

教师教育

“基于证据”教育政策研究的评估与整合

——以英国EPPI与美国WWC的经验为例

洪成文¹ 莫蕾钰²

(1.北京师范大学 高等教育研究所, 北京 100875; 2.中国矿业大学(北京) 管理学院, 北京 100083)

摘要: 近期,“基于证据”的研究与政策模式受到英美等国的推崇,对提升包括教育政策在内的公共决策的科学性和社会效益起到了切实有效的作用。深入理解“基于证据”的政策理论,分析英美代表性教育证据评估机构对“基于证据”政策研究的评估理念与实践经验,不仅能为我国教育政策研究的“证据化”提供思路,还能为形成符合我国实际、具有中国特色的教育政策研究评估标准提供参考。

关键词: 教育政策研究; 基于证据的政策; 评估与整合; 英美经验

中图分类号: G510

文献标识码: A

文章编号: 1005-9245(2015)06-0121-07

如何提高包括教育政策在内的公共政策的质量、优化决策的科学性是当前世界各国政府管理部门关注的焦点。近二十年,被称为“千禧年的最大理念”的“基于证据”的研究与政策模式^[1]受到英美等国的推崇,为政策研究的实用化及决策的科学化提供理论与技术支持,已成为影响目前西方医疗保健、农业、科技、教育等各领域公共决策的基本范式。学者们认为实施“基于证据”的政策切实提高了具体政策方案的针对性、有效性和时效性^[2],提升了政府的公信力,而且产生了更大的政策影响与社会效益。

1993年起,我国教育政策追求的目标是教育决策的科学化^[3]。然而,我国教育研究仍无法为教育决策提供科学的证据,无法实现真正意义上的“基于证据”的政策决策。深入理解“基于证据”的政策理论,分析

英美代表性教育证据评估机构对“基于证据”政策研究的评估理念与实践经验,能为我国教育政策研究的“证据化”提供思路,为形成符合我国国情、建设有中国特色的教育政策研究评估标准提供参考。

一、基于证据的政策理论内涵

基于证据(Evidence-based)的理念最早出现于20世纪90年代的医学领域,当时的循证医学(Evidence-based Medicine, EBM)旨在通过科技手段(及实验方式)采纳最优可得证据,用于最终临床和政策决策^[4]。近20年来该思想方法逐渐扩展到社会科学领域的研究与决策之中。联合国在《千年发展目标报告》中将基于证据的政策(Evidence-based Policy,以下简称EBP)定义为:一种帮助规划者在更充分信息条件下决策

收稿日期: 2015-05-01

基金项目: 本文系全国教育科学规划教育部青年项目“教育决策者的科研采信动力分析: 基于知识管理的视角”(EFA140356)的阶段性成果。

作者简介: 洪成文,北京师范大学高等教育研究所教授;莫蕾钰,中国矿业大学(北京)管理学院讲师,博士。

的政策过程,该过程以最佳可得证据为核心^[6]。尽管当前人们对EBP的理解难以达成一致,但一些定义还是可以帮助我们作较为深入的理解。从政府的角度看,EBP为决策提供理想、客观指导的指南,且不受传统、私人利益、意识形态和个人偏见等力量的影响;从学者的角度看,EBP是政策开发和实施的理论方法,使用严谨的技术开发和维持政策开发、决策和实施所需的证据基础^[9]。无论侧重点如何,EBP理论都会涉及三个核心问题:什么是证据、如何生产优质证据、如何评估与整合证据。

(一)(政策)证据的内涵是什么

证据是使得政策目标变得更明确,能获取更优政策效率效果的系统信息。学者们普遍认可“重要的是什么在起作用”,具体涵盖内容存在不同层次:有学者将证据局限于事实、趋势、调查信息等硬数据,亦有学者将符合逻辑和技术标准的分析与推理也包含进来。最为宽泛的定义还囊括相关的观点与经验,如客户或选举人的观点(反馈)以及“使用者经验”等,既包含帮助理解问题的证据,也有提供有效解决方式的证据^[7]。1999年,英国政府内阁办公室对在公共政策中可采用的证据定义较为全面,即:“专家的知识、现有的国内外研究、现有的统计资料、利益相关者的咨询意见、以前的政策评价、网络资源、咨询结果、多种政策方案的成本估算、由经济学和统计学模型推算的结果”^[7]。

