

数据驱动的教学决策效果总览与视野突破

——基于十项元分析的伞型评价

刘晋¹, 张学波², 林书兵³

(1.华南师范大学 信息化建设管理办公室, 广东 广州 510631;

2.华南师范大学 教育信息技术学院, 广东 广州 510631;

3.北京师范大学 人文和社会科学高等研究院, 广东 珠海 519087)

[摘要] 数据驱动的教学决策应用作为当前基础教育实践的新趋势, 仍然存在着证据不足、效果不明和路径不清等问题, 即教学数据干预的作用范畴涉及哪些方面, 数据使用具体给学校带来哪些实际成效, 如何评价其实际效果, 掣肘数据使用的主要因素有哪些, 学校具体如何构建教育数据使用的实践机制, 这一系列问题反映着教育数据应用的主要困惑, 也预示着数据驱动教学决策的深入探究方向。鉴于相关专题元分析数量不断增长但存在角度和范围的局限, 本研究采用伞型评价的方式, 提取十项基于元分析的研究, 从作用范畴、周期、措施和路径等角度对数据驱动的教学应用干预效果进行了系统性再回顾, 并提出了数据驱动的主要影响因素、决策方法、评估机制和长远目标等关键议题, 以期为后续相关研究提供全面分析结论和未来突破方向。

[关键词] 数据驱动的教学决策; 效果评估; 伞型评价

[中图分类号] G434

[文献标志码] A

[作者简介] 刘晋(1984—), 男, 湖北武汉人。助理研究员, 博士研究生, 主要从事多模态教学行为分析、智慧校园建设研究。E-mail: liujin@scnu.edu.cn。张学波为通讯作者, E-mail: zhangxb@126.com。

一、研究背景

在过去的20年, 促进“基于数据的决策”(Data-Based Decision Making, DBDM)、“数据驱动的决策”(Data-Driven Decision Making, DDDM)和“基于数据的个性化”(Data-Based Individualization, DBI)等改革举措迅速增加。教育决策者们越来越强调通过数据分析告知教师学生的学习需求, 以改进教育教学措施, 达到提高学生成绩的目的。

数据作为当今社会现代性的一个典型隐喻, 与科学性、客观性和有效性等特征天然地联系在一起。数据使用往往携带着美好的假设, 所有促进数据使用的干预措施都根植于这一信念, 即认为只要正确收集和

分析数据, 便能为关键的教育问题提供答案, 随后将产生更好的教育结果^[1]。基于数据的教学决策似乎具有不证自明的先进性与合理性, 但对于具体的数据应用过程, 学校仍需验明其实际成效和厘清其具体实现机制。尽管DBDM在提高学生成绩方面被赋予很高的期望, 但其干预措施的科学性和有效性仍然相当有限, 目前关于DBDM的实施和效果的建议并没有建立在强有力的实证基础上。虽然有些开展了支持数据使用的各种干预研究, 但缺乏对支持学校数据使用的专业发展干预措施效果的系统研究^[2], 因此也就无法回答影响数据使用的因素, 以及这些因素之间如何相互影响。

在国内, 以数据驱动精准教学命名的研究和实践

基金项目: 2018年国家社科基金重大项目“信息化促进新时代基础教育公平研究”(项目编号: 18ZDA334); 2022年华南师范大学哲社社会科学重大培育项目“数据驱动的教学决策理论与实践研究”(项目编号: ZDPY2207)

不断涌现,主要集中在精准教学策略、活动、模式、模型和操作路径等方面^[3-5],但数据驱动是否就一定能确保“精准教学”?教师的决策机制和过程如何对精准教学产生影响?哪些因素会左右其精准性?精准的程度如何判定?这些问题并没有得到准确回应。当前相关教育数据应用常陷入精准、智慧、深度等模糊不定、似是而非的空泛表述语境,需要来自教育一线数据应用实证研究的科学结论,对数据应用理论、方法和技术应用于教学的实际效果作出准确判定。如何让数据驱动的变革从短期的课题、项目或活动式推进方式转变为学校教育的科学决策和理性思维的第一选择,从而成为学校变革的长期稳固范式,需要对当前数据驱动的教学决策效果评估和机制构建层面深入研究。

此外,有关教育数据应用效果的实证研究以及元分析数量已有很多,但既往的相关文献评述在数量、时效、检索程序范围和效果评估等方面都有不同程度的局限和偏颇,现有综述仍然没有告诉我们为什么基于学校的决策在不同的条件和环境下会产生积极或消极的影响。因此,需要站在更高层面对教育数据应用于教学决策的整体效果进行审查和概览,系统回应数据驱动教学决策的应用实效性和实现机制之问,并透过一般层面数据干预研究,总结教育数据应用的普遍规律和策略,为更深入地研究构建教育数据应用理论和机制提供理论基础。

二、伞型评价程序

伞型评价^[6](Umbrella Reviews)是一种基于循证的分析方法。伞型评价来源于系统评价和元分析,但又与二者存在差异,其相互关系和比较分别如图1和表1所示。系统评价及元分析是以大量的单个研究为基础进行的综合研究,而伞型评价通过对某项研究主题的所有系统评价和元分析进行再次系统评估,从而得出更可靠的结论。当某项研究主题的系统评价和元分析达到一定数量以后,且这些系统综述和元分析得出的研究结论存在一定争议时,才有必要进行深入伞型评价。近年来,有关DBDM的元分析数量不断增加但主题范畴和结论启示各不相同,有必要对此进行更高层次的总结概括。Hennessy等提出的伞型评价具体步骤^[7]得到了学术界普遍认可,为开展相关研究提供了指南。

