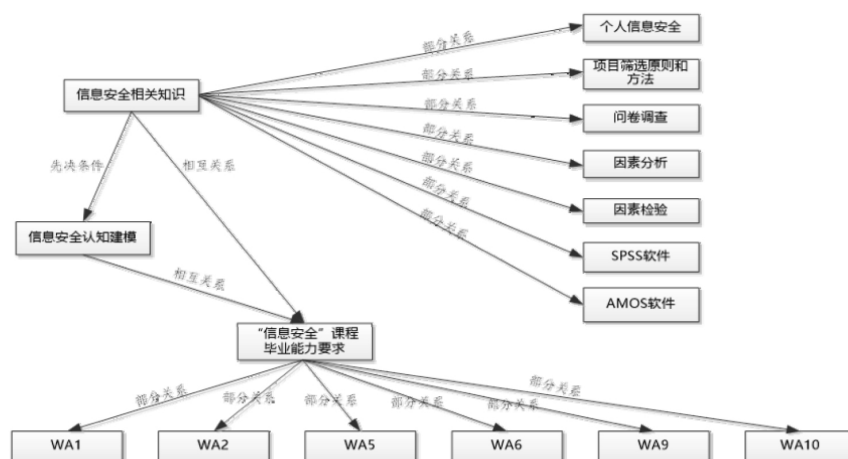


图2 信息安全之“大学生信息安全认知分析”能力概念化模型(图谱)

表1 学习者在解决复杂工程问题的能力要求的不同方面

序号	区别特征	华盛顿协议毕业素质要求的属性简介
WA1	工程知识	应用数学建模、概率论等数学和自然科学知识解决探索性因素分析问题; 对大学生信息安全认知系统的、基于理论的理解; 对不同统计方法进行数学建模; 掌握 t 检验、相关系数法、因子分析法、极大似然法等; 理解在工程实践领域, 采用不同信息处理方法对分析大学生信息安全认知情况的影响。
WA2	问题分析	能够应用统计学、概率论、信息安全等工程原理对大学生信息安全认知分析中包含的子问题进行分析; 依据重要性大、敏感性高、独立性强、确定性好的原则制作结构优良的调查问卷; 以概率论知识为基础, 进行探索性因素分析; 依据探索性结果进行验证性因素分析; 检验大学生密码设置、大学生信息安全意识、行为与人格变量的关系。
WA5	现代工具的使用	对大学生信息安全认知分析任务, 选用 SPSS、AMOS 等现代工具对调查问卷数据进行处理和分析, 并理解其局限性。
WA6	工程与社会	应用信息安全背景知识的论证分析, 评价大学生信息安全认知分析实践和复杂工程问题解决方案对社会、健康、安全、法律和文化问题的影响以及产生的责任。
WA9	个人与团队工作	在大学生信息安全认知分析活动中在各种团队和多学科背景下有效地发挥个人、成员和领导角色的作用。
WA10	沟通	能够在大学生信息安全认知分析中包含的子任务活动中与工程界和全社会以及同行之间进行有效沟通, 给出和接收清晰的指令。

2) 认知证据模型

以每项任务展示后 Rubrics 投票结果为基础,结合世界大学城学生提供的测试结果,设计证据模型。证据模型指示 CM 中不同的技能水平。表 2 列出了几个用于更新工程知识、问题分析和现代工具使用的 CM 变量的证据模型可观察数据。

这些可观察数据用来评价学生的知识、技能和能力,一方面教师可以通过这些数据使用更加详细和及时的方式即任务级反馈的形式进行教学;另一方面这些数据为学生和教师提供总结性反馈。

表 2 证据模型中观测值的例子

专业知识和推理	量化推理和解决问题能力	现代工具的使用
结果正确性 t 检验 相关系数法 因子分析法 极大似然法	过程正确性 结果正确性 项目筛选方案 数据分析原则 因素分析方法	过程正确性 SPSS 使用方法 AMOS 使用方法 结果正确性 数据处理方案

3) 复杂工程问题任务模型

识别了学生可能展示的证据后,我们将此项单个学习活动具体设计分配如表 3 所示,用于搭建复杂工程问题任务模型。

表 3 “大学生信息安全认知分析”子任务学习者分配表

学习者	具体活动
通信工程 1 组	Activity 1: 问卷制作
通信工程 2 组	Activity 2: 探索性因素分析
通信工程 3 组	Activity 3: 验证性因素分析
通信工程 4 组	Activity 4: 回归分析