(二)如何生产优质的证据

虽然学界对证据的解读不一,但均承认证据的品质存在差异。而产生差异的原因主要存在于三个方面:一是证据是否具有相关性,即效度问题;二是证据是否具有可靠性,即信度问题;三是证据是否具有可复制及可推广性,即政策方案可行性问题。那么生产优质的证据需要相应达到三个方面的标准:第一,优质的证据需要良好的信度,即证据与拟解决的问题在内容上应准确对应;第二,优质的证据要求证据信度高,而源于自然科学的证据概念更偏爱科学实验的结果,近年社会科学研究开始重视验证干预(Intervention)与结果(Outcome)之间的联系,尤其是随机控制试验、准实验研究等更得推崇;第三,证据需带来可行性方案,这些方案不但要具有一定的代表性,符合当前文化与经济需求,而且必须是翔实、可

以由此推进的。

(三)如何评估与整合证据

EBP的核心在于将证据运用于政策的开发与执行之中^[8]。证据的评估与整合是选择并利用证据的基础,也是当前我国推行EBP的基础环节和突破重点。从证据的评估来看,虽然政府官员与学者对优质证据的特点有基本共识,但在具体评估中还存在一些问题,且我国在此方面的研究与经验都较少。如在具体评估上应该秉承哪些理念,在评估标准上是否应当坚持应然的标准,抑或应依据我国政策研究的水平与具体政策问题进行相应调整等。而针对政策议题的证据群的整合更是盘活证据存量、获取全面决策依据的过程,是EBP执行的重点难点。为此,国外学者在社会研究领域沿用和开发了包括系统评价(Systematic Review)、Meta分析、叙述性评论(Narrative Reviews)和现实主义综合分析方法(Realist Synthesis)等富有技术性的方法来分析和处理各类在政策领域中出现的实验及经验证据。

当前我国虽强调将科学证据运用于决策,但具体到“基于证据”的政策实践的经验还有限。尤其在证据的评估与整合方面,亟须了解与借鉴英美等发达国家的经验,加强机构、理念及操作。英美国家在教育领域中EBP的经验,为优化我国教育政策研究与决策提供了思考。

二、英美教育中的EBP

(一)政府的重视

英国于1997年布莱尔政府开始倡导基于证据的政策活动,并在1999年英国工党政府的《政府现代化》白皮书中明确表示将“基于证据的政策”纳入政府制定政策的基本理念^[9],并设立了ESRC EBBP(经济和社会研究委员会“基于证据的政策和实践中心”)、EPPI(政策与实践信息协调证据中心)、Campbell(Cochrane协作干预评估中心)、SCIE(杰出社会保障研究与实践审查中心)等四个中心计划来系统深入地发展EBP的理论及实践,并且这些中心都涉及教育类研究的评估与整合。

美国于2002年布什总统“不让一个孩子落后”(NCBL)法案开启了美国教育“基于证据的政策”模式之门,此后的奥巴马政府被称为美国历届政府中

最重视决策证据的政府,在众多教育政策制定中都强调科研证据的作用。

(二) EPPI 与 WWC 的设立

将“基于证据”的理念具体落实在教育行业中,国内外学者和实践者都在证据的来源、分类与界定、评估、总结等方面展开了讨论与尝试^[9]。英国的政策与实践信息协调证据中心(The Evidence for Policy and Practice Information and Coordinating Centre, EPPI-Centre, 以下简称 EPPI)与美国的有效教育策略信息交流中心(What Works Clearinghouse, 以下简称 WWC)都获得其政府支持并在行业中得到充分的认可,其范式已走在国际前列。分析它们之间的理念与实践差异,选择符合中国特色和基础的方式,对我国建立和发展“基于证据”的教育研究与决策具有重要意义。

英国 EPPI 由伦敦大学教育学院主导于 1993 年成立,由英国经济和社会研究委员会(ESRC)、医学研究理事会、国家卫生和卓越护理研究院(NICE),欧盟、英国若干政府部门和慈善机构及其他国家和国际伙伴提供资助,在针对社会科学(含教育)与公共政策等领域,执行研究整合与文献评审工作,旨在为政策制定者或实务工作者的决策提供可靠的研究发现。研究领域包括教育和社会政策、健康促进和公共卫生、国际卫生制度和发展等,其在社会科学与公共政策领域的系统评价及其方法论方面一直处于领先地位。

