

# 不同教师群体对教育大数据的认知及影响因素

——基于全国5434名教师的调查

王学男

(中国教育科学研究院 教育发展与改革研究所, 北京 100088)

**[摘要]** 教师作为大数据与教育、教学、管理的交汇主体,其重要作用不言而喻。本研究通过SPSS25.0统计分析软件对全国5434名教师对教育大数据的认知和影响因素展开调查,并进行描述性分析、方差分析、皮尔逊相关分析、多元回归分析。结果显示:被调查教师群体对大数据特征的理解主要集中于“数量大、类型多、速度快”,东、中、西部教师没有显著差异;不同教龄、家校距离、区域(东、中、西部)、任教学校类型的教师群体对教育大数据的积极性存在显著差异;教师工作时间、接触数据权限、教学需求、学校软件配套等外部条件的局限对教师应用教育大数据有负面影响;评价方式、教师态度、学科适用、学校投入、传播效果能正向预测教师对教育大数据的积极性。最后,文章从政策实施、理论研究、培训设计和学校执行四方面提出了建议。

**[关键词]** 教师;教育大数据;认知;应用现状;影响因素

**[中图分类号]** G434 **[文献标识码]** A **[文章编号]** 1007-2179(2019)03-0081-11

通过教育大数据推进教育信息化、智能化,是教育信息化2.0持续发展的必要条件。教育大数据是在教育教学和校园活动中产生的(王正青等,2018a),并根据教育需要采集的一切用于教育发展并可创造巨大潜在价值的数据集(杨现民等,2015),它为提升教育质量、优化资源配置、实现个性化学习、推动科学决策提供了重大机遇。通过数据驱动教学变革、提高教育质量、借助资源共享促进教育公平,教师是关键。因此,本文以“教育大数据”为核心概念展开研究。为避免对数据规模的争议,特在问卷调查中将其具体化为可操作的概念,以便于对调查对象的统一理解。

在教育大数据不断深化和应用下沉的过程中,

教育信息化硬件建设已取得巨大成就,但在教学方面,无论是在教室还是互联网上,教师的教学模式从内容、手段、教育理念到教育方法都没有明显变化(杨现民等,2016)。如舍恩伯格(Schönberger)所说:“生活在两三个世纪前的人物走进今天的教室,可能会说:没有什么变化——即使外面的世界早已变得几乎面目全非。”(舍恩伯格等,2015)教育技术专业总希望使用信息技术改变以往的学习形态,但“滥用”信息技术的现象,恰恰破坏了学习规律的学习生态(李芒等,2017)。教育大数据应用面临两难困境:一方面,高呼信息技术改变教育,教育信息化应用推广太慢;另一方面,批判技术过度使用,违背教育规律。信息技术对教育的影响度,实质

[收稿日期]2019-03-28

[修回日期]2019-04-27

[DOI编码]10.13966/j.cnki.kfjyyj.2019.03.009

[基金项目]北京市教育科学“十三五”规划优先关注课题“大数据在提高学校教育质量方面的应用研究”(CEHA17067)。

[作者简介]王学男,博士,助理研究员,中国教育科学研究院教育发展与改革研究所,研究方向:教育政策、科学教育(bnuwxn@163.com)。

上不能以信息技术使用的数量为准则,而应以使用效果为最高标准。因此,本研究以学校一线教师为研究对象,切实了解教师对教育大数据的认知和应用现状并探究影响因素。

## 一、文献综述

教育大数据能洞察每个真实的学生,因此将成为学校最重要的资产(杨现民等,2016)。但真正利用大数据解决教育问题的实践活动和成功案例并不多。如何将大数据有机地与教学实践和教学管理结合起来,还需要方法论方面的探索(祝智庭等,2017)。纵观相关研究,国外研究已从政策倡导、理论模型构建深入到教育实践和操作应用层面,注重从师生视角对教育大数据进行实证研究;国内研究处于政策先行、学者引领、理论阐释及本土化阶段,主要集中于教师数据素养标准的构建,较少涉及教学关系中教师主体对教育大数据认知与应用的影响和主观能动性的实践研究。

美国教育部 2007 年发布的《教师使用学生数据系统改进教学》报告指出,仅靠政策颁布不足以推进教育大数据的教育实践,需要切实了解阻碍相关政策执行的原因,针对教师的大数据使用情况进行全美国范围调查结果表明,教师对数据系统的来源、应用、成效均存在不同程度的问题,需有的放矢进行培训(Means et al., 2008)。伊耶(Iyer, 2008)也强调教师作为教育大数据利益相关群体,对教育大数据的应用需求应给予重视。教师作为学校教育潜在的数据分析师,应具备为学生提供有效的基于数据分析的问题解决和个人发展的技能和敏感性(Joshua, 2017)。2014 年美国教育部发布的《通过数据挖掘与学习分析改进教与学》报告指出,对大数据的运用主要是创造了“学习分析系统”——一个数据挖掘、模化和案例运用联合框架,旨在向教育工作者提供了解学生“怎样”学习得更多、更好、更精确的信息(Marie et al., 2012)。海蒂(Hattie, 2009)通过综合分析史上最大量的数据指出,对学业成就影响最大的因素是教师;教师需注意每个学生所思所知,并以此建构意义和有价值的经验,精通他们所教的内容并进行深度理解,提供有意义和适当的反馈,使每个学生通过不同水平的课程学习获得持续进步。通过数据挖掘实现教育大数据帮助解

决教育问题,首先要明确教育数据挖掘涉及的目标用户和参与者群体,以及每个群体使用数据信息的目的。对教师而言,大数据可揭示学生的特点,预测未来的表现,提供教学帮助(Yu et al., 2016)。这为我国从教师视角开展教育大数据研究提供了启发。

