

一种全新的尝试:高等教育 质量测评的国际动向

——OECD“高等教育学习成果测评”的设计与实施

陈 涛^{1,2}

(1.厦门大学教育研究院,福建厦门 361005 2.比利时鲁汶天主教大学心理与教育科学学院)

[摘要] 经济合作与发展组织历时5年研发了一套旨在测评高等院校本科毕业生学习成果的国际比较项目——“高等教育学习成果测评”。从“提出设想”到“面对挑战”,从“研发工具”到“投入实践”,该项目在测评理念、测评方式、测评内容和技术方面突破了传统质量评估的学院模式。作为一种“以学生为中心”的综合性的国际测评方式,无疑具有重要的现实意义和拓展价值。

[关键词] “高等教育学习成果测评”;高等教育质量;经济合作与发展组织

中图分类号:G649 文献标识码:A 文章编号:1003-7667(2015)02-0030-08

长久以来,教师都在争论大学毕业生应该掌握什么样的知识和技能。^[1]事实上,这个看似颇有争议的话题却是大学人才培养的基本命题,更是体现大学核心职能履行与否的关键所在。因此,考察大学的学习成果已经成为当今大学、政府乃至国际组织衡量高等教育质量的最新手段。然而,“如何衡量”成为其问题的症结。经济合作与发展组织(OECD)历时五年,研发并实施了一套旨在测评大学本科(三或四年学制末期)学习成果的国际比较项目——“高等教育学习成果测评”(Assessment of Higher Education Learning Outcomes, AHELO)。目前,已有17个国家和地区、248个高等教育机构、4,900名教师和23,000名学生参与其中。^[2]尽管从一开始,项目研发就存有争议,但是这丝毫不能掩盖一个事实:学习成果的确是意义教育之关键,亦是预示诊断和提高教学过程和学生学习之必备。^[3]

本文系教育部重大攻关课题“我国本科人才培养质量研究”(课题编号:10JZD0038)的成果。

作者简介:陈涛,男,国家留学基金委资助厦门大学教育研究院与比利时鲁汶天主教大学心理与教育科学学院联合培养博士研究生。

一、背景与趋势:高等教育学习成果测评的合理性

AHELO的研发并非是偶然的创想,而是随着高等教育发展之势应运而生。

近四十年来,世界高等教育发生了巨大变化,规模扩张便是重要的表现。数字激增的背后反映出学生背景和需求的多元化,新型高等教育机构的出现恰恰填补了这些缺口,随之办学质量日益成为关注的焦点。这些外在变化引发了高等教育范式的内在转变,“以学生为中心”的教育理念必然带来质量评估方式的革新。

(一)扩张性与多样性:世界高等教育发展之走势

1.世界高等教育规模的急剧扩张

从世界范围来看,规模扩张已经成为近五十年来高等教育发展最为显著的特征。根据联合国

教科文组织有关高等教育入学的统计,从1970年3 250万人到2010年17 800万人,年度平均增幅为4.3%。特别是1990年以后,这一扩张势头呈现加速状态,年平均增幅上升至5.9%,预计2025年将达26 300万人。^[4]高等教育规模扩张已蔓延至世界各地,20世纪60年代美国和加拿大率先进入大众高等教育阶段,80年代西欧和日本也完成了高等教育的大众化,^[5]如今东南亚和非洲大陆的扩张之势也愈加明显。数据刷新背后蕴含了多种因素:公共需求、技术革新、城市化进程和民主化浪潮等,但这都指向一个不争的事实:高等教育系统急剧扩张。

2. 新型高等教育机构的出现

为了迎合高等教育大众化、学习者需求多样化,新型高等教育机构如雨后春笋般涌现而出。然而,这些迅速增长的机构基本上是紧密联系当地劳动力市场需求、雇主导向的非大学部门,可以说这是高等教育体系近年来最具重大意义的结构性变革,^[6]如法国的技术科技大学(IUTs)、澳大利亚的技术和继续教育学院(TAFE)、德国的应用科学大学,以及美国职业导向的初级学院等。事实上,新型高等教育机构的出现不仅是办学类型、产权和教育课程的多样化,而且它们的使命、社群目标(女性、少数族裔、特殊需求的学生、成人学习者、国际学生等)、服务当地之需、专攻领域,以及与特定产业和公司的合作关系也更加与众不同。^[7]