本研究使用Web of Science、ERIC、PsychINFO、Education Full Text、ProQuest学位论文和学术论文等电子数据库进行搜索,关键词主要分为三类:(1)使用数据为决策提供信息的过程常用描述语包括数据驱

动的决策(Data-Driven Decision Making, DDDM)、课程本位测量(Curriculum-Based Measurement, CBM)、基于数据的个性化(Data-Based Individualization, DBI)、基于数据的项目调整(Data-Based Program Modification, DBPM)、基于数据的决策(Data-Based Decision Making, DBDM)、反馈干预(Feedback Interventions)、形成性评价(Formative Assessment)等;(2)常用于干预结果范畴包括学生成绩、学术成绩、学术成果、学生学习、教师专业发展、学校提升、学校改进等;(3)文献类型包括综述、元分析、系统综述等。

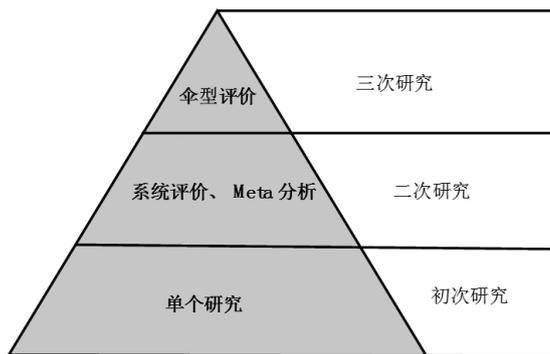


图1 三种研究关系图

表1 系统综述和伞型评价的方法比较

比较维度	系统综述	伞型评价
目标	总结原始研究的证据	从系统综述中总结证据
选择标准	描述纳入和排除标准,主要研究对象是原始研究	描述纳入和排除标准,主要研究对象是系统综述
搜索	相关原始研究的综合搜索	全面搜索相关系统综述
排除	包括符合资格标准的所有原始研究	包括符合资格标准的所有系统审查
方法质量/偏差风险评估	评估纳入原始研究的偏倚风险	评估纳入的系统评价的系统回顾和原始研究的方法质量/偏倚风险
数据收集分析	使用元分析、网络元分析和/或叙述性总结分析等方法对所有原始研究进行分析,得出相关重要结果	包含所有系统性综述,综合每个重要结果或重新分析纳入系统综述中的结果数据

根据伞型评价的基本标准:仅纳入与研究主题相关的系统评价和元分析,不纳入其他类型的文献;包括数据应用的结果,如学生成绩变化、教师素养提升或学校改进;纳入的元研究原则上建议包括以下纳入标准:(1)使用了随机对照试验、准实验或单一受试者设计;(2)包括至少七周的干预和进度监测;(3)包括效应大小描述值。图2以PRISMA图^[8]的形式概述了本研究搜索过程和结果。

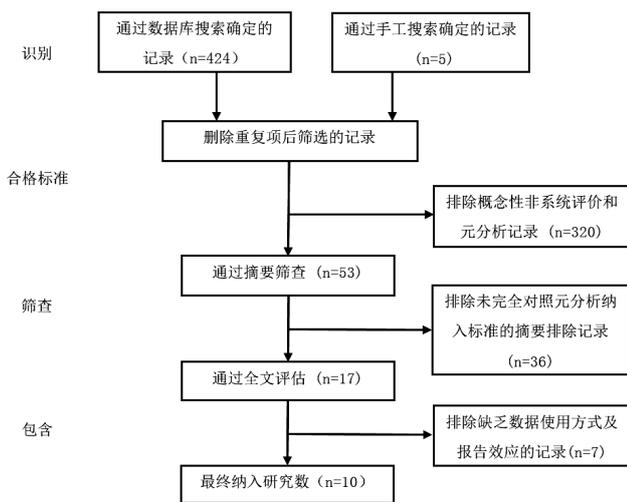


图2 文献搜索和选择过程 PRISMA 图

数据提取内容都来源于原始系统评价和元分析,三名作者团队评审员独立收集相关数据,然后对这些信息的准确性进行交叉核对,具体收集关于纳入研究的关键特征见表2。数据提取过程提取可靠性通过让两名研究人员独立编码约50%的选定文献,获得80%的一致性率,科恩 kappa 系数为0.62,证明了令人满