我国已有相关研究聚焦教师数据素养的概念界定、模型构建、指标体系、评价标准等,指出数据素养是教师在接触教育数据时所体现出来的一种综合能力,包括数据意识与态度、数据基础知识、数据核心技能以及数据思维方法四方面(刘雅馨等,2018)。教师数据素养评价的一级指标包括数据知识、数据技能、教学应用、道德意识(李青等,2018)。王正青等(2018a, 2018b)通过分析教育大数据在美国教育决策、学校管理、教学创新、教育科研等方面的广泛应用及其实施与推进的策略,建议培育数据至上的学校文化与教学传统,出台发展教师数据素养的政策与规划,建立支撑教师使用教育数据的纵向数据库,开发发展教师数据素养的资源与工具,并通过在教师专业标准中纳入数据素养要求,在教师职前培养阶段增加数据素养训练,在日常教育教学中推广使用教育数据,在职后教师专业发展中强化数据素养培训等方式加以落实。《互联网+教育蓝皮书:中国基础教育大数据》和《中国基础教育大数据 2016-2017:走向数据驱动的精准教学》年度蓝皮书,分别对我国基础教育大数据进行了全景式与专题式论述。前者勾勒出教育大数据在基础教育应用中的五大模式:教育政策科学化、教育评价体系重构、区域教育均衡发展、学校教育质量提升和师生个性化发展(杨现民等,2016);后者针对数据驱动的教学变革给教师提出五条建议,包括培养使用数据的习惯、就学生数据进行家校沟通、对大数据改善教学及实现工作减负保持积极心态、加强数据科学的学习及数据处理能力训练、结合学科特点等积极开展数据驱动教学的实践探索或大数据在教学中的应用策略(杨现民等,2018)。

然而,少有研究关注教师的教育大数据需求。目前仅 2 篇论文从认识、有用性和易用性等维度展开调查,发现绝大多数中小学教师非常期望能利用大数据支持的学习分析与评价方法和技术对学生学习信息实现全方位和全程化收集、整理与分析,从而系统评价学生的学习过程与结果,同时也期望得到

相关的培训与指导(高巍等,2018)。采集、处理、分析和应用传统课堂的线下数据与互联网平台的线上数据所构成的学生学习行为的教育大数据,能为教师的教学评价与教学指导提供依据(郑燕林等,2015),但都存在样本量较小(小于200)的局限。

## 二、研究设计与方法

### (一) 研究问题与假设

本研究旨在了解学校教师对教育大数据的认知和应用现状,探讨影响其应用教育大数据的因素(学校投入、教育大数据适用的评价方式、教育大数据理念与应用的传播效果、教师的开放态度、教学或学科适用性),研究问题包括:

- 1) 教师对教育大数据的认知和应用现状如何?
- 2) 不同教师群体对教育大数据应用积极性的差异如何?
- 3) 评价方式、教师态度、学科适用、学校投入、传播效果与教师对应用教育大数据的积极性有什么关系?

基于上述三个基本问题,本研究提出如下研究假设:

- 1) 教师对教育大数据的认知和应用达到普及水平;
- 2) 不同性别、民族、教龄、职称、毕业专业类型(是否属于师范类专业)、主要上网途径、家校距离、任教学校类型、区域(东部、中部、西部)的教师群体应用教育大数据的积极性存在显著差异;
- 3) 评价方式可操作性越强、教师开放态度越高、教育大数据的传播越广泛有效、学校投入越充分、学科适用越强,教师应用教育大数据的态度越积极。

### (二) 研究工具

本研究采用定量的实证研究方法,问卷是主要的数据收集工具,分两大部分:第一部分是教师对教育大数据的认知和应用,涉及被调查者的社会人口统计学等客观变量;第二部分结合美国《教师使用学生数据系统改进教学》报告内容(Teachers' use of student data systems to improve instruction),依据我国教育大数据在学校应用的实际修订而成,主要分教师应用教育大数据的积极性及其影响因素(见图1)。第二部分包括15个题目,每道题采用李克特

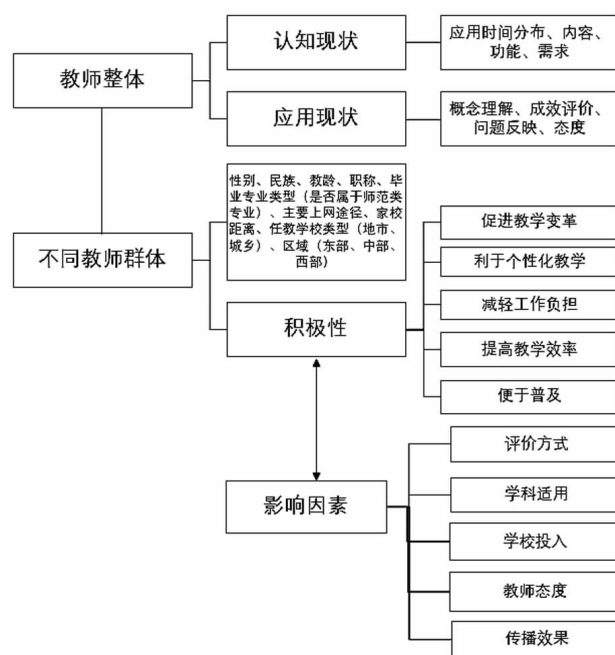


图1 研究框架

(Likert)五级量表法表示,即1-5分,从“很不符合”到“非常符合”。第二部分问卷整体信度Alpha值为0.821,表示问卷整体可靠性较强;结构效度分析结果显示,KMO值为0.869,高于0.8,说明问卷结构效度高。本研究先在北京市、贵州省、深圳市各选取20名教师进行试测,并对问卷进行区分度、信度和效度检验,删除和修改部分题目后向全国开放施测,进行正式调查。

### (三) 研究对象及数据采集

本研究使用网络问卷方式共发放问卷7000份,实际回收5876份,回收率为83.94%。其中,有效问卷为5434份,有效率为92.48%。本研究通过重点委托<sup>①</sup>北京、广东、重庆、贵州、云南、甘肃、天津、青海、陕西、内蒙古、浙江11个省、市、自治区教育部门,按小学:初中2:1的比例取样,同时通过滚雪球取样的方式面向全国30个省、自治区、直辖市<sup>②</sup>的中小学教师开展全面调查。其中,来自东部地区的教师有1999人,占36.79%,中部地区教师1250人,占23%,西部地区教师2185人,占40.21%。问卷调查时间从2018年7月20日至10月30日。参与调查教师的基本情况如表一所示。

### (四) 研究方法

本研究把教师应用教育大数据的积极性作为因

变量,教师群体特征作为控制变量,评价方式、教师态度、学科适用、学校投入、传播效果作为自变量,用 SPSS25.0 作为统计分析软件,统计学分析方法包括描述性分析、方差分析、皮尔逊相关分析、多元回归分析。描述性统计分析主要用于统计教师整体对教育大数据的认知和应用现状,方差分析主要用于发现不同教师群体对教育大数据应用积极性的差异程度,皮尔逊相关分析用于测量因变量和自变量之间的关联水平,多元回归分析用于发现教师应用教育大数据积极性的影响因素及其水平。