美国 WWC 由美国教育部出资于 2002 年成立。该中心由美国教育科学院(Institut of Education Science, IES)负责,委托数学政策研究公司(Mathematical Policy Research, Inc.)建立。WWC 旨在透过发展与运用标准程序来评阅与整合教育相关研究,寻找有效、高质量的研究,为教育决策提供可靠的科学证据来源。WWC 希望通过高效模仿“循证医学”的方式,解决教育政策中的问题,让国会议员、民众清楚地知道他们花的钱取得了哪些成效。WWC 文献综合评估的主题含成人识字、初学阅读、品格教育、高等教育、择校、教师教育与有效领导、教师激励等 18 项主题。

三、EPPI 与 WWC 的评估理念、产品及范围

英美两国推进 EBP 评估的总体目标是一致的,

即为教育政策制定提供更具可靠且更易获取的证据,而英国 EPPI 与美国 WWC 均为政府出资、具有相当社会影响力和技术优势的专业评估机构,它们都提供了决策者和实践者可理解的各类评估产品,为推进本国教育决策的科学化贡献着不容小觑的力量。然而在评估实践中,这两个中心在理念上存在不同,同样在评估范围及产品上也有所差异。

(一) 评估理念

英国 EPPI 在证据评估中强调权变性与人性化,将评估看作是开放性的研究过程,给予评估专家充分的自主权来制定符合评估论题的研究方法,且可以随着研究的深入调整评估计划。为了顾及利益相关者的意见,将用户纳入到评估过程的各个环节中去,保证服务使用者、教育实践者、研究者、雇主及公众成员及包含不同性别、年龄、种族的各类代表均能参与。

美国 WWC 在证据评估中强调标准化与独立性,将评估过程看作是专业性的评判过程,有着统一、公开、明确的严苛标准,评估过程有清晰的抽样标准或数据计算程式,强调技术专家的独立性,尽量避免非专业人员(如政治人物)的涉入而造成“证据污染”。

(二) 评估的产品

EPPI 主要采用系统评估的方式来对证据进行评估与整合,强调研究方法成系统,最终获得可信的结果。评估过程本身就是一种研究,研究最后会对现有该议题的研究发现进行合成,并以决策者和实践者可理解的方式发布。其产品注重不同的阅读对象,包括三种形式(见表 1)。

表 1 EPPI 评估产品

产品形式	内容	特点
小结 (Summary)	篇幅为 1 页左右,总结评估研究的结论	重实用、易读
短报告 (Short Report)	篇幅略长,对研究的结论背景描述更详细	强调研究结论及背景,便于经验复制
技术报告 (Technical Report)	篇幅约为 100 页,包括评估的细节及方法论述	强调技术方法分析,提供信效度支持,便于未来的研究方法讨论及更新
附属产品:研究数据库、评估人工作软件系统等		

WWC 的评估产品由评估目标与证据类型来决定。既包含对某干预因素的检验,多见于对某研究议题的证据评估与整合,又包括针对实践经验、个别研

究项目的其他报告形式,主要有以下四种(见表 2)。

表2 WWC评估产品

产品形式	内容	特点
教育干预报告	审阅特定时间段里已经出版的相关教育干预效果的所有研究。对于达到 WWC 标准的研究,会综合所有的研究发现来估计教育干预的效果值。	根据研究的数量、样本大小、数据的显著性和量值来阐明干预率(Intervention Rating)和研究结果的有效程度。
实践指导报告	针对一线实践教育工作者的具体问题提供实践建议。	基于评估总结和国家公认的专家组(含研究者、教育专家)的专业判断。
个别研究评估	对个别教育项目、教育政策、教育实践研究的评估。	辨别研究是否达到 WWC 标准,如达到,将会总结其研究结果的有效程度。
快速评审	获取媒体足够重视的研究将被优先评审。	研究会快速审核,并发布三页的快速审阅小结。
附属产品:研究数据库		

(三)评估的范围

1.评估的内容

英国 EPPI 一般使用系统评估方法对证据进行评估与整合。传统系统评价更关注针对于干预性研究的评价,但对非干预性研究的查找、评价以及合成是 EPPI 目前面临的主要挑战和工作重心^[6]。而对于教育研究议题的评估并没有具体的约束,该中心正积极拓展教育研究以及英国现代教育研究数据库。

