

中国教育财政

怀仁怀朴 唯真唯实

北京大学中国教育财政科学研究所

2015年第2期（总第101期）

2015年2月3日

以增值评价助推评价改革和教育公平

王允庆*

“增值评价”是发展性评价的一种方式，与惯常以升学率高低和平均分排序评价学校不同，“增值评价”是以评价学生在一段时间教育过程后的“成长”和“变化”，取代原来对学生在某一个特定时刻的结果的关注。形象点说，增值评价以一段时间内“谁长得快”评价学校。

教育部2013年6月出台的《关于推进中小学教育质量综合评价改革的意见》（教基二〔2013〕2号）（简称《意见》）提出改进评价方式：“……注重考查学生进步的程度和学校的努力程度，改变单纯强调结果不关注发展变化的做法。”这一改进应以增值评价工具予以落实。

“增值评价”以评价学生在原有基础上进步的程度来评价学校教育教学工作的效果，这为评价学校教育质量提供了一种公平的视角。这种评价必须面向所有的学生，关注学生的进步与变化，以此反映学校教育努力的真实情况。在各级政府为推进教育公平和义务教育均衡配置资源而付出真实努力的现阶段，引入增值

* 作者简介：王允庆，国家督学，大连现代学习科学研究院院长，大连市教育局原党委书记。

评价意义深远而重大。它不仅可衡量出教育公平是否真正落实在每一所学校和每一个学生身上，还可以作为“杠杆”推动评价改革，促进教育质量层面的“公平”得以落实，使政府真实了解每个区域每所学校的教育效能与质量状况。

增值评价自 20 世纪 80 年代开展以来，已经在越来越多的国家和地区得到了实施，形成了以推行增值评价助推教育公平落实、教育质量提高的教育共识。

一、增值评价的实施

增值评价的实施包括基线测试（起点测试）、问卷调查和增值分析三个主要组成部分。

基线测试和问卷调查在初、高中入学后的第一学期进行，是整个学段学习的起点测试。增值分析在学生 2-3 年学习期间所在地区进行大规模测试（当地自己进行的测试，如学业水平测试、地方统一测试等，增值分析不要求进行独立的专门的测试）后利用测试的数据进行。

1. 基线测试

基线测试属认知评价，测量的是当前学生基础学习能力，测试的重点不是了解学生学了多少知识，而是了解影响学生学习未来的基础学习能力。当然能力不能完全离开知识而独立表现出来，所以测试还是依据若干学科进行，初中生进行语文、数学、英语三个学科基本能力的测试，再加一个认知能力测试；高中生则还要加上科学学科测试。

2. 问卷调查

问卷调查属非认知评价，有学生问卷、教师问卷和校长问卷三项调查。

学生问卷主要针对学生的学习背景、儿童人格，包括学习和教学的相关问题，例如：自然状况、学科偏好、家庭状况、对未来的预期、活动与社交、人际关系、师生关系、同伴关系等基本状况，还包括学习适应、学习方法、学习风格、学习动机、自我概念、自我效能感等学生基本特征方面。

教师问卷主要针对教师的教学背景，包括自然状况、专业状况、专业思想、教学状况、阅读状况、发展期望、学校教学环境等教师特征和学校特征方面。

校长问卷主要针对校长对学校的管理方式、管理行为、制度建设、教学与课

程领导、学校自主权等学校特征和校长特征方面。

3. 增值分析

进行基线测试和问卷调查后，得到了参与的学生和学校的基本测试和调查数据，经过数据分析就可以形成“学生起点成绩报告”、“学生学习背景调查报告”，对每个学校每个地区的学生成绩和学生学习背景进行详尽的分析。通过数据挖掘，还能得出三项重要的结果：一是对学生学习成绩增值的预测，就是对学生学习可能、学习潜力的预判；二是各区域之间、学校之间学生学习能力的均衡分析，将学生成绩的差异分解为学校之间的差异和学生个人的成绩差异，并可以对不同地区的两个差异的占比进行比较；三是把测试成绩和学习背景，包括教师问卷和校长问卷提供的学校特征、校长特征和教师特征都联系起来作相关分析，以此来探索影响学生学习成绩的因素和各个因素对学生成绩的影响程度。这后两项结合起来就叫作“学生成绩差异分析”。