表一 参与调查教师的基本特征(单位:人)

|        |        |      |         |      |         |          |            |      |        |  |
|--------|--------|------|---------|------|---------|----------|------------|------|--------|--|
| 性别     | 男      |      |         |      |         | 女        |            |      |        |  |
|        | 1238   |      |         |      |         | 4196     |            |      |        |  |
| 年龄     | 30 岁以下 |      | 30-40 岁 |      | 41-50 岁 |          | 50 岁以上     |      |        |  |
|        | 2622   |      | 1296    |      | 1156    |          | 360        |      |        |  |
| 教龄     | 5 年以下  |      | 6-10 年  |      | 11-15 年 |          | 16-20 年    |      | 20 年以上 |  |
|        | 1680   |      | 1022    |      | 740     |          | 771        |      | 1221   |  |
| 户籍     | 农业户口   |      |         |      |         | 非农业户口    |            |      |        |  |
|        | 840    |      |         |      |         | 4594     |            |      |        |  |
| 毕业专业类型 | 师范类    |      |         |      |         | 非师范类     |            |      |        |  |
|        | 4402   |      |         |      |         | 1032     |            |      |        |  |
| 任教学段   | 幼儿园    |      | 小学      |      | 初中      |          | 高中         |      |        |  |
|        | 705    |      | 3035    |      | 1490    |          | 204        |      |        |  |
| 学历     | 大专     |      |         | 本科   |         |          | 研究生及以上     |      |        |  |
|        | 1096   |      |         | 3955 |         |          | 383        |      |        |  |
| 职称     | 未定级    | 小学一级 | 小学二级    | 小学三级 | 小学高级    | 中学初级(一级) | 中学中级(二、三级) | 中学高级 | 特级教师   |  |
|        | 793    | 1001 | 664     | 137  | 831     | 794      | 678        | 509  | 27     |  |

相应增加。幼儿园阶段的教师上课和班级管理的时间最长;小学和初中阶段的教师上课和备课的时间分别位列前两位;高中阶段的教师备课时间高于上课时间。可见,随着学段的升高,教师备课占用的工作时间会相应增加。从各学段教师“登录并使用学校(学生)数据系统”的均值和占比看,其数值最小(幼儿园除外,因为幼儿园无需批改作业),幼儿园、小学、初中、高中阶段的教师平均每天应用数据系统的时间分别为 0.38 小时、0.55 小时、0.45 小时和 0.56 小时,分别占工作时间的 3.27%、3.90%、3.08% 和 3.78%。由此可判断,教师常态的教学和管理中,教育大数据相关工作所占比例最小且孤立,数据驱动的教学、备课或者教研变革更难实现。

表二 教师教育教学活动时间构成(单位:小时/周)

| 教育教学活动          | 学段     |        |        |        |
|-----------------|--------|--------|--------|--------|
|                 | 幼儿园    | 小学     | 初中     | 高中     |
| 上课              | 12.102 | 19.011 | 17.371 | 11.52  |
| 备课              | 7.159  | 10.653 | 12.142 | 14.973 |
| 批改作业及试卷         | 1.67   | 9.168  | 9.091  | 9.758  |
| 管理及参加班级活动       | 12.825 | 6.973  | 7.596  | 9.409  |
| 对学生进行个别教育及辅导    | 4.943  | 4.359  | 4.453  | 6.058  |
| 业务进修            | 4.154  | 3.963  | 4.779  | 4.274  |
| 学校事务            | 5.43   | 5.615  | 6.162  | 8.063  |
| 学校教研相关工作        | 3.416  | 3.394  | 4.241  | 4.086  |
| 登陆并使用学校(学生)数据系统 | 1.918  | 2.758  | 2.255  | 2.807  |
| 其他工作            | 4.965  | 4.775  | 5.293  | 3.263  |
| 总计              | 58.582 | 70.669 | 73.383 | 74.211 |

### 三、数据分析与研究发现

#### (一)教师应用大数据现状的描述性统计

1. 不同学段教师在校教育教学活动的时间构成  
对不同学段教师工作时间均值进行调查和统计,可以间接了解他们应用教育大数据的时间构成特征。由于直接用“教育大数据”表述较抽象和笼统,因此选择“登录并使用学校(学生)数据系统”这一日常工作触手可及且概念理解较统一的表述开展调查。调查发现,随着学段的升高,教师工作时间会

#### 2. 教师在校使用大数据的时间、内容和功能

第一,39%的被调查教师“在校下班后加班时使用数据系统或者关注学校教育数据”,37.43%选择“在校临时挤出时间”,22.65%选择“在校专门预留出时间”。这说明超过七成的被调查教师在校工作时没有专门时间用于使用数据系统或者关注学校教育数据,也再次印证表二的调查结果,教师日常教学时间构成的客观条件,限制了教师对教育大数据的关注与应用。

第二,对于现有的学校教务系统或学生数据系

统中最常使用的功能,选择“查询”功能的占 49.02%,选择“通知”“记录”和“初步统计分析(如排序、计算百分比等)”功能的分别占 15.38%、13.36%和 12.68%,选择“比较”功能的占 6.59%,选择“高级统计分析(相关分析、建模等)”的占 1.75%。这也反映出教师对数据系统的应用仍处在初级、传统水平,未达到挖掘数据潜在价值层面。但这也可能源于学校数据系统的架构与设计存在局限,限制教师应用大数据的更多可能性。

第三,对于在什么情况下教师更愿意使用数据这一问题,选择“与同事讨论教学与管理时”占比最高,达 25.08%,其次分别是“与同事共同完成学校任务时”,占 24.14%;“独立工作时”,占 23.74%;“完成学区及以上行政部门的共同任务时”,占 12.94%;“指导学生时”,占 7.42%;“与家长沟通时”,占 5.93%。也就是说,与同事共同协作、完成学校工作任务时,教师更倾向于有数据支撑。

第四,对多选题“目前学校的教育大数据应用”和“最希望得到大数据支持”的响应率<sup>③</sup>表明,教师对教育大数据的应用,在供需两侧存在差距。虽然目前针对课堂教学的数据收集与分析能力已有所突破,以人脸识别、语音识别和在线学习行为分析为代表的教育大数据应用,可在一定程度通过智能分析帮助教师实时、客观、准确地了解学生学习行为并促