在尝试解决困难的设计任务过程中给予学生的实际任务级反馈:

- (1) 检查问卷的有效性
- (2) 消除不相关数据之间的不平等
- (3) 注意基于实际情况处理实验结果
- (4) 命名因素的简单化

3.2 基于能力-证据-任务模型的形成性反馈

在教师主持下,全体学生按照设计的评价量规(Rubrics)对课堂展示任务完成情况的学生表现出的能力进行集体评价。本文选取通信 2 组的评价情况,并对本组代表王建辉同学表现出的部分能力情况进行案例说明。

1) 任务级反馈

此时的任务级反馈不以课堂听课为主,也不仅仅关注教师的课堂表现,而是以学生表现为主,关注学生在课堂上对任务完成情况的展示,明确了质量

监控与毕业要求达成的关联。

下面以第 4 项任务级反馈“命名因素的简单化”为例进行阐述。

采用 SPSS 软件对大学生信息安全意识与行为问卷进行探索性因素分析。

本组学生经探究后发现,问卷问题答案的多元化导致通信 1 组制作的问题不能良好的体现填表人的想法。因此除了归一化处理之外,本组额外做了二值化处理,将问题转化成了肯定性问题;经实验发现二值化结果与原先结果基本一致,但却方便了命名因素的处理。

最后该组学生得到的结果:参照碎石图确定抽取的因素数目,分析表明大学生密码设置倾向是一个四因素的结构,方差解释率达到 71.982%,因此大学生信息安全意识与行为问卷有较好的构想效度。表明该组学生完成任务情况良好。

2) 总结性反馈

当结束这一阶段所有的信息安全课程任务之后,基于之前的各项数据,对学生所有的能力展示做出总结性评价。此时的总结性评价是准确、真实并且有迹可循的,能聚焦到课程目标的达成以及对相应毕业要求指标点的支撑。

通信 2 组学生王建辉能力评价得分如表 4 所示,表中数据为票数。

表 4 通信 2 组的学生王建辉能力评价得分表

评价元素 元素等级	Rubric1	Rubric2	Rubric3	Rubric4	Rubric5	Rubric6
A	89	101	80	41	42	21
B	16	7	21	45	56	85
C	3	0	7	22	10	0
D	0	0	0	0	0	2

建立成绩-能力的转换模型:

A 每票分值为 $Q_A = \frac{1}{108} = 0.0093$ 分; B 每票分值为; $Q_B = \frac{0.5}{108} = 0.0047$ 分; C 每票分值为 $Q_C = \frac{0.3}{108} = 0.0028$ 分; D 每票分值为 $Q_D = \frac{0}{108} = 0$ 分

能力指标点 1-1 权值 $\gamma_1 = 0.15$; 能力指标点 2-4 权值 $\gamma_2 = 0.25$; 能力指标点 5-1 权值 $\gamma_3 = 0.15$; 能力指标点 6-3 权值 $\gamma_4 = 0.1$; 能力指标点 9 权值 $\gamma_5 = 0.2$; 能力指标点 10-2 权值 $\gamma_6 = 0.15$

Rubric 得分为 S_1 权值为 0.7; 网络学习空间得

分为 S_2 , 权值为 0.3; i 为各指标点。

$$\text{第 } i \text{ 项能力指标点的达成值: } \alpha_i = \sum_j S_j \rho_j \quad (1)$$

$$\text{Activity 的 Rubric 达成值: } \alpha_a = \sum_i \alpha_i \gamma_i \quad (2)$$

$$\text{Activity 的学习空间达成值: } \beta_a = \frac{S_2}{100} \times 100\% \quad (3)$$

$$\text{Activity 的能力达成值: } A = 0.7\alpha_a + 0.3\beta_a \quad (4)$$

网络学习空间的证据和评分见图 3。将数据输入式(1)算出学生王建辉的活动能力达成值,如表 5 所示。



[已提交]
2018-06-27 20:16

王建辉 答:

1、选取4个成分是因为碎石图的关键就是找拐点,也就是找图中陡坡和缓坡的临界点,其本质就是找出特征值明显较大的因子,而且特征值大其方差解释率也大,由图可见显然前面4个点符合。

2、由图二结果可见,其区别在于项目、成分的不一样,载荷矩阵数值不一样;通过这次对比结果,可见分析社会实际问题时应实事求是,因人而异,不能盲目借鉴;同时结果显示较多使用密码与自身相关性较大,总体而言是不安全的,容易发生密码泄露,因此我们在信息时代应该注意个人信息安全,密码等关键性信息应该设置得较复杂难以破解;最后对比项目,其问题不一致的,但是分析出的结果又一定的一致性,这告诉我们不同地区文化存在差异性,但其本质不变性,我们通信人在处理社会问题应该是入乡随俗,尊重当地文化的基础上为社会提供更好的服务。