表2

元分析统计

作者	研究数量	学科范畴	作用范畴	平均效应	干预措施	干预时长	参与者年龄	数据收集频率
Scholin S E, Burns M K, 2012 ^[9]	18	阅读	阅读流畅性	0.30~0.40	课程本位测量	1 学年	1~5 年级	未明确
Ansyari M F, Groot W, De Witte K, 2020 ^[10]	17	数学	成绩、教师专业发展	0.62	教师专业发展干预	大于1 年	小学、初中、大学	未明确
Jung P G, McMaster K L, Kunkel A K, et al, 2018 ^[11]	14	阅读、数学和拼写/写作	成绩	0.37	基于数据的个性化	大于7 周	未明确	未明确
Peters M T, Förster N, Hebbecke K, et al, 2021 ^[12]	6	阅读	阅读流畅性、阅读理解	0.10~0.13	学习进度评估	1 学年	小学表现不佳的学生	每3 周8 次
Visscher A J, 2020 ^[13]	6	数学、阅读理解和拼写	成绩	0.37	基于数据的决策	大于1 年	4~5 年级	未明确
Gesel S A, LeJeune L M, Chow J C, et al, 2021 ^[14]	26	教师知识、技能和自我效能	教师知识、技能和自我效能	0.57	基于数据的决策	大于15 周	职后或职前的中小学教师	未明确
Stecker P M, Fuchs L S, Fuchs D, 2005 ^[15]	12	阅读和数学	成绩	部分差异	课程本位测量	大于7 周	1~6 年级	每周2 次
Hendriks M A, Scheerens J, Slegers P, 2014 ^[16]	26	语言、数学、科学和其他科目	成绩	0.07	形成性评价	至少7 周	未明确	未明确
Fuchs L S, Fuchs D, 1986 ^[17]	21	未明确	成绩	0.70	形成性评价	小于3 周, 3~10 周, 大于10 周	特殊教育学生	每周至少2 次
Filderman M J, Toste J R, Didion L A, et al, 2018 ^[18]	15	阅读	成绩	0.24	基于数据的决策、课程本位测量	未明确	幼儿园到12 年级	每周至少1 次

意的编码器间可靠性。采用 AMSTAR^[19]标准对入选文献进行评分,依据11个条目检测系统评价/元分析控制偏倚的情况。若同一个主题有多篇元分析,纳入最新发表、纳入原始研究最多、AMSTAR 评分高的元分析。基于高质量的研究的综述考虑,论文不低于7分才能被纳入研究。在团队两个作者筛选了最初的搜索标题和摘要之后,如若两名作者不能达成一致,将引入第三个作者参与决定。筛选后,完成全文筛选,以确定包含在审查中的文本。最终,10项总括审查符合进一步分析的纳入标准,见表3。

三、数据驱动的教学决策效果综合分析

基于文献筛选,本研究从以下角度对元分析搜索结果进行总体定性分析:(1)数据驱动教学决策有效作用范畴,以定量研究数据作为主要汇总指标,从范围、主题和程度多方面分析数据驱动的教学决策应用状况;(2)数据干预有效作用周期,分析数据干预作用时长,频率和周期变化等影响情况;(3)数据干预有效措施,分析不同干预措施的设计背景和有效特征因素;(4)数据干预有效路径,分析相关数据应用干预的

表 3

纳入总括审查的标准和评估

作者 标准	Scholin S E, Burns M K, 2012 ^[9]	Ansyari M F, Groot W, De Witte K, 2020 ^[10]	Jung P G, McMaster K L, Kunkel A K, et al, 2018 ^[11]	Peters M T, Förster N, Hebbecke K, et al, 2021 ^[12]	Visscher A J, 2020 ^[13]	Gesel S A, LeJeune L M, Chow J C, et al, 2021 ^[14]	Stecker P M, Fuchs L S, Fuchs D, 2005 ^[15]	Hendriks M A, Scheerens J, Slegers P, 2014 ^[16]	Fuchs L S, Fuchs D, 1986 ^[17]	Filderman M J, Toste J R, Didion L A, et al, 2018 ^[18]
是否提供了前期设计方案	是	是	是	是	是	是	是	是	是	是
纳入研究的选择和数据提取是否具有可重复性	否	是	否	是	是	是	否	是	否	是
是否实施广泛全面的文献检索	是	是	是	否	否	是	是	是	是	是
发表情况是否已考虑在纳入标准中	否	是	是	是	是	是	否	是	是	是
是否提供了纳入和排除的研究文献清单	是	是	是	否	是	是	是	是	否	是
是否描述纳入研究的特征	是	是	是	是	是	是	是	是	是	是
是否评价和报道纳入研究的科学性	是	是	是	是	是	是	是	是	是	是
纳入研究的科学性是否恰当地运用在结论的推导上	是	是	是	是	是	是	是	是	是	是
合成纳入研究结果的方法是否恰当	是	是	是	是	是	是	是	是	是	是
是否评估了发表偏倚的可能性	否	否	是	是	是	是	是	是	是	是
是否说明相关利益冲突	是	是	否	是	否	是	否	否	否	是
得分	7	9	8	8	8	11	7	9	7	11

作用方式和途径。

(一)数据干预有效作用范畴

从数据驱动的教学干预效果来看,效应大小是相关实证研究干预效应最直接的证据,代表着实验组和对照组之间的标准差。从文献总体来看,教育数据干预的总体效应大小普遍低于 0.8,位于小到中等程度,与事先人们所期待的“立竿见影”和“包治百病”式的作用效果有一定落差,也预示着教育数据应用并非一蹴而就的过程。由于实验条件、范围和规模的限制,很少有研究涉及了不同条件下 DBDM 的效果对比,如与小学阶段的干预相比较,中学阶段干预则比较缺乏。DBDM 实施过程保真度调查也较为少见,数据驱动决策干预效应因此无法有力支撑因果推论,限制了