进有效的师生互动,但课堂教学的深度挖掘仍是教师最关注的。据图 2 可知,供大于需的六个方面数据分别是课程管理、教务管理、安全管理、考勤管理、教师考核、报表填报,这可能与新高考改革的“选课走班”有关;供小于需的六个方面数据主要是课堂教学数据、家校沟通、作业布置与批阅、学生自我评价与管理、家长对学校教育和师生教学过程的了解、学生的思维过程。由此可见,学校数据系统提供的功能主要是服务于学校管理,教师对教育大数据的关注和需求更集中于课堂教学和评价沟通。

### 3. 教师对大数据的基本认知

大数据的 4V 特征,已成为大众对大数据的基本共识。本研究通过多选题调查教师对这一概念的认知,发现教师对大数据的认知主要是“3V”特征,即数据量大(21.2%、67.6%)、数据类型多(20.5%、65.2%)、数据处理速度快(20.2%、64.3%),而对于数据价值、数据来源、存储等价值维度还未形成广泛共识。这与大数据从 3V 到 4V、6V 等的发展历程及人类对其认知的过程一致。

本研究进一步利用交叉分析考察东、中、西部省份的教师对大数据的理解,结果显示,东、中、西部区域差异与各选项的选择没有明显关联,百分比选择分布相对均匀。也就是说,教师对大数据的理解和认知不存在区域差异。

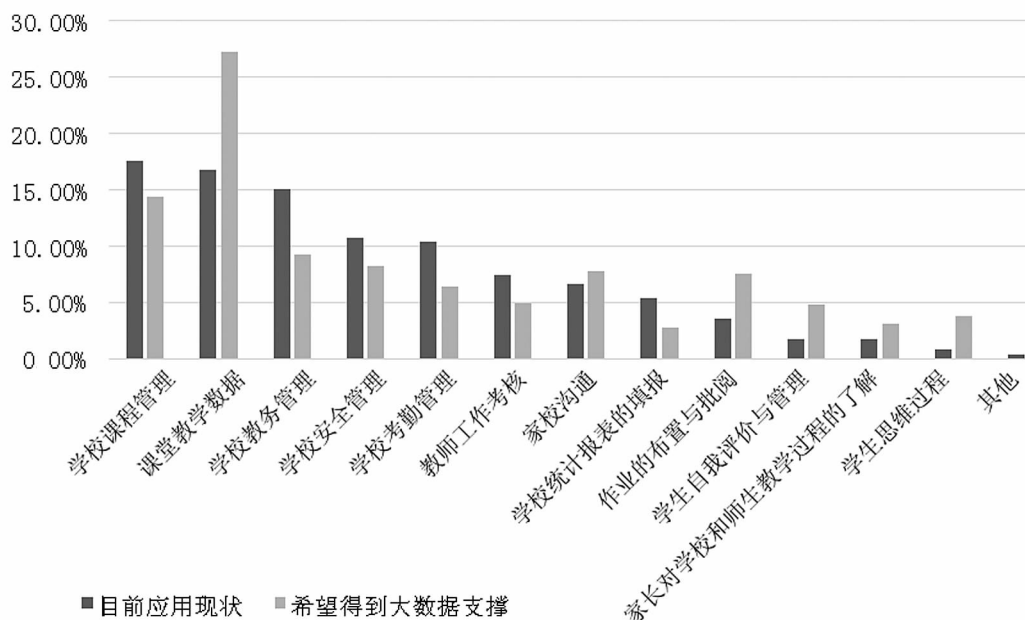


图 2 教师对学校教育大数据的供需对比

表三 东、中、西部教师对教育大数据的认知

| 选项      | 东中西部编码(%)          |                    |                    | 汇总<br>(N=5434) |
|---------|--------------------|--------------------|--------------------|----------------|
|         | 1.0(东)<br>(N=1999) | 2.0(中)<br>(N=1250) | 3.0(西)<br>(N=2185) |                |
| 数据数量大   | 1393(69.7)         | 876(70.1)          | 1403(64.2)         | 3672(67.6)     |
| 数据类型多   | 1305(65.3)         | 832(66.6)          | 1405(64.3)         | 3542(65.2)     |
| 数据处理速度快 | 1290(64.5)         | 855(68.4)          | 1350(61.8)         | 3495(64.3)     |
| 数据价值大   | 954(47.7)          | 595(47.6)          | 981(44.9)          | 2530(46.6)     |
| 数据来源多   | 831(41.6)          | 544(43.5)          | 1042(47.7)         | 2417(44.5)     |
| 数据存储量大  | 446(22.3)          | 307(24.6)          | 653(29.9)          | 1406(25.9)     |
| 数据价值小   | 59(3.0)            | 41(3.3)            | 103(4.7)           | 203(3.7)       |
| 其他      | 7(0.4)             | 5(0.4)             | 8(0.4)             | 20(0.4)        |

#### 4. 教师应用教育大数据的动力与评价

首先,对于使用学生数据最大的支持和动力来源的调查表明,选择“学校教学工作必须使用”的教师占比最高(23.2%),其次分别是“大数据能帮助改进教学”(占20.9%),“同事都在用,以便于交流”(12.7%),“想尝试一下”(11.9%),“专业发展规划”(11.7%),“校长要求”(8.0%),“大数据比较热”(6.7%),“学校有相关的培训”(5.4%)。由此可见,教师对教育大数据的态度受外部因素和内部因素的共同影响,既有学校要求,也有自我驱动。

其次,教师日常教育教学中易于接触数据是其应用教育大数据的基本前提。调查发现,响应率排在前五位的便利可接触的数据由高到低依次是学生基本信息(16.3%)、学校课程安排(15.9%)、学校教务信息(13.8%)、学校人员信息(10.9%)和学校统计报表(10.9%)。响应率约占5%的分别是学生作业、考勤、考试信息和学习过程记录。低于5%的为课堂学习行为、教师考核信息、家长信息。这也反映了数据应用水平和开放程度,教师便利可接触的数据类型和内容有限,会影响教师应用教育大数据的积极性与实效性。

再次,本研究调查了教育大数据推广中存在的问题。统计发现,增加教师工作量(18.9%)、增加学校管理工作量(13.50%)、数据应用有限(12.40%)、学校硬件不投入(10.90%)、增加学校财政支出(10.20%)、学校软件建设不投入(7.60%)影响较大;小于5%的分别是不会数据统计与分析、关注的教学数据

无法收集、已有数据的使用不符合教学需求、不适应新的数据平台或系统、无法评价和比较数据、校长及所在区(县)教育局不重视。这些调查结果再次反映了教育大数据在落地推行过程中,面临着外部环境、制度和主体因素的多重挑战,导致供需不一致,对教育大数据应用产生了负面影响。