对学生成绩的预测有助于学校基于学生的学习可能和学习潜力作进一步改善教学的参考，为学生制定恰当的目标。

差异分析中学生学习能力的均衡性分析对于地方教育行政部门和学校非常重要，深刻、全面地了解本地区学校之间的学生学习能力差异对于促进学校的质量均衡发展和推进教育公平有非常重要的现实意义。差异分析的另一影响——影响学生学习成绩的因素和各个因素对学生成绩的影响程度也非常重要。它引导教育领导者和教师不仅要关注成绩，更要关注与成绩相关联的“要素”。这是对学校全面关注学生发展的有力引导，也相当于按《意见》指出的，对“学生的品德发展水平、学业发展水平、身心发展水平、兴趣特长养成、学业负担状况”的相互联系进行了评价，对“学校教师队伍、设施设备、教育教学管理等影响教育质量相关因素的数据资料”进行了分析，也是对教育质量的实证性归因分析。以增值评价落实《意见》中的这些举措，必将有力推进教育评价的改革，再通过评价不断改善地区、学校影响学生学习水平发展的因素，促进学生学习水平的不断提高，进而全面提高质量。

因参与增值评价的是初一和高一的学生，都是刚进入到该学段的学生，因此学生成绩的差异分析就表现出两个特点：一是表现出他们所在的上一学段的学校的教育特征。学习者特征的形成不是一日之功，一般而言，应该是前一学段的教

育结果，或者与前一学段的教育结果密切相关，因此学生学习能力差异在一定程度上反映了地区之间、学校之间的教育教学差异，因而指明了地区和学校的努力方向，这对地区和学校教育教学的均衡发展起着重要的推进作用。二是学生的成绩差异深刻反映着学生的学习特征，可以对此进行学生学习的综合特征地区间的比较研究，使各地区了解本地区学生学习的长处和不足，从而发扬长处，弥补不足，这必将促进更大范围内学生学习质量的均衡提高。这正是教学评价的根本目的所在。

增值评价中无论是基线测试还是增值分析，其具体的评价过程都是运用开发的专业化软件进行数据分析的过程。我们已经处在一个大数据时代，基于数据改变教育，让数据说话的教育需求日益增加。对此《意见》中指出：“注重全面客观地收集信息，根据数据和事实进行分析判断，改变过去主要依靠经验和观察进行评价的做法”。这就要求我们将学生发展的所有信息收集、整理、分析并得出结论性的认识。在信息技术日益深入学校教育应用的今天，这是可能做到的，但还没有现成的工具，所以要主动去设计、创造和实践。此外，由于涉及每一位学生的数据是大量的、充分的、多元复杂的，也只有运用现代信息技术进行数据分析才能满足关注每一名学生的需要。与此同时，要注意自觉地将数据分析的结果融入各个教学环节中，帮助教育行政管理机构和学校建立基于数据改善教育的机制。这一方式可以被称作：数据驱动改善教学。增值评价是数据驱动改善教学的具体实施。增值评价的结果（各种报告）都是通过数据可视化的方式具体展现出来的。例如：关于学生学习特征、教师特征、学校特征以及学生学习水平均衡性、差异性都是通过数据分析的图表展现出来的，在大数据时代，读懂数据图表、理解数据的统计意义业已成为教师新的专业能力之一。通过参与增值评价并体会大数据方法，弄懂数据分析图表的内涵和意义，则成为学校参与增值评价的一个额外收获。

应该说《意见》和落实《意见》的努力对于开发信息技术系统、运用大数据的方式对学生发展作出全面客观的评价具有重大的指导意义。在全面了解学生学习和发展数据基础上进行的增值评价客观上大大助推教育行政部门全面了解学生，基于数据给予学生发展以精确化、针对性的指导。