其进一步扩展推广的可能性。由于研究者发表偏倚原因,在相关综述当中也较难看到无影响及负面影响的研究个案。

从数据干预的作用群体来看,虽然 DBDM 旨在最大限度地提高所有学生的成绩,但针对不同层次、经济地位和年级的学生干预效果差异较为明显。相关研究显示效应大小与学习者初始水平有关,学习困难的学生往往从干预措施中受益最大,这可能是由于教师更关注学习表现特别典型的学生^[20]。数据使用的互动效应也表明,对于社会经济地位较低的学生比例较高的学校来说,数据干预效果也较为明显^[21],因为相对于经济地位高的学生,他们更需要通过学校的实践项目和辅助资源获得帮助。此外,干预效果也并非对

所有年级都同样显著,高年级又比低年级段的干预效果大^[21],可能与其是否能够理解和接受干预措施有关。当前 DBDM 实施效果研究主要集中在学校、班级和小组层面,后续应加强针对个别学生的干预研究。

干预规模与效应大小也存在相关关系,随着学生数量的增多,效应随之逐渐减少^[23]。因为随着干预规模扩大时,相关干扰和随机因素逐渐增多,而这些因素并不是都可以精准控制,这也导致数据应用预设与现实情况之间差距的主要原因。从数据使用的主题来看,现有的 DBDM 干预主要集中在“以学习者为主体”的数学、阅读理解、词汇或拼写成绩的影响上,而对教师的知识、技能和态度的影响较少探讨。为了拓宽我们对 DBDM 影响的理解,有必要对多主体和多学科的 DBDM 效应进行全面研究。

(二)数据干预有效作用周期

对于数据驱动的最佳干预周期,并没有定论,从上述文献统计数据来看,大多数数据应用干预的时间都在7周以上,这也是适用标准决策规则时,可以进行教学调整和评估的最短时间。对于学生而言,增加干预的总时间或使用更个性化的教学方法支持学生学习来加强干预,不会有助于阅读成绩的增加,仅靠频繁的进度监控(没有教学修改)似乎并不能提高学生的成绩^[15]。但对教师而言,持续时间是有效专业发展的一个重要特征,较长的教师专业发展干预有助于教师和学生的学习。因为学习和改变实践需要时间。改变教师的行为并不容易,他们在工作面临许多任务,根据 Supovitz 和 Turner 的观点,改变教师行为需要80个小时^[24]。而 DBDM 要求所需的主题内容知识、教学知识和技能以及所需的评估和测试能力太多,教师无法在如此短的时间内掌握。因此,建议向教师提供持续的教师专业发展数据,而不是传统的一次性教师专业发展干预。

现实中教学数据应用干预效果不明显或是无效可能只是暂时的,干预的影响往往需要持续一段时间才能出现。相关研究显示数据干预影响是在一个学年后发现的^[25-26]。在数据干预项目中,教师需要学习和实践一套复杂的数据素养,干预项目完成后,他们可能仍然需要更多的时间在教学中使用数据。因此,干预后不能马上看到对学生成绩的影响。这表明持续时间在数据使用过程中的重要性,以及需要有足够时间对学生成绩进行现实评估。此外,学校教师和学校管理人员流动性也会 DBDM 持续发展产生一定影响。建议采用全校整体参与的方法,规避人员变动风险,使学校工作人员能够相互支持实施 DBDM,并促进成

为学校的教学常规举措,而非个人行为。

数据干预的持续成效需要教师教学内容知识的介入,干预存在一定的上限效应,即学生学习成绩在最初的增长后趋于平稳,只有当数据使用与教师教学内容知识相结合时,才会取得更加可观的增长^[27]。因为数据干预可以带来一定的收益,之后需要专注于所分析的内容领域(如识字、算术),当数据与内容领域的知识一起进行时,数据干预的效力才是最佳的。

(三)数据干预有效措施

从干预形式来看,当前教育数据应用的干预主要包括评估、形成性评价、课程本位测量、基于数据的个性化和数据驱动的决策,虽然名称不同,但相关概念之间存在着一定继承关系,其本质原理都是以学生成绩数据为基础,通过分析告知教师学生的个人需求,以此制定教学的基础,从而系统地致力于教育改进。其中 DBDM 是对数据应用规律不断深化所延伸出的概念,与其他干预方式仅强调数据分析过程相比,它更强调依据数据分析作出决策和反馈调整的过程,即在不同的时间为不同的学生调整教学,或者确定哪些变化对不同类型的学习者构成潜在合理的调整机会,这些决策过程是数据干预对教学现实产生影响的最为重要的方式,也是教师专业发展在数据应用领域面临的最大挑战。

从数据干预过程来看,DBDM 是作为一个系统程序来实施的,包括评估和分析结果,设定具体、可衡量、可实现、相关和及时目标,确定目标实现战略以及最终执行目标实现战略四个组成部分^[28],它们相互关联,且执行过程具有一致性和连贯性要求。因此,为了使数据干预措施有效,它必须包括完整的 DBDM 组件,即在分析和评估数据步骤之后,教师必须要有后续的跟进步骤,才能在整个 DBDM 周期中实现其意义。教师通过基于数据分析获得的见解,设定具有挑战性的目标,选择并执行合适的策略来实现目标。数据用于监测和评估所实施策略的效果和评估目标实现的程度,以及作出新的基于数据的决定。虽然识别和分析数据为有效改进学生学习奠定了基础,但由此产生的行动和进度监控将最终决定 DBDM 干预的成效^[29]。