#### (二) 不同教师群体应用教育大数据积极性的方差分析

为了探究不同教师群体应用教育大数据的积极性有无差异,本研究通过方差分析,对不同性别、民族、教龄、职称(初级、中级、高级)<sup>④</sup>、毕业专业类型(是否师范类专业)、主要上网途径、家校距离、任教学校类型(所属行政区划级别及城乡类型)、所在区域的教师群体应用教育大数据的积极性进行显著性检验,研究发现,不同性别、民族、职称、毕业专业类型和主要上网途径的教师群体应用教育大数据的积极性没有显著差异。然而,不同教龄、不同家校距离、不同任教学校类型、不同区域的教师群体应用教育大数据的积极性有显著差异。6-10年教龄的教师应用教育大数据的积极性最高,16-20年教龄的教师应用教育大数据的积极性最低。家校距离5-10公里之间的教师应用教育大数据的积极性最高,家校距离5公里以上的教师应用教育大数据的积极性明显高于家校距离5公里以内的教师。东部教师应用教育大数据的积极性高于中部教师,中部教师又高于西部教师。不同行政区划层级和城乡教师应用教育大数据的积极性也呈现0.01水平显著性差异( $F=5.62, P=0.00$ )。事后检验分析发现,地级市及以上的城区学校教师应用教育大数据的积极性高于县城或县级市的城区学校、乡镇学校和村完小教师;县城或县级市的城区学校教师应用教育大数据的积极性高于乡镇学校教师(见表五)。

#### (三) 教师应用教育大数据积极性的影响因素回归分析

本研究将评价方式、教师态度、学科适用、学校投入、传播效果与教师应用教育大数据的积极性进行相关分析,结果发现,五个自变量之间、五个自变量与因变量之间存在显著相关关系,这说明五个自变量和因变量之间可能存在预测关系,需通过回归分析验证假设,分析变量之间的因果关系。

表四 不同教师群体对教育大数据积极性的单因素方差分析

|        |              | 平均值 ± 标准差   | F 值    | P 值      |
|--------|--------------|-------------|--------|----------|
| 性别     | 男            | 3.25 ± 0.74 | 3.112  | 0.078    |
|        | 女            | 3.29 ± 0.69 |        |          |
| 民族     | 汉族           | 3.28 ± 0.71 | 3.67   | 0.055    |
|        | 少数民族         | 3.23 ± 0.67 |        |          |
| 教龄     | 5 年以下        | 3.28 ± 0.70 | 2.893  | 0.021 *  |
|        | 6 - 10 年     | 3.33 ± 0.70 |        |          |
|        | 11 - 15 年    | 3.29 ± 0.71 |        |          |
|        | 16 - 20 年    | 3.23 ± 0.69 |        |          |
|        | 20 年以上       | 3.25 ± 0.71 |        |          |
| 职称     | 初级           | 3.29 ± 0.70 | 1.639  | 0.194    |
|        | 中级           | 3.26 ± 0.69 |        |          |
|        | 高级           | 3.24 ± 0.71 |        |          |
| 毕业专业类型 | 师范           | 3.28 ± 0.70 | 3.398  | 0.065    |
|        | 非师范          | 3.24 ± 0.71 |        |          |
| 主要上网途径 | 手机网络         | 3.26 ± 0.75 | 1.642  | 0.161    |
|        | 电脑网络         | 3.26 ± 0.68 |        |          |
|        | 通过 Wifi 连接   | 3.31 ± 0.69 |        |          |
|        | 用手机和电脑网络差不多  | 3.29 ± 0.62 |        |          |
| 家校距离   | 住校内          | 3.21 ± 0.69 | 2.869  | 0.014 *  |
|        | 住在学校所在的村里    | 3.25 ± 0.64 |        |          |
|        | 1 - 5 公里之间   | 3.28 ± 0.71 |        |          |
|        | 6 - 10 公里之间  | 3.32 ± 0.72 |        |          |
|        | 11 - 20 公里之间 | 3.30 ± 0.68 |        |          |
| 任教学校类型 | 20 公里以外      | 3.31 ± 0.73 | 5.618  | 0.000 ** |
|        | 地市级以上的城区学校   | 3.32 ± 0.72 |        |          |
|        | 地市级以上的农村学校   | 3.25 ± 0.80 |        |          |
|        | 县城或县级市的城区学校  | 3.26 ± 0.66 |        |          |
|        | 县城或县级市的农村学校  | 3.29 ± 0.69 |        |          |
|        | 乡镇学校         | 3.18 ± 0.67 |        |          |
|        | 村完小          | 3.20 ± 0.65 |        |          |
|        | 教学点          | 3.20 ± 0.66 |        |          |
| 区域     | 东部           | 3.32 ± 0.72 | 12.139 | 0.000 ** |
|        | 中部           | 3.31 ± 0.71 |        |          |
|        | 西部           | 3.22 ± 0.67 |        |          |

注: \* 代表  $p < 0.05$ , \*\* 代表  $p < 0.01$ 。

本研究以教师应用教育大数据的积极性为因变量,以评价方式、教师态度、学科适用、学校投入和传播效果等影响因素为自变量,进行多元回归分析后发现,五个自变量均能正向预测教师应用教育大数据的积极性,教师应用教育大数据的积极性与影响

因素的判定系数  $R^2$  为 0.357,调整后  $R^2$  为 0.357,说明自变量可以解释因变量 35.7% 的变化原因。模型公式为:认知感受 = 1.008 + 0.155 \* 评价方式 + 0.182 \* 开放态度 + 0.104 \* 传播效果 + 0.096 \* 学校投入 + 0.115 \* 学科适用。针对模型的多重共线性检验结果表明,模型 VIF 值均小于 5,意味着不存在共线性问题;D-W 值在 2 附近,说明模型不存在自相关性,样本数据之间没有关联,模型较好。模型通过 F 检验 ( $F = 603.659, P < 0.05$ ),也说明评价方式、开放态度、传播效果、学校投入、学科适用不同程度地对认知感受产生影响(见表七)。多元回归分析结果说明,教育大数据应用成效的评价方式越可操作,教师的积极性越高;教育大数据相关的理念、知识、方法传播的渠道越广泛,教师的积极性越高;学校对硬件和软件投入越充分,教师的积极性越高;不同学科的应用差异化与适用性越强,教师的积极性越高;教师的态度越开放,积极性越高。