因此，增值评价也有助于促进教育评价改革的数据化评价方式，运用它可以

全面掌握一个区域内基础教育各学段学校学生学习状况的基础数据,对学生的基本学习能力、学生学习的环境因素、学生学习行为和个性心理因素进行数据分析,对学生学习的发展状况作出评价。值得注意的是,从大数据的运用来看,这里得到的是关于所有学生和所有学校的总体数据,不是抽样数据,以此为基础必将助推“办好每一所学校,教好每一个学生”;必将助推该区域各级教育行政部门和学校对学生的学习和学校教师教育教学付出的努力作出公正客观科学准确的评价,必将有效地推进教育公平的真正落实。同时采用大数据的方法对学生学习发展的相关因素进行分析(寻求相关性是大数据的另一个特点),以期改进有关的因素,能够促进学生更好更快的发展。

二、增值评价的研究与实践

笔者从 2011 年以来,一直带领团队进行增值评价的研究和实践,于 2011-2013 年在全国部分地区为 34.6 万学生实施了“增值评价”,受到当地教育行政部门领导和学校校长的欢迎。一些教育行政部门的领导认为这是一种真正助推教育公平的评价方式,是以评价为导向落实教育公平的宝贵经验。一些学校的校长强烈呼吁,只有采取此种评价方式,才能体现公平性和科学性,才能更有效地促进每一所学校的发展。相对薄弱学校的校长认为增值评价充分肯定了他们学校的教育教学工作,相对优秀学校的校长认为增值评价使他们看到了自己存在的问题,可以有针对性地予以解决,使自己的学生得到更好的发展。

2013 年辽宁省所有 56 个市区的 147,109 名初一新生参加了增值评价,当年完成了起点测试和问卷调查,对学生测试数据采用项目反应理论分析测验的特征,并使用 Conquest 软件进行项目参数估计和学生能力估计。将估计得到的能力值标准化后转换成均数 70、标准差为 15 的量尺分数,为本次测验的最终报告所有分数。由于本次基线测试各科均采用选择题形式,因此在模型选择上,采用 Rasch 模型。Rasch 模型适用于 0/1 计分题型中,即学生答对某个题目记为 1 分,否则记为 0 分。对于单一维度,在 Rasch 模型中,能力为 θ 的被试 n 答对难度为 δ 的题目 i 的概率 P 可用以下公式表示:

$$P_i(\theta) = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)}$$

其中， $P_i(\theta)$ 表示被试 n 答对题目 i 的概率，其取值范围为(0.0-1.0)； θ_n 表示被试 n 的能力水平，其取值范围为 $(-\infty, +\infty)$ ，通常情况下取 $(-3, +3)$ ； δ_i 表示项目 i 的难度，通常情况下其取值范围为 $(-2, +2)$ 。 θ_n 和 δ_i 均需根据作答数据估计得出。

由此得出的结果形成了《辽宁省初中生起点评价报告》，包括起点测试成绩报告和学习背景调查报告。随后又形成了《辽宁省学生起点成绩差异分析报告》。

这里以《差异分析报告》(以下简称《报告》)为例，说明增值评价对于促进教育评价改革助推教育公平的意义。

《报告》主要阐述了以下三方面问题的数据分析结果：一是不同学校间学生能力的差异状况，即全省、各市、各区的校际均衡状况；二是学校层面、学生个体层面各类因素对学生能力的影响；三是对各市、各区的学生发展状况的综合情况分析。

1. 不同学校间学生能力的差异状况

(1) 辽宁省学生各项能力均存在显著的学校间差异。其中，学生认知能力差异的21.99%、学生语文能力差异的21.84%、学生数学能力差异的21.73%、学生英语能力差异的25.84%，是由于学校间差异所造成的。

(2) 在市级层面，通过具体数据列举出认知、语文、数学、英语各科成绩高位均衡(校际成绩差异小于均值而成绩高于均值)、一般均衡、一般不均衡和低位不均衡的市。需要指出的是，有一个被认为是教育强市的市竟然有两科成绩出现低位不均衡的现象。

(3) 在区级层面，通过具体数据列举出认知、语文、数学、英语各科成绩高位均衡(校际成绩差异小于均值而成绩高于均值)、一般均衡、一般不均衡和低位不均衡的区。应该指出的是，这样的区并不完全属于前面指出的各该类型的市。