[已评分]

评分: 90

评语: 该同学的报告提交完整,符合主题,分析较为透彻,较好的完成能力指标点的要求。

图3 网络学习空间作业提交和评分界面

表5 通信2组的学生王建辉的能力指标点达成值

任务	指标点	评价得分				达成值
		A	B	C	D	
基于问卷调查数据,通过SPSS统计分析软件分析考察大学生密码设置倾向和大学生信息安全意识与行为的结构和特征。	1-1	0.827	0.075	0.008	0	0.910
	2-4	0.939	0.033	0	0	0.972
	5-1	0.744	0.099	0.019	0	0.863
	6-3	0.381	0.212	0.062	0	0.655
	9-1 9-2 9-3	0.391	0.263	0.028	0	0.682
	10-2	0.195	0.399	0	0	0.595

由式(2)得出 Rubric 的达成值 $\alpha = 0.800$

由式(3)得出网络学习空间达成值 $\beta = 0.9$

由式(4)得出该学生的最终能力达成值 $A = 0.83$

下面举例对表中能力达成值中的部分结果进行总结性评价解释:

(1) 学生王建辉在能力指标点 2-4 上得分很高,表明他在能力指标点 2-4 这一方面表现优异,这是基于他能用数学、工程基础和专业知识对问卷调查结果进行推演和正确性分析以获得有效结论。

(2) 在学生现场评分环节,能力指标点 10-2 得分为 B。这是因为该学生在对通信 1 组制作的问卷进行修正后并没有同他们进行沟通交流,最终呈现出的问卷依旧是有问题的问卷,这影响了整个任务的顺利完成。沟通的能力上得分较低,而这种沟通能力在许多工程实践中是很重要的,缺少沟通的

后果很有可能会导致整项工程的失败。因此该学生的此项指标点的达成值偏低。

经过形成性评价,可判断学生各项能力的长处和短板,并扬长避短。用形成性评价获得的反馈数据来了解教学活动中学生能力状态的变化,进一步支持学生的学习,为专业教学的持续改进提供依据。

4 结语

传统的教学模式已经不能满足社会对高校人才的需求,因此如何在课程中开展形成性评价和持续改进模式是实施工程教育认证亟待解决的难点问题。培养创新型人才,如何将形成性评价发展成为以证据为中心的评价体系,本文以“信息安全”课程的探究性教学内容“信息安全认知分析”学习成果评价为例,基于能力—证据—任务模型做了有益的探索。为进一步完善该评价设计,还需在今后大胆尝试新的技术和方法,形成良好的认知诊断、收集证据、评价反馈学生学习状态和能力达成度,突破工程教育界的固有观念和模式。

参考文献:

- [1] 顾佩华,胡文龙,林鹏,等. 基于“学习产出”(OBE)的工程教育模式——汕头大学的实践与探索[J]. 高等工程教育研究, 2014(1): 27-37.
- [2] 林健. 如何理解和解决复杂工程问题——基于《华盛顿协议》的界定和要求[J]. 高等工程教育研究, 2016(5): 17-26.
- [3] 何克抗. 关于形成性评估与隐性性评估——美国《教育传播与技术研究手册(第四版)》让我们深受启发的亮点之三[J]. 中国电化教育, 2017(6): 24-29.
- [4] Scott D. Globalisation Mechanisms [M]// New Perspectives on Curriculum, Learning and Assessment. Springer International Publishing, 2016.
- [5] Shute V J, Kim Y J. Formative and Stealth Assessment [M]// Handbook of Research on Educational Communications and Technology. Springer New York, 2014: 311-321.
- [6] 彭熠. 量规在教师教育中的作用与设计原则[J]. 大学教育科学, 2014(3): 55-60.
- [7] Mislevy R J, Haertel G, Riconscente M, et al. Assessing Model-Based Reasoning using Evidence-Centered Design: A Suite of Research-Based Design Patterns [J]. 2017.
- [8] Jonassen D H. Assessing Problem Solving [M]. Springer New York, 2014.
- [9] 张敏,郝素冰,龚子捷,等. 大学生信息安全认知的测量及其与人格特征的关系[J]. 高等工程教育研究, 2016(6): 144-148.