表五 任教学校类型与教师应用教育大数据的积极性事后检验分析

|      | (I) 名称 | (J) 名称 | (I) 平均值 | (J) 平均值 | 差值 (I-J) | LSD 统计量 | P 值      |
|------|--------|--------|---------|---------|----------|---------|----------|
| 认知感受 | 1      | 2      | 3.321   | 3.247   | 0.073    | 1.427   | 0.154    |
|      | 1      | 3      | 3.321   | 3.265   | 0.056    | 2.053   | 0.040 *  |
|      | 1      | 4      | 3.321   | 3.288   | 0.033    | 0.623   | 0.534    |
|      | 1      | 5      | 3.321   | 3.18    | 0.14     | 5.274   | 0.000 ** |
|      | 1      | 6      | 3.321   | 3.204   | 0.116    | 2.901   | 0.004 ** |
|      | 1      | 7      | 3.321   | 3.203   | 0.118    | 1.383   | 0.167    |
|      | 2      | 3      | 3.247   | 3.265   | -0.017   | -0.309  | 0.757    |
|      | 2      | 4      | 3.247   | 3.288   | -0.04    | -0.564  | 0.573    |
|      | 2      | 5      | 3.247   | 3.18    | 0.067    | 1.219   | 0.223    |
|      | 2      | 6      | 3.247   | 3.204   | 0.043    | 0.688   | 0.492    |
|      | 2      | 7      | 3.247   | 3.203   | 0.045    | 0.455   | 0.649    |
|      | 3      | 4      | 3.265   | 3.288   | -0.023   | -0.409  | 0.682    |
|      | 3      | 5      | 3.265   | 3.18    | 0.084    | 2.509   | 0.012 *  |
|      | 3      | 6      | 3.265   | 3.204   | 0.06     | 1.336   | 0.182    |
|      | 3      | 7      | 3.265   | 3.203   | 0.062    | 0.703   | 0.482    |
|      | 4      | 5      | 3.288   | 3.18    | 0.107    | 1.9     | 0.057    |
|      | 4      | 6      | 3.288   | 3.204   | 0.083    | 1.304   | 0.192    |
|      | 4      | 7      | 3.288   | 3.203   | 0.085    | 0.86    | 0.39     |
|      | 5      | 6      | 3.18    | 3.204   | -0.024   | -0.537  | 0.591    |
|      | 5      | 7      | 3.18    | 3.203   | -0.022   | -0.256  | 0.798    |
|      | 6      | 7      | 3.204   | 3.203   | 0.002    | 0.016   | 0.987    |

注: \* 代表  $p < 0.05$ , \*\* 代表  $p < 0.01$ 。

表六 相关性分析结果

|      |             | 评价方式     | 开放态度     | 传播效果     | 学校投入     | 学科适用          | 积极性 |
|------|-------------|----------|----------|----------|----------|---------------|-----|
| 评价方式 | Pearson 相关性 | 1        |          |          |          |               |     |
| 开放态度 | Pearson 相关性 | 0.139 ** | 1        |          |          |               |     |
| 传播效果 | Pearson 相关性 | 0.144 ** | 0.436 ** | 1        |          |               |     |
| 学校投入 | Pearson 相关性 | 0.170 ** | 0.373 ** | 0.395 ** | 1        |               |     |
| 学科适用 | Pearson 相关性 | 0.408 ** | 0.116 ** | 0.127 ** | 0.178 ** | 1             |     |
| 积极性  | Pearson 相关性 | 0.379 ** | 0.409 ** | 0.374 ** | 0.366 ** | 0.350 ** ** * | 1   |

注: \* 代表  $p < 0.05$ , \*\* 代表  $p < 0.01$ 。

表七 教师应用教育大数据的积极性多元回归分析

|                | 非标准化系数 |       | 标准化系数 | t      | p        | VIF   | R <sup>2</sup> | 调整 R <sup>2</sup> | F                 |
|----------------|--------|-------|-------|--------|----------|-------|----------------|-------------------|-------------------|
|                | B      | 标准误   | Beta  |        |          |       |                |                   |                   |
| 常数             | 1.008  | 0.043 | —     | 23.655 | 0.000 ** | —     | 0.357          | 0.357             | 603.659(0.000 **) |
| 评价方式           | 0.155  | 0.008 | 0.223 | 18.528 | 0.000 ** | 1.223 |                |                   |                   |
| 开放态度           | 0.182  | 0.01  | 0.232 | 18.606 | 0.000 ** | 1.317 |                |                   |                   |
| 传播效果           | 0.104  | 0.008 | 0.159 | 12.615 | 0.000 ** | 1.344 |                |                   |                   |
| 学校投入           | 0.096  | 0.008 | 0.146 | 11.833 | 0.000 ** | 1.284 |                |                   |                   |
| 学科适用           | 0.115  | 0.007 | 0.186 | 15.462 | 0.000 ** | 1.219 |                |                   |                   |
| 因变量: 积极性       |        |       |       |        |          |       |                |                   |                   |
| D - W 值: 1.995 |        |       |       |        |          |       |                |                   |                   |

## 四、讨论与建议

### (一) 相关讨论

被调查教师群体对大数据特征的认知没有显著区域差异,东、中、西部的教师认知水平相当,对“数量大、类型多、速度快”达成了普遍共识,但对大数据价值层面的认知还存在一定局限。究其原因,主要在于教育大数据的应用尚不普遍,导致理念先导和实践滞后间的割裂。实践应用与政策倡导和学者研究仍有较大差距。调查发现,教学工作(备课、上课和批改作业)和非教学工作(处理学校、班级及其他工作)占据了教师大部分工作时间,少有专门时间用来接触或处理数据,这说明教育大数据的应用尚不普及,也不深入,并没有融入教师日常教学与管理工作,仍是附加在教育之外的“任务”。这与学校的信息化环境、制度管理和教师数据素养密切相关。学校的硬件设备、网络环境均已基本配置达标,但在应用数据平台或系统时,仍有一半教师主要使用

“查询”“比较”等初级功能,来自学校管理要求的外部动机大于工作需要或专业发展的内部动机,与同事协作或完成更高行政级别的任务时更倾向于使用或研究数据,同时他们也认为教育大数据在推广过程中增加了工作量。这反映了大数据在学校层面的应用实践与政策、理论和理念之间存在一定差距,政策制度、理念宣传与理论研究已基本实现了相对公平,没有明显的区域差异,但在实践层面还未能实现预期的“减负提质”。