从干预的精细和深入程度来看,数据干预既要考虑学校层面的组织,也要考虑教师层面的实施。相关研究显示,学校或学区积极推广 DBDM,但教师们可能对使用数据毫无准备。即使他们学会了如何分析和解释数据,也没有改变课堂实践的意愿^[30]。当建议使用学生数据来改进教学时,教育政策的倡议往往过于

笼统和抽象,没有以清晰和详细的方式描述期望的变革行为,干预内容和教育实践之间缺少明确的联系,也就不可能据此系统地培训教育者。

(四)数据干预有效路径

教育领域中“数据路径”指数据在学校中的传播和作用路线,主要研究不同类型数据如何在不同级别进行流转以及不同层次的参与者如何就数据的使用进行相互交流。数据路径受学校文化以及负责人领导风格的影响,在相同数据干预框架下,数据干预的路径不尽相同,数据应用效果也呈现出一定的复杂性和异质性,这也凸显了我们对数据使用过程进行深入探究的必要性,其内容应主要包括使用了什么类型的数据,数据在系统的不同层次(地区、学校和课堂)具体是如何使用的,以及这些层次是如何相互作用的。

首先,从数据干预的目的来看,开展数据驱动教学决策的教育工作者可以根据数据采取的行动主要分为三种不同的类别:问责、指导教学和学校发展的行动,但它们之间并不是相互排斥的关系,用于问责目的的数据行动也可用于学校发展目的。此外,用于教学目的以及学校发展目的的数据行动也都可以带来学校改善。这些使用方式的区分更多体现的是数据聚集的程度的差别。如教师可以使用个人学生成绩数据在课堂中设置目标,学校领导也可以使用学校级别的学生成绩数据在学校层面设置目标。学校应根据这些不同成员角色设立相应层级化目标,以有效指引数据使用程序^[31]。

其次,先进和复杂的数据使用方式并不总是能给学校带来改进,学校应根据自身学校文化、学校领导方法、教师发展水平、学校环境和运作水平等因素选择最为适合的数据使用路径。Archer 等研究发现,学校的数据路径主要包括团队、级联和自上而下三种方式^[32]。不同级别参与者以不同方式使用不同类型的数据。例如,学校领导可能对学校的整体成绩水平感兴趣,以此作为决策的基础。教师可能会对学生的成绩水平更感兴趣,从而根据这类数据改变教学策略。从教师到负责人,数据分析、解释和应用过程中通常涉及多个角色,学校应为不同的数据角色设定相应的数据访问许可权限,提供的数据也必须满足不同人员在不同数据使用复杂程度上的个别化需求,这些数据也被用来促进不同年级和角色教师就课程和个性化支持计划进行对话。

四、数据驱动的教学决策效果研究的新视角

传统的 DBDM 研究主要从理论上将数据干预的

作用过程看作是一个由数据直接施加影响的线性作用过程,而没有从实践应用的实际情境中对影响数据应用的关键因素和关键环节进行重点审视,更没有从构建数据应用规律和机制的角度去探讨稳定长远提升 DBDM 效果的可能性,需要从以下视角进一步研究和实践。

(一)重视数据干预影响效果研究,抓住教师这一关键因素

数据使用过程不是孤立发生的,而是一个复杂的非线性过程。数据使用涉及人为解释过程,教师必须获取、收集和分析数据,才能将其转化为信息,并与理解和专业相结合,成为有意义和有用的行动基础^[1]。教育数据干预领域的长期实践证明,教师是教学改进的主导因素,只有改变教师才能改变学生,这是学校改善必要的第一步^[33]。考虑到具体现实情境中基础资源基本不变的前提下,教师个体是唯一调整可变的空空间,教师因素也是影响数据干预效果最重要的因素,包括教师知识、技能和态度等在内的教师数据素养整体状况决定了教育数据使用的最终成效,而通过各种教师专业发展措施(团队会议、教学教练、教学培训、教学辅导等)提升教师数据素养也是最直接、经济的提升数据干预效果的途径^[11]。因此教师专业发展应特别重视那些旨在提高教师和学校领导提供反馈和根据学生评价数据调整教学的能力的环节,强调使用数据与实践行动相互联系的重要性,将数据驱动教学决策能力培育作为教师专业发展方向^[34]。应倾向于解决诸如制定研究问题、解释结果以及有效开发和课堂评估等技能,并且为教育工作者提供讨论数据和使用真实数据的机会,而不是假设的案例和问题。为了提高学生的成绩,教育者需要实际运用他们的知识和技能来使用数据,使之成为专业发展的一部分。当教师能够有意识地选择满足学生需求的策略时,学生的成绩就有可能得到提高。

(二)探究数据驱动教学决策规律,提升数据应用效能

决策环节是数据干预的最终应用环节,也是决定其是否能够发挥效用的关键环节。在决策进度控制方面,大多数研究使用嵌入课程的措施^[35]、诊断图^[36]、策略使用规则^[37]等决定何时以及如何调整教学。在 DBDM 团队行动控制方面,Crone 建议 DBDM 团队重点考虑一些重要问题,以确定团队是否达到了关键目标,包括:作出了哪些决定,这些决定能够提高学生的成绩吗?是否有作出决定所需的所有相关数据?决策的实施过程是否明确(如指定人员、设定目标、确定时间