美国 WWC 目前主要审核四类研究设计,包括:随机对照试验(RCT)、有匹配对照组的准实验设计、回归不连续性设计(RDD)以及单案例设计(SCD)。并为符合标准的研究或主题提供因果证据报告。即便如此,WWC 需要评估的研究仍较多,故而对评估对象有具体的评分式预挑选,具体准入及优先指标见表 3 所示。

2.数据的搜集范围

EPPI 为了避免数据搜集的价值倾向和遗漏,其搜寻对象包含未发表及难获得的研究结果,并根据系统评估项目的具体需求,选择并综合使用多种搜索源,以达到全面性和针对性的均衡。基本搜索源涵盖:电子数据库搜索、手动核心期刊检索、专业网站搜索、搜索引擎(如“谷歌”和“谷歌学术”)、个人联系丰富搜索范围(访问已收录研究的作者和该领域专家)。在搜寻过程中注意技巧,其中包括评估问

题与数据库的选择、关键词的变换及研究方法的评估等,要求评估组成员接受相应训练。

WWC 的证据搜索范围一般包括已出版的文献和虽未出版但可以找到的公开文献资料。接受评估的研究必须同意被公开获取,且个别研究评估和快速评审必须达到一定程度的公众关注才可纳入评估。符合 WWC 资质的评估人会按照具体评估草案中的关键词,在 WWC 常用综合搜索数据库(13 种)和组织网站(25 种综合网站,32 种专题搜索数据库及网站,54 个媒体网站)中进行搜索。

四、EPPI与WWC评估的过程与质量控制

(一)过程控制

1.基本流程控制

EPPI 强调评估方法明确且透明,严格遵循标准化步骤:组建评估及顾问组→确定评估计划(范围及策略)→搜集数据→评估数据处理→撰写报告→评估结果的传播。其中直到评估数据处理阶段,EPPI 均允许评估专家在详细的论证之下调整评估标准与纳入评估的具体研究数量,但对每个过程均需要

表3 WWC评估研究的准入标准及优先指标

	准入及优先标准	具体指标	评分
教育干预报告	内部效度标准	随机控制试验	3
		回归间断设计或单例设计	2
		准实验设计	1
		不符合以上标准	0
	规模标准	清晰指出被试数量达到 250 个以上儿童或 10 个以上班级	加分
实践指导报告	实验干预权重	无前研究	3
		有前研究	$[1+0.1 \times (\text{当前年} - \text{首报告年})]^2$
个别研究评估	准入及优先标准	IES 请 WWC 就特定研究进行评估 研究来自网站帮助平台的公众意见,或获 IES 资助的研究 研究符合 WWC 快速审查的标准	
快速评审	及时	必须近期发布	
	曝光性	在全国各大新闻源或重大教育新闻出版物上发表	
	主旨明确	必须已经或有意探讨教育干预直接或间接提高学生学业的有效性(符合 WWC 标准的快速评审研究可再进行个别研究评估)。	

详细的记录,可供相关研究人员随时查证,以此来保证评估过程的严谨与可复制性。EPPI 在评估结果的传播上也十分重视,不仅鼓励对研究证据在现实中的应用提供开放式解读与讨论,还为评估结论解读提供详细的背景信息,为未来实践应用与评估更新提供参考。

WWC 的评估步骤包括:拟定评估草案→确定评估(文献)范围→按照 WWC 标准初步筛选合格研究→评估合格研究→撰写报告。WWC 比 EPPI 多一个初步筛选过程,后期会在网络上公开最终报告文本。其中筛选流程控制图如下(见图 1)。