(二)透明性与责任性:高等教育范式变迁之释义

1. 超越学院式的测评方式

高等教育系统的扩张与突变使其测评方式也发生了转变——从学者共同体水平治理的学院式转变为具有自上而下自治特征的新模式,比如:增加了透明性和问责性。^[8]就传统意义而言,学习成果测评是高等教育机构的内部事务,但当前一些迹象表明,传统的学院式测评方式已经不能满足现有需求,因为内部测评更多是量身剪裁,缺乏外部监测。这并不是说地方性的评估不重要,但作为一个补充,更富综合性的测评形式(常规考试、研究生申请测试,以及专业的证书考试等)的出现

已是大势所趋。^[9]AHELO的研发理念反映的正是目前这类综合性测评最先进的表现形式。

2. 日益关注学生学习成果

高等教育范式的第二大转变是从基于“输入”的概念(授课数、学时数和工作量)转变为基于“输出”的概念。这一转变的有力证明来自欧洲的《博洛尼亚宣言》(the Bologna Declaration),其明文规定2020年高等教育学习成果模块和项目都要在欧洲高等教育区(EHEA)内得到认可。^[10]因此,47个成员国和许多外围国家都努力使其高等教育体系与博洛尼亚进程(Bologna Process)相匹配,旨在促进资格描述、学位互认和学生流动。^[11]为了巩固高等教育区所建立的成果,欧洲的部长们重申了学习成果方式实施的重要性,而且将其作为欧洲学分转换系统(ECTS)、文凭补充(DS)、认可、资格框架和质量保证成功实施的关键。^[12]

3. 以学生为中心的教学过程

从“讲授范式”(instruction paradigm)到“学习范式”(learning paradigm)可以看作是高等教育范式的第三大转变,教师应该意识到自己的任务并非只是讲授,而是和每一个学生共同努力创造学习。^[13]如今,不论是在欧美国家,还是在亚太地区,将学生视为高等教育中心已基本形成共识。尽管在观念、实践中还存在诸多困难,但这的确提出了一个新的理念——增强学生的学习成果。相关研究已经证实,以学生为中心的项目可以提高学生的学习成果。就此而言,AHELO的研发通过提供有关学习成果的客观测试以及大量的语境和背景信息,拓展了一系列新的研究途径,这与“以学生为中心”的高等教育新范式相一致。^[14]

二、挑战与抉择:高等教育学习成果测评的可行性

OECD研发AHELO面临着严峻的挑战,由于高等教育机构的特殊性,较之PISA更为艰难。然而,不论是“质疑之声”,还是“操作之困”,从事AHELO研发的专家和工作人员始终秉持他们的理念——通过测评学生学习成果来提高高等教育质量。由此,一年之内3次专家会议的相继召开为

PISA(Programme for International Student Assessment) 国际学生评估项目,OECD研发的一套旨在评估接近15岁完成基础教育学生的知识和技能项目。

此工作绘制了路线图。

(一)质疑之声:排名的风险与结果的误用

自 OECD 提出 AHELO 构想以来,就遭到高等教育界的种种质疑。其中,最为担忧的是政策制定者很有可能利用 AHELO 的数据进行排名,最终导致测评结果的误用,尽管 OECD 从未有开发排名系统的意图。^[15]然而,各国政府的财政支持至少表明政府认为这是一个有用的工具,院校领导及组织担心政府的努力是将其结果作为政策制定的参照而非分析和改进的依据。^[16]AHELO 的数据结果很可能会扭曲公共政策,致使教育资源重新分配、鼓励精英教育和旗舰型大学甚至重组高等教育系统,这与 AHELO 加强高等教育质量的目标是背道而驰的。^[17]

(二)操作之困:对比的复杂与标准化测试

由于高等教育机构的使命、属性和学生主体的多样性,导致了公平比较的复杂性,特别是跨文化、跨地区的多国比较更是困难重重。针对院校的多变性,批评者认为用一套标准化的测试以求做到“一体通用”(one-size-fits-all)是不切实际的。^[18]然而,教育国际组织(Education International, EI)认为,标准化测试将在教育背景、资源、项目和教育过程方面提供有用的信息。^[19]其实,最令高等教育机构不安的是,AHELO 会妨碍大学自治、学术自由,甚至会引发高等教育趋同化,这是最为潜在的影响。^[20]但教育国际组织相信该项目不会削弱,反而会提升学术自治。^[21]