表)?在多大程度上满足了最高优先级的学生的需求?数据团队通过定期自我评估,形成 DBDM 建议框架中规定的迭代过程^[38]。在决策模型方面,目前已经有众多学者提出了相关决策概念模型,如 DBDM 组件模型^[25]、DBDM 八个步骤^[32]、教师数据素养^[39]和多成分模型^[40]等。由于决策涉及不同的时间尺度,支持决策所需的证据应该与其时间尺度相关。一些决策需要在教学实践进行过程中提供信息的数据;而另一些决策,如资源分配或规划专业发展,发生在较慢的时间尺度上,则可以从更大粒度的数据中受益。未来的 DBDM 研究不仅需要描述数据用于指导决策的过程,还要详细说明决策过程中使用了什么数据,使用了什么规则来作出决策,多久收集一次数据,多久使用一次数据作出决策。加强数据驱动教学应用规律的探究,对决策进程的持续关注和对干预保真度的跟踪,提供深入的决策细节描述,最终提升 DBDM 应用效率和效力。

(三)扩展数据作用效果分析思路,构建评估核心机制

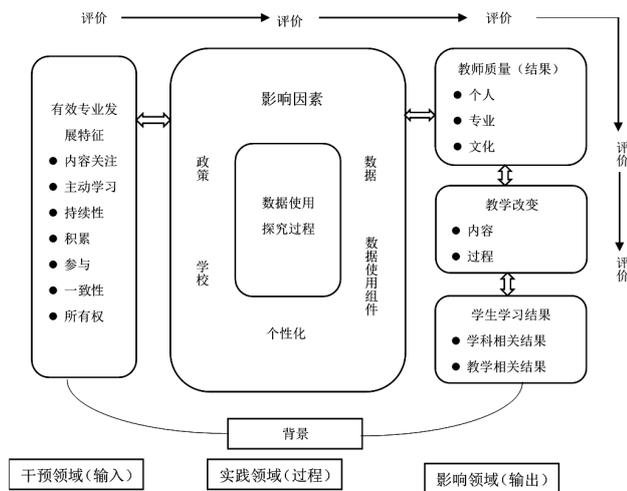


图3 评估数据使用专业发展干预对教师和学生结果的影响的一般框架^[41]

当前的数据干预效果验证思路,主要是通过数据干预随机对照组实验效应大小或是数据使用过程质性描述的方式来开展的。实验研究关注最终效应大小,缺少针对数据干预过程和机制的详细描述,质性研究又因研究者通常存在社会期望效应,在结论推广上缺乏说服力。因此,有必要突破现有的数据作用效果分析思路,基于一般的数据使用通用路径提出综合性的评估框架,来对数据干预的效果进行全面评估。Merchie 等通过研究教师专业发展干预对学生影响的一般过程,认为教师变量、数据使用实践和教学条件真正相关,提出了一个包括干预(输入)、数据使用的

实践(过程)以及最终的结果(输出)的通用路径模型(如图3所示),用于评估专业发展干预对教师成果、课堂实践和学生成果的影响^[41]。但在某些现实情况下,很难评估数据使用项目的实际效果,因为不同的数据使用流程决定了不同的数据绩效模型,不同的数据干预之间很难就数据使用绩效评估框架达成一致。因此,有效性问题的重点将需要再次转移到更广泛的学习或组织环境,以及它在多大程度上有足够的资源来支持基于证据的专业实践,以加强学生的学习。未来的研究应侧重检查与成就效益相关的不同项目开发过程,以确定数据使用过程的哪些方面似乎与更高的效益最相关。

(四)改变数据应用短视行为,致力长远发展目标

目前 DBDM 应用过程中普遍的短视行为表现为对数据的来源渠道和最终目标的狭隘界定,特别是由于不同教育背景下的问责压力,导致教学领域大多数时候只关注学生的考试成绩和分数,并且片面地以成绩和分数数据作为依据,从而作出较为肤浅的教学理论假设。虽然以学生成绩数据作为数据源符合教学实践领域的便利性原则,但基于成绩和分数的来源数据,教师无法做到全面客观地评价学生。数据干预的短视还表现在比较关注干预的当下效应,长期是否能够促成学校数据文化的形成,数据团队是否会继续使用数据,如何能够一直延续数据干预的有效性等问题亟待进一步关注,也是通过数据应用促进学校师生长远发展必须解决的难点。一般而言,数据干预项目结束后,学校的数据文化很难维持下去,学生成绩提升也得不到持续保障。此外,多数数据干预项目并没有评估在取消研究者支持后教师继续参与 DBDM 的程度,也没有评估在不太理想的条件下 DBDM 对教师的影响,因此也就无法得出关于数据应用的一般性应用结论。未来的研究应尽量在自然教学情境中,检验 DBDM 数据干预行动给师生带来的实际影响,并应在研究完成后对 DBDM 继续使用状况进行后续跟踪评估。通过研究这些背景下 DBDM 行动促进教师专业发展和提升学生成绩的经验 and 规律,研究人员可据此构建基于真实自然情境下的数据干预理论,从而真正实现常规教学的精准指导。

五、结 语

近年来,数据驱动的教学决策的理论和实践研究不断增长,数据驱动教学决策理念和行动在世界教育范围内蔚然成风,利用数据来改善学校教育实践,已经成为一种流行的学校变革趋势,并为探究教育教学