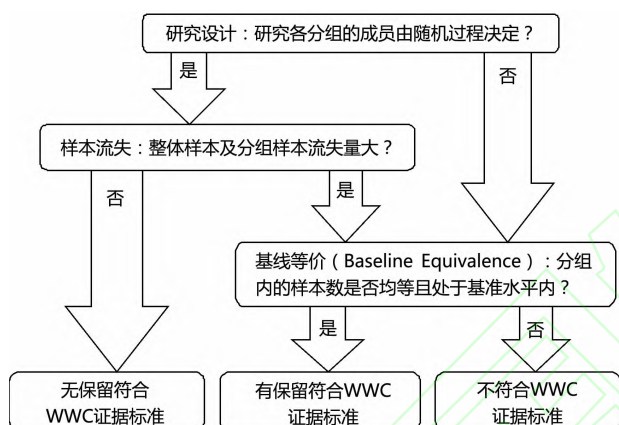


图1 WWC的研究设计筛选评估流程控制

2. 标准控制

虽然 EPPI 与 WWC 都强调评估人员应受过专业的训练,获得过资质或相关学位,但就具体评估来看,EPPI 评估仅在方法质量、方法適切性、主题相关性和证据总体权重的确定方面进行原则性规范,而给予评估专家更大的自由度。而 WWC 标准更趋于明确,即便对于主观性较强的实践指导报告也给出了相当清晰的评估分档标准(见表 4)。

表4 EPPI与WWC的证据评估标准比较

国别	问题来源	评估团队	搜索范围	量化评估标准控制	质性评估标准控制
英国 EPPI	主要为政策议程 ^①	评估组、咨询组	依据需要选择数据库与网站,必要时依据重要作者和专家访谈添加文献	标准由方法论专家主导,重视控制性实验分析	为专家评估提供基本框架列表,但主要依靠专家的主观判断
美国 WWC	主要来自研究者、传媒议程及公共议程	评估组	有固定数据库和网站库进行查验	重视随机控制试验和准实验研究,注意内外部效度及信度分析	限于实践指导的证据评估,逻辑推理与实践经验的价值低于实验结果。

(二)质量保证

1. EPPI 的质量保证

首先,内部审查评估。包括聘请一组具有不同专长的专家合作进行评估,确保参与评估的研究具有相关性和恰当的质量,最终的报告也具有高相关性和高质量。并且在关键阶段可得到咨询小组支持。

其次,外部审查评估。系统评估不但查阅主要的研究计划和学术出版物,还在评估协议制定和最终报告阶段参考同行评议。尤其欢迎同行专家对评估的质量和相关性作出独立评估。

再次,良好的过程控制。在评估的不同阶段,都查证评估人对评估过程与标准的理解和贯彻。一方面,EPPI 在评估正式开始前提供任务模拟服务;另一方面,EPPI 通常要求对每个研究的数据提取和整合都至少由两个评审人独立评审,再将他们的结论进行比较。

最后,维护数据库记录。每个阶段都有详细的数据记录。这样不但保证了评估的透明性(可讨论监督)、可持续性更新(过程及方法的优化及该领域研究更新),还为对特定研究感兴趣的用户提供详细的信息。

2. WWC 的指标评估控制

WWC 的标准严苛、清晰,具有较强的可操作性,且由于指标系统较为全面和详尽,对于不同的评估者来说,产生差异化结果的可能性已被大大降低。具体如表 5、6 所示。

总的看来,英国 EPPI 的系统性评估方式沟通成本高、耗时较长,但开放性较好,鼓励学者讨论与交流;WWC 标准严苛,损失了部分证据的人文及其他细节性考量,但评估效率较高,激励了政策研究者以更类似于自然科学的方式开展研究。

五、对我国教育政策研究的建议

(一)为智库提供服务平台,创造更高质量的政策“证据”

2014 年初,教育部印发《中国特色新型高校智库建设推进计划》通知,提出高校要为“教育现代化建设”提供智力支持,充分体现出政府对提高教育研究服务政府决策能力的希望和重视,也为未来“教育证据”的生产提供了政策支持。此前,教育部还专

表5 WWC干预报告的评估指标

评估内容		分类	具体解释
研究幅度评估		连续性结果	$g = \frac{\omega(y_i - y_c)}{\sqrt{(n_i - 1)s_i^2 + (n_c - 1)s_c^2}} \cdot \frac{1}{\sqrt{n_i + n_c - 2}}$
		二分性结果	$d_{\text{obs}} = \omega \left[\ln \left(\frac{p_i}{1 - p_i} \right) - \ln \left(\frac{p_c}{1 - p_c} \right) \right] / 1.65$
	改进指数	实验组均值 - 控制组均值	
统计意义	累计统计	$t = g \sqrt{\frac{n_i n_c}{n_i + n_c}}$	
干预有效性评估	积极作用	显示强积极效用,无有力反证	
	潜在积极作用	显示积极效用,无有力反证	
	无明显效果	对效果无肯定性证据	
	混合效应	显示不一致的影响证据	
	潜在的负面影响	显示负效应的证据,无压倒性反证	
证据程度	负面影响	显示有力负面影响证据,无压倒性反证	
	大中型	含的研究多于1个(必要)	
		含1个以上的背景设置(必要)	
		样本量多于350个学生(必要)或含有14个班×每班25名学生的样本量	
	小型	仅含1个研究(可能)	
		仅含1个背景设置(可能)	
		样本量少于350个学生或每班25名学生的情况下,样本量低于14个班	