(三)绘制路标:从“华盛顿会议”到“巴黎会议”再到“首尔会议”

2007年4月,首次专家会议在美国华盛顿举行,会议主要考虑了 AHELO 的有用性和需求性问题。专家们认为研发高等教育学习成果测评的国际性比较虽然困难重重,但他们都肯定可靠的学习成果信息将会增强高等教育多样化和国际化深入发展的重要性。^[22]同年7月,在法国巴黎召开了第二次专家会议,专家们主要确定了 AHELO 的概念性问题,如 AHELO 研究如何组织、政策问题如何应对。专家们回顾了已有的高等教育学习成果测评的方法,并试图探索国际比较背景下测评工作的可行性研究设计,特别是其透明性与参与方式的设计。^[23]同年10月,在韩国首尔举行了第三次

专家会议,这次会议重点讨论了可行性研究的设计与实施工作及研究框架,特别就研究的科学性、有效性、连续性和国际性以及样本选择、组件选定、时间选设等做了更为细致的安排。^[24]

三、设计与管理:高等教育学习成果测评的概念性

AHELO 研发工作通过多方“证据”(如已有测评工具、相关资料等)建构 AHELO,并确定了测评范围和样本,提出了4项测评组件(distinct strands)概念,使其抽象的研发设想有了有形的抓手。

(一)测评范畴:4项迥异却连贯的测评组件

AHELO 由4项不同的测评组件构成,这4项组件能够单独实施且具有一定的关联性。前3项组件来自3个不同的领域,即“通用技能”组件(Generic Skills Strand)(主要涉及批判性思维、分析推理、解决问题和书面表达能力)、“经济学”组件(Economic Strand)和“工程学”组件(Engineering Strand)。最后1项测评组件聚焦增值测量问题(Value-Added Measurement Strand),但它并非直接涉及学习成果测评和增加值。相反,它是在可行性研究的范畴内,通过方法论的途径,增值策略的数据需求以及不同途径利弊的三重对照下,以研究为视角处理增值测量问题。^[25]前3项测评组件可以单独实施,不同国家可以任选其一。总之,为了使不同组件间的配合成效最大化、使沟通简单化以及形成经济规模,在整个过程、测试管理和分析方面,可行性研究要求4项组件都须连贯地实施。^[26]

(二)测评样本:基于多因素均衡分布的国家

为了确保每项组件国际化变量的充足,OECD 将以下3种要素作为其样本选择的标准:一是地理分布,理想的国家样本应来自世界五大洲:非洲、美洲、亚洲、欧洲和大洋洲;二是语言,理想的样本国家至少包含一个英语语言国家、一个亚洲语言国家、一个拉丁语语言国家和一个欧洲语言国家;三是文化背景,理想的集合应是融合拉丁、北欧、欧洲、盎格鲁—撒克逊以及亚洲文化。^[27]事实上,参与国的均衡分布覆盖了所有操作组件项目,以确保展现一幅合理而均衡的图景。目前,共有17个国家或地区决定参加这一国际化的测评项目。它们分别是:参与“通用技能”组件项目的哥伦比

亚、埃及、芬兰、韩国、科威特、墨西哥、挪威、斯洛伐克和美国(康涅狄格州、密苏里州、宾夕法尼亚州)9个国家,参与“经济学”组件项目的比利时(弗兰德)、埃及、意大利、墨西哥、荷兰、俄罗斯和斯洛伐克7个国家,参与“工程学”组件项目的阿布扎比、澳大利亚、加拿大(安大略)、哥伦比亚、埃及、日本、墨西哥、俄罗斯和斯洛伐克9个国家。