学校特点对学生各科成绩差异的解释率的分析如表1所示。

表 1 学校特点对学生各科成绩差异的解释率比较

	认知	语文	数学	英语	关注点
学校基本特征	11.21	10.98	12.06	10.66	学校类型、学校学生就读方式与四科成绩显著相关。建校年限对四科成绩的解释率较高,学校类型对语文、数学成绩的解释率较高
校长基本特征	14.51	13.50	14.31	13.53	校长性别、校长学历与学校间四科成绩均有显著的相关关系。其中,校长性别的解释率较高
学校教师特征	8.76	8.11	8.78	8.12	生师比与学校间认知、语文成绩均有显著的相关关系,且对学校间四科成绩差异的解释率均较高
学校课程开设	8.76	8.40	8.27	10.76	3年来开设校本课程门数与学校间四科成绩均有显著的相关关系。教育行政部门对校本课程开发的看法、3年来开设校本课程门数的解释率较高
学校自主性	11.78	11.01	12.01	9.38	人事权与学校间四科成绩均有显著的相关关系,且解释率较高

2. 学校层面、学生个体层面各类因素对学生能力的影响

辽宁省学生各项能力均存在显著的校内学生个体差异。其中,学生认知能力差异的 78.01%、学生语文能力差异的 78.16%、学生数学能力差异的 78.27%、学生英语能力差异的 74.16%,是由于校内学生个体差异所造成的。

学生特点对四科成绩解释率的分析如表 2 所示。

表 2 学生特点对四科成绩解释率比较

	认知	语文	数学	英语	
学生基本特征	4.19	6.84	4.42	7.47	学生性别、是否独生子女、是否住校、流动留守与学生个体间四科成绩均有显著的相关关系。其中,学生性别的解释率较高
父母基本特征	5.74	7.64	6.70	5.97	父母受教育程度、父母职业与学生个体间四科成绩均有显著的相关关系。其中,父亲受教育程度的解释率较高
自我教育期望	10.36	13.01	12.56	11.10	学生自我教育期望与学生个体间四科成绩均有显著的相关关系,且解释率较高
学生学习习惯	5.64	7.33	6.81	5.86	睡眠时间、平时做作业时间、周末做作业时间、每周课外学习与学生个体间四

					科成绩均有显著的相关关系。其中，睡眠时间对学生个体间认知、语文、数学成绩差异的解释率较高
学生上网行为	7.01	8.77	7.50	6.46	每周上网时间、上网花费时间精力最多的项目、对上网的看法、习惯性上网与学生个体间四科成绩均有显著的相关关系。其中，每周上网时间的解释率较高
心理发展状况	11.05	12.44	12.28	12.20	学习适应、学习方法、学习效能感、学习策略、心理健康、师生关系、同伴关系与学生个体间四科成绩均有显著的相关关系。其中，学习适应、学习方法的解释率较高

《报告》对相关的问题作了分学科、分市分区的详尽数据分析。运用这些数据，省级教育行政部门能够更详尽而精确地把握省域内各市各区学生学习水平或者说学校教育教学效能和质量的均衡程度和差异，同时充分了解影响学生学习的各种因素，从而制定更有针对性的指导方案，推进全省教育的均衡发展和全面提高。市、区两级则可以在两个方向上运用这些数据：一个是向上的方向，了解本市、本区的学生能力在全省（或全市）中的状况，能够比较准确地把握本市本区教育教学效能的特点，例如与兄弟市、区比较有哪些优势有哪些困难，从而制定方案发扬优势克服困难，向高位均衡发展；一个是向下的方向，能够更详尽而精确地把握本市各区、本区各校学生学习水平或者说学校教育教学效能的均衡程度和差异，同时充分了解影响学生学习的各种因素，从而制定更有针对性的有区别的指导方案，推进全市、全区教育教学的均衡发展和全面提高。学校和教师也可以在两个方向上运用数据，更重要的是关注影响学生学习的各种因素，由此改善本校的教育教学，促进学生学习水平的均衡发展和全面提升。

3. 对各市、各区的学生发展状况的综合分析

《报告》描述了学生的认知能力、学业成绩（语文、数学、英语）、学习适应、学习方法、学习效能感、学习策略、心理健康、师生关系、同伴关系等 11 个方面的指标，对各市、各区进行排序，形成综合印象，为各市、各区深入了解本市、本区学生的学习状况，制定引领策略，推进学生学习水平均衡发展和全面提升提供了一致性的依据。