门设立“国家教育咨询委员会”,为教育部决策科学化提供组织保障。然而,落实有证据的决策还需要高校转变观念,提供具体战略规划予以支持,且高质量地实现该计划更需要高校具体的政策定位与评估系统的配合。根据上文研究,我们可参考英国的经验^②,建立与高校研究声誉与资助水平直接挂钩的配套激励制度。

当然,智库的水平也与智库的独立性有关联。与政府距离近的决策研究和咨询机构,其独立性会受到一定的影响。因此,在发展政府系统内的政策研究智库的同时,我们也不能忽视其他智库的作用。因此,一个比较合理的主张是:整合行政性智库和民间智库的作用,建构有效而科学的智库系统。

(二)设立清晰的评估标准

“基于证据”的教育研究与决策也经常受到批评,很多学者认为其严苛的统计化倾向和方法论有脱离教育本质与过程的嫌疑。但我们无法否认,更严谨的实验和数据分析所提供的证据已得到包括欧盟、世界银行等国际组织与欧美先进国家的充分认可并广泛应用于决策之中,且常得成功。一味地批判有可能使得我们在此领域落后于人。因此,理性的思考是建立清晰的评估标准。

表6 WWC实践指导报告的证据强度分析标准

标准	坚实证据基础	中等证据基础	微弱证据基础
效度	研究符合 WWC 标准,且有高度的内部和外部效度	研究具有高内部效度但中等外部效度;或高外部效度和中等内部效度	证据不能达到中等或高度证据的研究标准
对相关结果的影响	研究表明一致的正向影响,研究内部效度高,且不存在不一致的证据	研究显示正向影响的证据占优势。研究中的矛盾证据,必须在指导的相关范围内进行讨论,并将推荐强度视为评估干预的组成部分。	可能有效果微弱或矛盾的证据。
相关性范围	研究范围内显示出直接相关性——相关内容、抽样、比较及结果都将被评估	研究范围内相似的相关性可能变化。至少有一些研究能在范围内显示直接相关关系。	研究也许在实践指导的范围之外
研究与建议的关系	研究中直接对建议进行测试,或是建议是被干预中的主要成分	作为被干预的组成部分,“建议”的推荐强度不同。	作为被干预的组成部分,“建议”的推荐强度低,或反映专家意见的建议是由研究合理推断出来的。
专家组信心	专家组对这项实践的有效性有高度信心	专家组判定研究不能达到高强度水平,但比最低程度的证据有说服力。专家组不能确定在大部分或全部情境下该实践能切实有效。	根据专家组意见,这些建议只能被当做实践指南的一个部分;且专家组不能指出研究中显示出中度或高度的证据。
专家意见的作用	不适用	不适用	根据理论合理解读的专家意见。
重点推荐的评估	评估标准参见“教育与心理学测试标准”。	证据的样本效度并不足以代表建议所关注的人群。	不适用

尽管美国倡导的随机控制试验和英国重视的控制实验都存在范围广、样本大、控制难、单项目费用高等问题,在现实中推进有一定的困难,但其有效范围和可复制程度都远远高于普通的调查研究,也非小规模调研叠加能比拟,明显提高决策与实践的价值,是未来研究发展的趋势。我们需要及时学习随机控制试验的研究模式,并依此建立相应的高标准,如确保研究设计的程序信度、内外部效度^[1],虽然在推进的初期会有一定的难度,但对未来教育研究水平的整体提高起着标杆作用。当然,鉴于当前研究水平及客观实验条件和经费的限制,参考 Sanderson 的观点,将证据的标准从“什么才是有效的(What Is Work)”转换为“什么才是适当的(What Is Appropriate)”^[13],以此来为研究设计提供一定的灵活度和次优标准,也不失为我国“教育证据”研究发展设定的阶段性目标。