活动规律,提升教育教学质量提供了新视角、新动能和新路径。但从数据到决策不是一蹴而就的,现实教育决策难题也不能简单转嫁数据驱动解决,数据干预应用效果需要审慎评估,盲目而动、急功近利、缺乏有效评估的数据驱动极易导致教育教学陷入“数据泥

沼”。本研究围绕数据驱动教学决策效果,从作用范畴、周期、措施和路径等角度进行了系统性再回顾,提出了数据驱动教学决策的主要影响因素、决策方法、评估机制和长远目标等关键议题,以期对相关研究提供新的参考和帮助。

[参考文献]

- [1] COBURN C E, TURNER E O. The practice of data use: an introduction[J]. *American journal of education*, 2012, 118(2): 99-111.
- [2] MARSH J A, FARRELL C C. How leaders can support teachers with data-driven decision making: a framework for understanding capacity building[J]. *Educational management administration & leadership*, 2015, 43(2): 269-289.
- [3] 彭红超,祝智庭.面向智慧学习的精准教学活动生成性设计[J].*电化教育研究*,2016,37(8):53-62.
- [4] 郭利明,杨现民,张瑶.数据驱动的精准教学五维支持服务框架设计与实践研究[J].*电化教育研究*,2021,42(4):85-92.
- [5] 万力勇,黄志芳,黄焕.大数据驱动的精准教学:操作框架与实施路径[J].*现代教育技术*,2019,29(1):31-37.
- [6] PAPATHEODOROU S. Umbrella reviews: what they are and why we need them[J]. *European journal of epidemiology*, 2019, 34(6): 543-546.
- [7] HENNESSY E A, JOHNSON B T, KEENAN C. Best practice guidelines and essential methodological steps to conduct rigorous and systematic meta-reviews[J]. *Applied psychology: health and well-being*, 2019, 11(3): 353-381.
- [8] MOHER D, LIBERATI A, TETZLAFF J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement[J]. *Plos medicine*, 2009, 6(7):1-6.
- [9] SCHOLIN S E, BURNS M K. Relationship between pre-intervention data and post - intervention reading fluency and growth: a meta-analysis of assessment data for individual students[J]. *Psychology in the schools*, 2012, 49(4): 385-398.
- [10] ANSARI M F, GROOT W, DE WITTE K. Tracking the process of data use professional development interventions for instructional improvement: a systematic literature review[J]. *Educational research review*, 2020,31: 100362.
- [11] JUNG P G, MCMASTER K L, KUNKEL A K, et al. Effects of data-based individualization for students with intensive learning needs: a meta-analysis[J]. *Learning disabilities research & practice*, 2018, 33(3): 144-155.
- [12] PETERS M T, FÖRSTER N, HEBBECKER K, FORTHMANN B, SOUVIGNIER E. Effects of data-based decision-making on low-performing readers in general education classrooms: cumulative evidence from six intervention studies [J]. *Journal of learning disabilities*, 2021,54(5):334-348.
- [13] VISSCHER A J. On the value of data-based decision making in education: the evidence from six intervention studies [J]. *Studies in educational evaluation*, 2020,69:100899.
- [14] GESEL S A, LEJEUNE L M, CHOW J C, et al. A meta-analysis of the impact of professional development on teachers' knowledge, skill, and self-efficacy in data-based decision-making[J]. *Journal of learning disabilities*, 2021, 54(4):269-283.
- [15] STECKER P M, FUCHS L S, FUCHS D. Using curriculum-based measurement to improve student achievement: review of research [J]. *Psychology in the schools*, 2005, 42(8): 795-819.
- [16] HENDRIKS M A, SCHEERENS J, SLEEGERS P. Effects of evaluation and assessment on student achievement: a review and meta-analysis[J]. *The influence of school size, leadership, evaluation, and time on student outcomes*, 2014: 127-174.
- [17] FUCHS L S, FUCHS D. Effects of systematic formative evaluation: a meta-analysis.[J]. *Except child*, 1986, 53(3):199-208.
- [18] FILDERMAN M J, TOSTE J R, DIDION L A, et al. Data-based decision making in reading interventions: a synthesis and meta-analysis of the effects for struggling readers[J]. *The journal of special education*, 2018, 52(3): 174-187.
- [19] SHEA B J, BOUTER L M, PETERSON J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR) [J]. *Plos one*, 2007, 2(12): e1350.
- [20] ESPIN C A, WAYMAN M M, DENO S L, et al. Data-based decision-making: developing a method for capturing teachers' understanding of CBM graphs[J]. *Learning disabilities research & practice*, 2017, 32(1): 8-21.
- [21] VAN GEEL M, KEUNING T, VISSCHER A J, et al. Assessing the effects of a school-wide data-based decision-making