究其原因,本文认为现有大数据的应用未能真正扎根于教育实践,造成供需不匹配。学校关注管理和教务工作,教师更看重课堂教学和有效的沟通评价,不同群体对数据需求不一致,而真正与课堂教学、学习本质融合的大数据在技术和教育实践中都更具挑战性,很难立竿见影而受到市场冷落;学校信息化建设评估与考核的主要指标仍然是以硬件设备配置与课堂使用率为主,同时也造成配套软件、培训和教学实践的应用指导相对匮乏;学校对教师应用



数据的要求和权限多集中于学生基本信息、课程安排、考勤安全、考试成绩、相关反馈等报表信息。国内外研究均已发现,教师对大数据在教育领域中的应用充满好奇,期望大数据的应用能够全面提高教育质量、期待得到有针对性指导和培训(Joshua, 2017; Means et al., 2008; 高巍等, 2018; 郑燕林等, 2015; 杨现民等, 2018),但在现实中受到环境和制度的影响,“所用非所需”一定程度上限制了教师对教育大数据的认知程度和应用的积极性、易用性和有效性。

具体而言,教师应用教育大数据的积极性,在性别、民族、职称、毕业专业类型和上网途径方面没有显著差异,作为一个相对新的理念与技术,它的受众非常广泛,这也再次证明大数据时代信息传播的广泛性与全纳性。但不同教龄、家校距离、区域(东、中、西部)、城乡的教师应用教育大数据的积极性存在显著差异,教龄与教师的职业倦怠和职后专业发展动力有关,家校距离决定了人与大数据连结的可能性,区域和城乡差异涉及大数据能否促进教育公平。美国《教师使用学生数据系统改进教学》报告还专门对贫困程度不同地区的学校教师应用大数据情况进行了比较分析,发现高贫困学校的教师使用学生数据系统的频率至少与低贫困学校的教师相同,二者之间的差异在 11 个用途的调查中只有 3 个具有统计学意义,即高贫困学校教师比低贫困学校教师更倾向使用学生数据系统跟踪学年考试成绩,确定学生的考试需求。此外,学校的贫困水平往往与数据系统的使用频率和数据收集内容有一定的关联(Means et al., 2008)。这也为我国推进大数据在教育领域的应用和教育精准扶贫提供了新思路,优先发展贫困地区、农村地区,可能更富实效性地促进教育信息化对于促进教育公平和提升教育质量的积极作用。

因此,可以初步判断,大数据的价值实现更需要差异化和多元化,需要在教师群体特征上关注教龄和区域两个重要维度;在途径与内容方面,更多关注教育大数据应用的评价方式的可操作性、教师态度的开放性、学科教学的适用性、学校投入的有效性与持续性、传播效果的触达性和广泛性,才能切实提高教师应用教育大数据的积极性,进而不断推进教育与大数据的融合与创新。

## (二) 建议

1. 加强各级政策体系的配套与联动,并保证落实到教育的最基层

国家政策的制定通过顶层设计和宏观指导,为教育发展与改革指引方向,明确政策执行中的原则和框架。地方政策制度应在遵循国家政策的原则和边界内,进行本地化和可操作化的政策解读与政策研制,省级政策和县(市)级政策是确保国家政策落实的支撑点,也是国家政策灵活性和创造性的体现。政策体系应呈叶脉状拓展,不能仅通过“转发”的形式用政策落实政策。对于大数据在教育领域的应用,地方政府应落实并赋权执行主体于教育部门和工信部门,同时协调其他相关部门通力协作,规定具体、详实的与政策文件配套的人、物、权,以分步骤、可操作为宜,忌大目标、难实践、难评价。各级政府联动、政策执行系统连贯,随着政策的下沉,落实也随之加强,确保有效落实到一线师生的教学层面。例如加强区域教育大数据的标准建设和数据积累,学校数据库建设的具体要求、硬件配套和数据使用的报告公布等。

2. 加强以教育为核心的理论与应用研究,促进教育与大数据的深度融合

大数据与教育领域的深度融合,核心体现在课堂教学的融合与转变。目前的研究多集中于政策研究、理论引介和模型建构,但同时更急需关注的是服务于教学实践的实验性、实证性研究。大数据与教育的深度融合,应如金融、交通、医疗等行业的成熟可推广式应用,不仅不会增加教师负担,反而可以促进教育公平的实现,提高教学质量和工作效率。

加强跨界研究合作,主要指学科领域、行业领域的“教育+X”合作模式,不同教育阶段的贯通协作,如教育+教育技术/技术,教育+金融/医疗/交通等,学校+行政部门+社会力量(PPP模式),中小学教师与大学学者的共同体研究等,既可以深入了解不同行业在应用大数据过程中的经验与方法,又可以广泛收集数据、研究数据分析与使用的科学方法。其中,最重要的是转变意识,将教育理念融于技术应用,而不是技术附加在教育上或教育被技术牵着走。学校应从被动接受者转变为主导者,从数据提供者和消费者转变为数据的所有者、数据分析与使用的引领者。很多公司提供的数据服务和产品,

不仅无法满足学校教育的管理或不适用于学校教育的实际,反而搜集了学校管理、师生教学互动行为的大数据为己所用,这样不但增加了师生和家长的负担,还增加了学校的管理成本、财政支出(王学男,2018)。除了宏大、抽象、艰深的理论建构外,跨界研究应更聚焦于大数据在教育领域的深化应用,如评价方式、教师态度及信息素养、学科适用、学校投入、传播效果等,逐一解决具体问题,促进其良性可持续发展。

3. 广泛宣传与定向培训相结合,帮助教师成为数据红利的践行者

大众传播对于宣传普及知识有其特殊的优势,如教师对大数据特征的认识没有明显的区域差异;教师对教育大数据的积极性没有显著的人口学特征差异(性别、民族),在职称、是否是师范类专业毕业、上网途径方面也没有体现出显著差异,较好地实现了信息传播的公平。但是在教龄、家校距离、城乡、区域(东、中、西部)变量上存在显著差异,也就为分类施策指出了方向。因此,定向培训成为按需设计、分类管理的实践载体。加强教师的信息化理念和技能培训是信息时代快速便捷、高效优质地提升教师信息素养、数据素养的必要途径,在培训的整体设计上要体现差异化的分类思想,培训者要针对不同省份的信息化程度、不同参训人员所在的学校类型、不同教龄教师的专业情意和态度设计相适的内容;在培训内容设计上,不再仅是 Word、PPT、Excel 和电子白板的基本操作,更需有意识地增加大数据的基本概念、学科教学与大数据融合应用、数据收集与数据初级分析内容等;在培训方法上,注重实践操作和分组协作,不仅是传统的听课式培训;在时间安排上,增加有效学时,包括理论学习与实践学习,其中实践学习又包括实际操作、实地考察和跟岗实习等。