(三)设立专业的评估组织及数据库

教育证据评估机构不仅需要本领域专家与方法论专家团队(包括质性研究和量化研究专家),而且也需要专业的软件管理与数据处理专家。此外,为了吸纳合适的议题、了解教育决策者与一线决策者的需求,评估机构还应有相应的调研方式及互动平台,这都需要人力和财力的支持。

建议政府成立有关机构,建设国家级的证据数据库。除建立国家级的系统性评阅组织外,亦可与国际学者合作,建立以证据为本的联盟,加强交流、合作,分享相关数据。在有限资源下,初期由独立研究者透过小型系统性评阅组织的成立,为建立证据本位的理念而努力,亦是国内教育学者们可参考的方向。这样不仅能为国内教育提供长期性的证据,有效解决现行研究元分析所面临的资料搜集费时、费力、连续性研究资料累积的困难,还能提高研究成果的可信度,为未来决策改善和研究资料的整合、更新提供便利。

注释:

- ①英国“基于证据”的教育研究秉承的是英国政府“决策现代化”的理念,主要为政府的决策工作服务。
- ②近年来,英国改革了大学研究工作的评价体系(RAE),建立了新的(REF)评价体系。核心是在传统研究质量评价的基础上增加了对社会影响力和学术环境两方面的评价。其中,在对社会影响力的评价中,重要的一项指标是研究证据对公共政策与服务的影响,以此来引导和激励科学家在学术研究的基础上增强对政策制定的支撑作用。该评价结果既决定了英国大学在研究领域的声誉,还决定了政府对大学拨款资助的规模,直接激发了高校生产高质量“证据”的动力。

Evaluation and Integration of Research in “Evidence-based” Education Policy ——Take British EPPI and American WWC Experience for Example

HONG Cheng-wen MO Lei-yu

(1. Institute of Research in Higher Education, Beijing Normal University, Beijing 100875; 2. College of Management, China University of Mining and Technology, Beijing 100083)

Abstract: “evidence-based” research and policy model admired by British and American governments take feasible effect on scientific public decision including education policy and social benefit. Deeper understanding of “evidence-based” policy theory and analysis of assessment notion and practical experience made by British and American leading education testimony assessment agencies offers not only thought for “evidence-based” research in our education policy but reference for assessment criterion for research in education policy with Chinese characteristics in line with China’s reality.

Key Words: Research in Education Policy; Evidence-based Policy; Assessment and Integration; Experience of U.K and U.S

【责任编辑:马瑞雪】

参考文献:

- [1]Ray Pawson.Evidence-Based Policy:A Realist Perspective [M].SAGE Publications Inc,2006.
- [2]Brown-Chidsey,R. & Steege,M.W. Response to intervention:Principles and strategies for effective practice [M]. New York:Guilfor,2005.
- [3]洪成文.完善知识动员机制促进教育决策科学化[J].教育发展研究,2014,(5).
- [4]Philip Davies.What Is Evidence-Based Education?[J].British Journal of Educational Studies,1999,(2).
- [5]United Nations International Children's Emergency Fund (UNICEF). Bridging the Gap.The Role of Monitoring & Evaluation in Evidence-based Policy -making[EB/OL]. http://www.unicef.org/ceecis/evidence_based_policy_making.pdf,2015-04-21.
- [6]马小亮,樊春良.基于证据的政策:思想起源、发展和启示[J].科学学研究,2015,(3).
- [7]Strategic Policy Making Team Cabinet Office.Professional Policy Making for the Twenty-first Century [EB/OL]. <http://dera.ioe.ac.uk/6320/1/profpolicymaking.pdf>,2015-02-13.
- [8]Davies,H.,Nutley,S.& Smith,P.What works? Evidence-based Policy and Practice in Public Services[M]Bristol:Policy Press,2000.
- [9]李晓轩,杨可佳,杨柳春.基于证据的政策制定——英国的实践与启示[J].中国科学院院刊,2013,(6).
- [10]阎光才.对英美等国家基于证据的教育研究取向之评析[J].教育研究,2014,(2).
- [11]钮文英.教育研究方法 with 论文写作[M].台北:双叶书廊,2007.
- [12]Sanderson,W.C.Are evidence-based psychological interventions practiced by clinicians in the field [J]Medscape mental health:a Medscape eMed Journal,2002,7(1).