(三)测评阶段 测评工具开发与试验的效度

1.第一阶段(2010.1~2011.6):工具开发与设想的最初证明

AHELO 首先聚焦在达成一种可行性——设计能够反映不同国家、语言、文化和院校背景且具有足够效度的测评框架和工具。第一步:适应、开发临时性的测评框架和适用于国际化背景的工具;第二步:通过参与国小规模测试以获取跨语言 and 文化的效度,并证实测评工具的有效性,^[28]相继开发通用技能工具(2010)、学科工具(2010),以及背景维度工具(2011)。^[29]

2.第二阶段(2011.7~2012.12):可用工具与结果的心理测量分析

该阶段旨在通过致力于测评学生学习成果的可用性以评价 AHELO 的科学性和实用性。不同的高等教育机构负责测评工具和背景调查,探索最佳途径以吸引、激发管理者、教师和学生参与其中,并建立一种解释性的学习成果情境,从而帮助确定导致学习成果强化的多种因素。该阶段既涉及实施、测试管理和学生分数,还包括数据分析及 AHELO 科学性和实用性的报告与评价。^[30]

3.第三阶段(2012.3~2013.3):增值测量的方法论与途径

最后阶段工作主要是方法论和途径的探索以获取“增值”,如高等教育机构对学生学习成果的贡献或“学习收获”。增值测量组件的目的并非是测量“增值”,而是检验、分析获取高等教育机构“学习收获”的合理方法。因此,研发者回顾了有关增值测量途径的文献,建构了能运用于学校层面的相似操作模式,接着,负责增值测量的专家就方法论提供专门的指导方案,最终将之运用于主体测试中。^[31]

(四)测评管理:可行性研究的管理与行动者

AHELO 可行性研究的设计与实施牵涉多方

协作:一是国际性管理,由 OECD 秘书处和 AHELO 联盟组成,主要负责操作和分析,特别是对整个测评项目的监管;^[32]二是国家和院校的协调,参与国需要建立国家中心,并且在联盟的指导下设立国家工程经理(NPMs),旨在联络联盟,确保 AHELO 与规定的技术标准和操作指南相一致。参与院校则需要任命院校协调员(ICs),负责联络 NPMs,并协助他们选取测评样本、提供院校信息,以及组织院校层面的测评管理;^[33]三是国际专家组,专家组主要分享专业知识,提供建议,主要包括:背景维度的专家组(为背景工具开发概念和分析框架)、技术检验组 TRP(检验技术方案)和技术咨询组 TAG(为工具研发、翻译、效度、成绩、评价等提供建议)。^[34]此外,教育政策委员会(EDPC)和高等教育院校管理董事会(IMHEGB)也起到了强化保障的作用。

四、工具与样本:高等教育学习成果测评的操作性

在 AHELO 的操作过程中,有三项工作极其重要:一是测评框架和工具的开发;二是测评样本的选择;三是学生的回复率。每一项工作都需要精心的设计、规范的操作和严格的管理,否则 AHELO 在失去意义的同时,还会带来严重的错误。因此,从某种角度而言,AHELO 项目的关键之处就在于它科学的操作性。

(一)测评的基本流程

1.开发测评框架

测评框架旨在建构测评目的、概述一致定义,且提供测评构想与测试结果的联系。良好的测评框架应包含:一个清晰的测评定义、项目及其类型的描述以及解释结果的基础。AHELO 除了三项测评组件框架外,还涉及背景维度框架,共分四个阶段:一是现有资料的审核;二是检验基于已有资料的框架草案版;三是形成国家咨询以求证实;四是在咨询结果的基础上进行测评框架的最终检验。^[35]

2.开发测评工具

该步骤是要创设与规格目录相匹配的选项,具体内容一旦确定,测评专家就需要对拟定工具进行检验,以保证每一道问题与测试的相关性,并且所有问题都要反映出规格目录和测评框架。AH-

ELO 涉及三项测评组件工具和三项所采集的背景信息(学生、教师和院校背景)调查工具。第一种题型是结构型的问答题,第二种题型是选择题。以上题目除了科威特是纸质答题外,其余都是电子答题。^[36]

3.测评工具的翻译和调适

翻译和调适(地方化)在多语言的比较中格外重要,由于文化偏见和课程差异,如果翻译不当,问题就会频出。^[37]翻译指南引导翻译和调适的整个过程,以确保工具在不同语言中都能忠实于英语源版本。^[38]翻译团队一方面进行初试以预估可能的歧义、文化和可译性问题;另一方面,通过《翻译和调适指南》和“文件管理系统”以保证翻译的一致性,使每一个国家的参与者都有共同的认识和理解。^[39]