- intervention on student achievement growth in primary schools[J]. *American educational research journal*, 2016, 53(2): 360–394.
- [22] KONSTANTOPOULOS S, MILLER S R, VAN DER PLOEG A. The impact of Indiana’s system of interim assessments on mathematics and reading achievement[J]. *Educational evaluation and policy analysis*, 2013, 35(4): 481–499.
- [23] CHEUNG A C K, SLAVIN R E. How methodological features affect effect sizes in education [J]. *Educational researcher*, 2016, 45(5): 283–292.
- [24] SUPOVITZ J A, TURNER H M. The effects of professional development on science teaching practices and classroom culture[J]. *Journal of research in science teaching: the official journal of the national association for research in science teaching*, 2000, 37(9): 963–980.
- [25] VAN DER SCHEER E A, VISSCHER A J. Effects of a data-based decision-making intervention for teachers on students’ mathematical achievement[J]. *Journal of teacher education*, 2018, 69(3): 307–320.
- [26] SLAVIN R E, CHEUNG A, HOLMES G C, et al. Effects of a data-driven district reform model on state assessment outcomes[J]. *American educational research journal*, 2013, 50(2): 371–396.
- [27] LAI M K, WILSON A, MCNAUGHTON S, et al. Improving achievement in secondary schools: impact of a literacy project on reading comprehension and secondary school qualifications[J]. *Reading research quarterly*, 2014, 49(3): 305–334.
- [28] STAMAN L L, TIMMERMANS A C A, VISSCHER A J A. Effects of a data-based decision making intervention on student achievement[J]. *Studies in educational evaluation*, 2017, 55: 58–67.
- [29] KAUFMAN T E, GRAHAM C R, PICCIANO A G, et al. Data-driven decision making in the K–12 classroom [M]//*Handbook of research on educational communications and technology*. New York: Springer, 2014: 337–346.
- [30] SCHILDKAMP K, KUIPER W. Data-informed curriculum reform: which data, what purposes, and promoting and hindering factors [J]. *Teaching and teacher education*, 2010, 26(3): 482–496.
- [31] EBBELER J, POORTMAN C L, SCHILDKAMP K, et al. The effects of a data use intervention on educators’ satisfaction and data literacy[J]. *Educational assessment, evaluation and accountability*, 2017, 29(1): 83–105.
- [32] ARCHER E, SCHERMAN V, HOWIE S. Approaches to effective data use: does one size fit all?[M]//*Data-based decision making in education*. Springer, Dordrecht, 2013: 91–112.
- [33] EVANS A. No child left behind and the quest for educational equity: the role of teachers’ collective sense of efficacy [J]. *Leadership and policy in schools*, 2009, 8(1): 64–91.
- [34] 张学波, 林书兵, 孙元香. 从数据到知识: 数据驱动教学决策的理论模型与能力提升[J]. *电化教育研究*, 2021, 42(12): 41–47.
- [35] COYNE M D, SIMMONS D C, HAGAN-BURKE S, et al. Adjusting beginning reading intervention based on student performance: an experimental evaluation[J]. *Exceptional children*, 2013, 80(1): 25–44.
- [36] VERNON-FEAGANS L, GALLAGHER K, GINSBERG M C, et al. A diagnostic teaching intervention for classroom teachers: helping struggling readers in early elementary school[J]. *Learning disabilities research & practice*, 2010, 25(4): 183–193.
- [37] ORTLIEB E, MCDOWELL F D. Investigating the effects of an experimental approach to comprehension instruction within a literacy clinic[J]. *Current issues in education*, 2016, 19(1): 1–15.
- [38] CRONE D A, STOOLMILLER M, BAKER S K, et al. The middle school intervention project: use of a regression discontinuity design to evaluate a multi-component intervention for struggling readers in middle school in six school districts [R]. Oregon: Society for Research on Educational Effectiveness, 2012.
- [39] REEVES T D, CHIANG J L. Online interventions to promote teacher data-driven decision making: optimizing design to maximize impact[J]. *Studies in educational evaluation*, 2018, 59: 256–269.
- [40] VAN KUIJK M F, DEUNK M I, BOSKER R J, et al. Goals, data use, and instruction: the effect of a teacher professional development program on reading achievement[J]. *School effectiveness and school improvement*, 2016, 27(2): 135–156.
- [41] MERCHIE E, TUYTENS M, DEVOS G, et al. Evaluating teachers’ professional development initiatives: towards an extended evaluative framework[J]. *Research papers in education*, 2018, 33(2): 143–168.

Overview of Effect of Data-driven Instructional Decision-making and Vision Breakthrough —An Umbrella Evaluation Based on Ten Meta-Analysis

LIU Jin¹, ZHANG Xuebo², LIN Shubing³

(1.Office of Information Construction and Management, South China Normal University, Guangzhou Guangdong 510631; 2.School of Information Technology in Education, South China Normal University, Guangzhou Guangdong 510631; 3.Institute of Advanced Studies in Humanities and Social Sciences, Beijing Normal University, Zhuhai Guangdong 519087)

[Abstract] The application of data-driven instructional decision-making, as a new trend in current basic education practice, still has some problems, such as insufficient evidence, unclear effects, and unclear paths. In other words, what is the scope of instructional data interventions, what are the actual effects of data use in schools, how do schools evaluate the actual effects, what are the main factors that constrain the use of data, and how do schools build the practical mechanism for the use of educational data. This series of questions reflects the main confusion in the application of educational data, and also foreshadows the direction of further exploration of data-driven instructional decision-making. Given the growing number but limited perspective and scope of relevant thematic meta-analyses, this study adopts an umbrella evaluation approach, extracts ten studies based on meta-analysis and systematically reviews the effects of data-driven interventions for instructional applications from the perspectives of scope, cycle, measures, and paths. The key topics such as the main influencing factors, decision-making methods, evaluation mechanisms and long-term goals of data-driven are also proposed, with a view to providing comprehensive analysis conclusions and future breakthrough directions for subsequent related studies.

[Keywords] Data-driven Instructional Decision-making; Effect Evaluation; Umbrella Evaluation