4. 加强学校制度建设与软件配套,切实支持和鼓励教师实践探新

教育大数据并非孤立的,而是与教育、教学、管理融为一体,学校管理者在学校范围内,应兼顾多方利益群体的共同需求及管理与教学的需求,从而实现人的发展的最终目标。第一,确立自身的专业角色,作为教育实践和教育管理者,学校和教师应有专业自信和专业视角,能与技术专家对话,并指导公司

进行教育大数据的数据收集与分析以及开发、设计、探索等,如教材中的知识点分析、教学环境对学生专注度的影响、学生的移动轨迹与学校布局、学生搜索的频次与时间反映出的学习困难点和学习活跃时间、学生购书和阅读的现状与语文成绩的关联等。第二,着力提升学校全体教师的素养和骨干教师的数据分析能力,为学校教育大数据挖掘与使用,建构结构合理的教师团队和“帮带”机制。第三,在工作时间、工作内容和形式上,给予实质性的支持和鼓励,每周指定专门的工作时间用于教师处理与教育大数据有关的工作,适当开放数据使用权限,结合大数据的特点,安排团队协作式工作,让教师有更多机会接触和使用数据,并通过分享、帮带的形式激励教师。

#### [注释]

①重点委托是指委托这一地区教育行政相关人员,根据具体要求,向本地区发放问卷,填写问卷的教师人数不低于 200 人。

②除宁夏回族自治区以外,其他省份均有被试填写调查问卷。

③多重响应应用于多选题分析,分析多选题选项的选择比例情况等,主要涉及到两个概念:响应率和普及率。响应率用于对比各个选项的相对选择比例情况;普及率用于某项的选择普及情况,即整体上看多选题各选项占所有选择的比例情况。

④初级职称包括:未定级、小学三级、小学二级、小学一级、中学三级、中学二级;中级职称包括小教高级、中教一级;高级职称包括中教高级、特级教师。

#### [参考文献]

- [1]高巍,刘瑞(2018). 基于教育大数据的学习行为分析研究:教师评学与导学的角度[J]. 教师教育论坛, (7):38-42.
- [2]Hattie, H. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement [M]. Routledge, 237-238, 238-239, 258.
- [3]Iyer, K. V. (2018). Big data analytics: The stakes for students, scientists & managers-management perspective [A]. Accepted for oral presentation at EeL 2018 conference, Singapore, (9):1-9.
- [4]Joshua, E. (2017). Teaching future big data analysts: Curriculum and experience report [R]. Parallel & Distributed Processing, (7): 346-351.
- [5]李芒,孔维宏,李子运(2017). 问“乔布斯之问”:以什么衡量教育信息化作用[J]. 现代远程教育研究, (3):3-10.
- [6]刘雅馨,杨现民,李新,田雪松(2018). 大数据时代教师数据素养模型构建[J]. 电化教育研究, (2):109-116.
- [7]李青,赵欢欢(2018). 教师数据素养评价指标体系研究[J]. 电化教育研究, (10):104-110.
- [8]Means, B., Gallagher, L., & Padilla, C. (2008). Teachers' use

of student data systems to improve instruction [R]. U. S. Department of Education; 35.

[9] Marie, B., Mingyu, F., & Barbara, M. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief [R]. U. S. Department of Education, (10): 9-51.

[10] 王正青, 徐辉 (2018a). 大数据时代美国的教育大数据战略与实施 [J]. 教育研究, (2): 120-126.

[11] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶 (2015). 与大数据同行——学习和教育的未来 [M]. 赵中建, 张燕南译. 上海: 华东师范大学出版社.

[12] 王正青, 张力文 (2018b). 大数据时代美国发展教师数据素养的基础与路径 [J]. 比较教育研究, (2): 68-75.

[13] 王学男 (2018). 从大数据中提升学校教育的获得感 [J]. 教学与管理, (12): 31-34.

[14] Yu, X., & Wu, S. (2016). Typical applications of big data in

education [J]. Educational Innovation Through Technology, (4): 103-106.

[15] 杨现民, 王榴卉, 唐斯斯 (2015). 教育大数据的应用模式与政策建议 [J]. 电化教育研究, (9): 54-61.

[16] 杨现民, 田雪松 (2016). 互联网+教育蓝皮书: 中国基础教育大数据 [M]. 北京: 电子工业出版社: 221-256.

[17] 杨现民, 田雪松 (2018). 中国基础教育大数据 2016-2017: 走向数据驱动的精准确教学 [M]. 北京: 科学出版社: 216-221.

[18] 祝智庭, 孙妍妍, 彭红超 (2017). 解读教育大数据的文化意蕴 [J]. 电化教育研究, (1): 28-36.

[19] 郑燕林, 李卢一 (2015). 对大数据支持的学习分析与评价的需求调查: 基于教师的视角 [J]. 现代远距离教育, (2): 36-42.

(编辑: 李学书)

## Teachers' Cognition on Big Data in Education and Its Influential Factors: A Survey of 5434 Teachers in China

WANG Xuenan

(Institute of Education Development and Reform, National Institute of Education Sciences, Beijing 100088, China)

**Abstract:** *The big data in education is a necessary condition for the sustainable development of educational informatization 2.0. As the intersection subject of big data and education, teaching and management, teachers play an important role. In this study, descriptive analysis, variance analysis, Pearson correlation analysis, and multiple regression analysis were carried out by SPSS to investigate the cognition and influencing factors of 5434 teachers on educational development data. The survey results show that the teachers' understanding of big data characteristics is mainly focused on "large quantity, large type and fast speed"; There is no significant regional difference in the east, middle and west regions; There are significant differences in teachers' enthusiasm for big data of education among different teaching age, distance between home and school, region (east, middle and west) and type of teaching school. The limitations of external conditions such as teachers' working hours, access to data, teaching needs, school software support have a negative impact on teachers' application of big data in education. Evaluation methods, teachers' attitudes, subject application, school input, and dissemination effects can positively predict teachers' activity for big data in education. Finally, it puts forward countermeasures and suggestions from four aspects: policy implementation, theoretical research, training design, and school implementation.*

**Key words:** *teachers; big data in education; cognition; application status; influential factors*