4.测评工具的小规模效度

该环节的任务称为“前测”或“试测”,即通过选取与目标群体及测试环境具有共同特征的学生,一般至少200人。经过测试得出的数据将是最终测评工具筛选与修改的基础,同时它也能收集一些师生对测评工具的反馈意见,如测试长度、难度及相关性。此外,该过程还包含认知实验程序,有助于检验学生的思维过程。^[40]

5.测评工具的最终检验

测评工具的最终检验来自于小规模的前测活动、试测者的意见反馈,以及利益相关者的咨询建议。尽管以上步骤酝酿出测评工具的最终版本,但是只有定量分析的内容决定着最终要用于解释结果的测评问题。虽然整个研发过程谨小慎微,但一些测评问题仍不能从数据分析中得到预期的处理,其结果也无助于测评。^[41]

(二)测评框架与工具

就“通用技能”组件而言,其测评工具改编自“美国大学学习评价”(CLA)工具中的结构型问答部分,修改后的CLA绩效任务(PTs)适用于国际化背景中通用技能测试。除了结构型问答任务(CRTs,亦称为PTs)以外,由澳大利亚教育研究委员会(ACER)研发的“课程综合能力测试——选择题(MCQs)是其CRTs的重要补充。该测评框架是建立在一致理解基础上的核心综合能力的概念化,其测评工具涉及2个GRTs,旨在考察学生的批判性思维(90分钟);若干MCQs,旨在处理

更为广泛的话题和多种思维类型,但不涉及综合思维和写作能力的测试(30分钟)。该项测评组件共有参与国家9个,参与院校98所,参与学生10,657人。^[42]

就“学科—专业”组件而言,分为经济学和工程学两类,每项测评仍包含以上两部分内容(CRTs和MCQs)。美国教育测试中心(ETS)负责经济学组件的研发工作,其测评框架重在考察学生的抽象思维、分析、归纳和演绎、量化与设计,以及问题架构能力。^[43]而AHELO联盟,特别是ACER、日本国家教育政策研究所(NIER)以及意大利佛罗伦萨大学则承担了工程学组件框架和工具的研发任务,测评工具仅限土木工程专业,CRTs更多是提供一个简单的数字或短文回答。^[44]

“经济学”组件项目的测评框架共包含五项内容:(1)学科知识和认识;(2)学科知识及其在现实问题中的应用;(3)有效利用相关数据和定量方法的能力;(4)与专业人士和非专业人士的沟通能力;(5)获取独立学习技能的能力。^[45]它的测评工具包括了1个GRTs,旨在考察学生的综合技能和沟通能力(30分钟);48个MCQs(分四组,每组12题),旨在考察学生的经济知识、认识和技能(60分钟)。^[46]该项测评组件共有参与国家7个,参与院校60所,参与学生6,000人。^[47]

“工程学”组件项目的测评框架则是基于运用基础工程和科学原理以及工程流程与基本技能解决问题的能力。^[48]它的测评工具包括了两个GRTs,主要考察学生的四项关键能力:工程基本技能、工程分析、工程设计和工程实践(30分钟);40个MCQs(分四组,每组10题),旨在考察学生的基本工程科学知识(60分钟)。^[49]该项测评组件同样有参与国家7个,参与院校70所,参与学生6,000人。^[50]

就“背景调查”而言,来自荷兰高等教育政策研究中心(CHEPS)和美国中学后研究中心(CPR)的专家负责其框架和工具的设计工作。^[51]背景维度测评框架(the Contextual Dimension Assessment Framework)作为AHELO背景维度的重要部分,一是通过在不同国家和院校的有效方式,帮助评价获取背景变量的可行性;二是帮助检验心理测量分析以证实长期发展的背景变量,且证明AHELO院校

改进的分析潜力;三是帮助管理样本、数据收集以及质量监控。^[52]