《辽宁省初中生起点评价报告》和《辽宁省学生起点成绩差异分析报告》得

到了省市教育行政部门和学校的不同程度的多方面的运用,产生了不少数据挖掘的新经验,运用报告的数据分析改善教育教学工作,推动教育的均衡发展,促进学生学习水平的全面提升已经成为有关部门和单位的共识。

由此可见,采用增值评价能够促进学校关注所有学生的进步幅度,因为学校是由每一位学生组成的,所以力争上游的学校必须关注每一位学生,从而促进每一个孩子的充分发展;采用增值评价能够更公正客观地评价教师的教育教学工作——如何促进了所有孩子的进步提高、充分发展;采用增值评价能促进每一所学校的发展,薄弱学校能看到自己教育教学努力的成效从而能增添信心,优质学校能看到自己存在的问题从而能更好的发展;采用增值评价能促使教育行政部门精确而客观地了解到本区域内学校的教育教学效能的差异,从而制定促使各学校均衡发展、全面提高的指导方案。总之,采用增值评价能够推进教育公平,促进教育均衡发展和全面提升。

三、政策建议

我们建议有关部门制定相应的政策,把增值评价作为落实《意见》的教育评价改革的举措,作为地方常规的教育教学评价方式。以省市教育行政部门为主体,在初高中入学之后的第一学期实施基线测试和问卷调查,产生《起点评价报告》和《差异分析报告》,在此基础上制定相关范围的教育教学指导意见;在后来的一定规模的统考之后实施增值分析,产生《增值分析报告》,对各地区、各学校这一时段(由基线测试到这次统考之间的时段)的教育教学效能进行客观公正的评价,形成新的指导意见。

参考文献:

- [1]谈松华,黄晓婷.我国教育评价现状与改进建议[J].中国教育学刊,2012(1):8-11.
- [2]徐丹,牛月蕾.教育增值评价先行者——美国田纳西州教育增值评价模式解析[J].教育科学,2012(1):83-87.
- [3]张亮.普通高中学生增值评价研究[D].山东师范大学博士研究生学位论文,2010.
- [4]辛涛,张文静,李雪燕.增值性评价的回顾与前瞻[J].中国教育学刊,2009(4):40-43.
- [5]马晓强.增值评价:学校评价的新视角[M].北京师范大学出版社,2012.
- [6]王允庆.增值评价研究——发展性评价的一种实施方式[M].辽宁师范大学出版社,2013.

上期回顾

2015 年第 1 期（总第 100 期）

中国西部农村教师集体性维权行动分析

摘要：本文对 2000-2006 年间农村教师集体性维权行动的变化特征进行了分析，并结合实地调查经验，尝试总结教师群体维稳的若干经验教训。研究发现，2002 年之后比 2002 年之前，教师集体性上访发生的地域范围在增加，规模在提高，请愿对象越来越倾向于更高级别的行政部门。相对于 2002 年以前，集体性罢课事件的频率有所下降，但与较大规模集体性上访相伴的罢课行动依然会发生。农村教师维权的诱因主要围绕教师待遇不合理的问题，也就是说代课教师相对于公办教师以及教师相对于行政部门人员的待遇差距逐渐成为上访的关键诱因，而因拖欠工资引起的上访行为明显减缓。在“以县为主”体制确立之后，农村教师的集体性上访活动之所以会加剧，源于两方面的因素，一是地方行政部门和教育体系内部的利益分配所产生的冲突更为复杂；二是维权教师集体行动的能力逐步增强。研究表明，单纯依靠增加财政投入，并不能够保证教师队伍的稳定。

《中国教育财政》由北京大学中国教育财政科学研究所主办；旨在反映本所最新的学术科研活动；相关内容仅体现作者本人观点，并不必然代表本所的立场。

文章内容仅供参考，如需转载须事先征得本研究所同意。

本期印发：1800 份

下载网址：<http://ciefr.pku.edu.cn>

主办单位：北京大学中国教育财政科学研究所

电子信箱：newspaper@ciefr.pku.edu.cn

责任编辑：毕建宏 **电 话：**010-6275-9700

传 真：010-6275-6183

微信公众号：中国教育财政

地 址：北京市海淀区颐和园路 5 号

北京大学教育学院楼 413 室（100871）