该测评框架基于四项内容:反映参与国的看法和首席专家的想法、考虑目标群体的特征、确定在背景调查中测量的结构和反映结构和范围的相对重要性。测评工具分为:学生背景工具(SCI)、教师背景工具(FCI)和院校背景工具(ICI)(15分钟)。测评组件覆盖17个参与国,248所院校,4,807位教师和22,977名学生。^[53]

(三)测试取样与交付

1.院校取样

为了获取能够反映高等教育系统多样化的参与院校组合(组合的标准:院校现状、环境、资金、类型、分布、特色等),AHELO联盟要求国家工程经理(NPMs)就每个组件组织10所院校自愿加入其中。据各国NPMs表示,所辖区域的高等教育机构对此项研究有着极高的热情,甚至一国能有30多所院校有兴趣参与。然而,由于实施阶段的推迟、预算的不确定性,以及与学校最繁忙时间的冲突,导致院校的参与数量比预计的要少,目前参与院校总数(三个组件)共计248所。^[54]

2.师生取样

就整个研究而言,师生样本的选取在保证其可比性和有效性方面起着关键作用。然而,在实际操作中面临的主要困难是由于总体名单难以确定,抽样难以被广泛用于高等教育背景中。因此,每所参与院校首先要提供一份人员名单以便随机取样,也就是说所有的抽样工作都取决于完备的、最新的样本框。学生样本框是参与院校所有本科毕业生,即使他们已修完了课程或在外地实习中也要设法参加,每所学校单项组件的样本量为200人。教师样本框则是在学科或跨学科部门中从事本科教学的所有教员,包括兼职、休假的教师,样本量控制在40人。^[55]

3.交付方式

一旦测评工具和背景调查的翻译、调适、证实工作结束,所有的测试材料都会被上传到AHELO测试系统中,届时学生测试、数据收集和得分都在网上进行。该系统是由两个测试子系统整合而成。

SoNET测试系统适用于3项测试组件的MCOs以及3组背景工具,涉及测试、管理和得分3个模块,CLA平台主要负责通用技能组件的CRTs管理和得分,设有监考员、学生和记分员3个界面。^[56]

(四)学生回复与激励

测评的回复率对于研究结果至关重要,从数据分析中可以看出,各国及所有测评组件的学生回复率呈现出显著的不同,学生的回复从3到273人不等。参与院校大致显示出双峰分布:大约1/4的院校保证20名~50名学生的回复,另外1/4的院校保持在125~175人之间。从学生平均回复率统计中可看出,“工程学”回复率(89.2%)>“经济学”回复率(74.1%)>“通用技能”回复率(52.7%)。可能群体效应是这一结果的最佳答案,即越小规模的群体更易被激发参与其中。尽管AHELO技术标准要求回复率要达到75%,但由于各国人口性别、年龄及抽样设计本身等问题,AHELO联盟建议在未来的实践中将50%设为回复率的门槛。^[57]

为了激励学生参加,各国的参与院校也都使尽了浑身解数,既有将其作为国家的强制性考试(如哥伦比亚),也有以学生自愿参加为原则。特别是后者院校,它们运用了各种刺激办法(使用频率由高到低依次为)——现金或凭证、证书、抽奖、礼物、食品等。然而,刺激与学生参与的关系总是复杂的、非直接的和语境化的。因此,院校策略(并非刺激因素)才是决定学生回复率的关键。^[58]

五、经验与启示:高等教育学习成果测评的拓展性

在OECD开发的四大国际教育测试项目中,AHELO项目争议可谓最大,但其关注并不亚于PISA。因为它不只是测评方式、内容的转变,更是测评理念的革新。从某种意义上而言,AHELO之于其他已有高等教育质量测评而言,无疑具有突破性和创造性,它的测评理念、方式、内容以及技术具有一定的拓展性。

(一)AHELO之理念:以学生为中心

从测评理念来看,OECD研发AHELO项目至少给世界高等教育传达一个信号,即高等教育必

PISA(国际学生评估项目)、AHELO(高等教育学习成果测评项目)、TALIS(国际教学调查项目)和PIAAC(国际成人能力测评项目)。

须要树立“以学生为中心”的思想。这一思想渗透在高等教育的方方面面,自然也包括质量及其评估问题。AHELO作为一个测量工具,其研发目的就是为高等教育机构提供所属学生综合表现的国际比较测评,以求支持、促进学生学习成果的改进。^[59]因此,唯有领会AHELO的真实目的,才能有助于后续工作的开展。

(二)AHELO之方式:国际合作比较

从测评方式来看,AHELO打破了传统的一国之内的测评模式,而是尝试采用国际合作的方式。测评思维转换的同时,也带来了测评工作的复杂和繁重。国际合作比较使AHELO从“图纸”跃入到“现实”,这一方面是由于高等教育国际化的持续推进已经筑造了良好的测评环境;另一方面,则归功于国际、国家和院校三方的紧密合作以及商讨、翻译、调适和证实等多个工序的协同研发。从某种程度而言,AHELO国际合作比较取向的测评方式将促进质量评估模式的国际化。

(三)AHELO之内容:基于学科知识

从测评内容来看,AHELO又开启了一项新的突破,即基于学科知识的测评。

1998年联合国教科文组织发布的《21世纪的高等教育:展望与行动》宣言中指出:“高等教育的质量是一个多层面的概念,包括高等教育的所有功能和活动。”然而,基于学科知识考察学生学习成果并未列入其中,可以说这是高等教育质量评估内容的重大转变。当前,OECD只是聚焦在“通用技能”、“经济学”和“工程学”三个领域,随着AHELO项目的日臻完善,相信将会扩大至更多的学科领域。

(四)AHELO之技术:研发框架工具

从测评技术来看,由于国家、文化和语言背景的差异,测评框架和工具的研发工作相当繁杂。丰富的经验交流和重要的信息反馈将促进工具研发的过程,这就意味着良好的质量控制和质量保证环节较之一国之内测试更为重要。^[60]特别是测评框架的国际一致性问题对于AHELO的后续研发工作将提供宝贵经验。在整个技术研发过程中,不论是选择题(MCQs),还是结构型问答题(CRTs),不论是背景维度,还是工具地方化,AHELO都将为相关测评技术的开发提供有益启示。

参考文献:

- [1][美]罗杰·盖格.大学与市场的悖论[M].郭建如,马林霞等译.北京:北京大学出版社,2013:127.
- [2]OECD'AHELO.Participation[EB/OL].<http://www.oecd.org/education/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdahelo.htm>.2014-04-20.
- [3][8][9][14][20][25][26][27][28][30][31][32][33][34][35][36][38][39][40][41][42][43][44][45][46][47][48][49][50][51][52][53][54][55][56][57][58][59][60]OECD.(2012),Assessment of Higher Education Learning Outcomes:Feasibility Study Report Vol.1[EB/OL].www.oecd.org/edu/ahelo.2014-04-20/05-06_9_34_35_38_61_80_81_83_88_91_95_96_100_99_107~109_109~110_111_135~136_111~112_112_112~116_116~117_121_117_119_121_122_123~124_128_128_131_134~135_151~152_153~156_157~158_161~165_165~169_182_186.
- [4]UNESCO Institute for Statistics (2011),Global Education Digest 2011 [EB/OL].<http://stats.uis.unesco.org.2014.04.20>; Daniel,J.S.(2009),Highlights of the UNESCO Global Forum on Rankings and Accountability:Uses and Misuses,closing presentation [EB/OL].www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/pdf/RANKINGS/Stamenka-JohnDaniel.pdf2014-04-22.
- [5][7]Philip G. Altbach,Liz Reisberg,Laura E.Rumbley (2009),Trends in Global Higher Education:Tracking an Academic Revolution,A Report Prepared for the UNESCO 2009 World Conference on Higher Education [EB/OL].<http://unesdoc.unesco.org/images/0018/001831/183168e.pdf.2014-04-22>.
- [6]W. Norton Grubb.(2003),The Roles of Tertiary Colleges and Institutes:Trade-offs in Restructuring Postsecondary Education [EB/OL].www.oecd.org/dataoecd/50/55/35971977.pdf.2014-04-25.
- [10]Bologna Secretariat (1999),Joint Declaration of the European Ministers of Education[EB/OL].<http://ec.europa.eu/education/policies/educ/bologna/bologna.pdf.2014-04-23>.
- [11]Kennedy,D.(2008),Linking Learning Outcomes and Assessment of Learning of Student Science Teachers. Science Education International,Vol.19(4):387~397.
- [12]Bologna Secretariat (2012),Making the Most of Our Potential:Consolidating the European Higher Education Area - Bucharest Communiqué of the European Ministers of Education [EB/OL].<http://www.ehea.info/Uploads/%281%29/Bucharest%20Communique%202012.pdf.2014-04-23>.

- [13] Robert B. Barr, John Tagg (1995), From Teaching to Learning - A New Paradigm for Undergraduate Education, Change Magazine, Vol.27 (6) :12~25.
- [15] Lederman, D.(2008), Quick Takes : Caution on Plans for World Test, Inside Higher Education[EB/OL].www.insidehighered.com/news/2008/04/04/qt#ixzz252NtfjCn. 2014-04-26.
- [16] Green, M. (2011), Lost in Translation : Degree Definition and Quality in a Globalized World, Change Magazine [EB/OL]. www.changemag.org/Archives/Back%20Issues/2011/September-October%202011/losttranslation-full.html. 2014-04-26.
- [17] Hazelkorn, E. (2011), It's Time to Move Beyond Rankings, Chronicle of Higher Education [EB/OL]. http://chronicle.com/blogs/worldwise/its-time-to-move-beyond-rankings-2/28830.2014-04-26.
- [18] Lederman, D. (2007), A Worldwide Test for Higher Education?, Inside Higher Education[EB/OL].www.insidehighered.com/news/2007/09/19/international.2014-04-27.
- [19][21] Education International (2007), Assessing Higher Education Learning Outcomes : "PISA" for Higher Education?[EB/OL]. http://download.eiie.org/docs/IRISDocuments/Education/Higher%20Education%20and%20Research/Higher%20Education%20Policy%20Papers/2008-00036-01-E.pdf.2014-04-27.
- [22] OECD (2007), Assessing Higher Education Learning Outcomes Summary of a First Meeting of Experts, Washington, 18 April 2007 [EB/OL]. www.oecd.org/edu/highereducationandadultlearning/39117243.pdf.2014-04-27.
- [23] OECD (2007), Assessing Higher Education Learning Outcomes Summary of the Second Meeting of Experts, Paris 5-6 July 2007[EB/OL].http://www.oecd.org/edu/highereducationandadultlearning/39117295.pdf.2014-04-27.
- [24] OECD (2007), Assessing Higher Education Learning Outcomes Summary of the Third Meeting of Experts, Seoul, 26-27 October 2007[EB/OL]. http://www.oecd.org/edu/highereducationandadultlearning/41059347.pdf. 2014-04-27.
- [29] AHELO Consortium (2012), Generic Skills Assessment Framework ;AHELO Consortium (2011), Economics Assessment Framework [EB/OL].http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne (2011) 19/ANN3/FINAL&doclanguage=en.2014-04-28 ;AHELO Consortium(2011), Engineering Assessment Framework [EB/OL].http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne (2011)19/ANN5/FINAL&doclanguage=en.2014-04-28 ; AHELO Consortium(2011), Contextual Dimension Framework [EB/OL]. http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne (2011)21/ann1/final&doclanguage=en.2014-04-28.
- [37] Harkness, J. A. et al. (2003), Questionnaire Design in Comparative Research. Cross-cultural Survey Methods, 19~34.

A New Attempt: The International Trend of Assessment of Higher Education Quality ——The Design and Implementation of OECD's Assessment of Higher Education Learning Outcomes

CHEN Tao

Abstract: The OECD has developed a comparative international programme - 'Assessment of Higher Education Learning Outcomes' which aims to evaluate undergraduates' learning outcomes of Higher Education Institutions during a period of five years. This programme makes a breakthrough on the collegiate modal of traditional quality assessment in terms of its assessment ideas, methods, contents and techniques from 'proposing assumptions' to 'meeting challenges' and from 'developing instruments' to 'putting into practices'. Undoubtedly, it has an importantly practical significance and expanding value as the comprehensive international assessment way based on 'students-centered idea'.

Key words: Assessment of Higher Education Learning Outcomes, feasibility study report, higher education quality, OECD

本文责编:曾晓洁