

基于大数据的学生综合素质评价系统的 开发与应用

——清华大学附属中学的创新实践

王殿军 鞠慧 孟卫东
(清华大学附属中学,北京 100084)

摘要:清华大学附属中学围绕学生综合素质评价的核心问题展开探索,借鉴现代教育理论,借助云计算、大数据等现代化信息技术,率先开发了基于大数据的学生综合素质评价系统。该系统在落实立德树人根本任务、遵循学生身心发展和教育教学规律的基础上,注重以学生行为为依据,关注学生成长的过程性与发展性,注重参与主体、评价体系、结果呈现和发展趋向的多元性,注重数据的动态量化和真实性,帮助学生真实地了解自我,明晰自身的优势特长和发展短板,让每一个学生都找到适合自身成长的方式。经过4年多的实践检验,该系统实现了学生发展和学校教育教学管理、各级学校人才选拔的双向共赢,发挥了综合素质评价育人导向与评价选拔的双重功能,具有过程性、多元性、实测性等创新点。

关键词:高考改革;综合素质评价;立德树人;多元智能理论

【中图分类号】G405

【文献标识码】A

【文章编号】1005-8427(2018)01-0046-8

DOI: 10.19360/j.cnki.11-3303/g4.2018.01.007

习近平总书记指出,时代越是向前,知识和人才的重要性就愈发突出,教育的地位和作用就愈发凸显。我国正处于历史上发展的最好时期,要实现“两个一百年”奋斗目标、实现中华民族伟大复兴的中国梦,必须更加重视教育,努力培养出更多更好能够满足党、国家、人民和时代需要的人才。人才的评价、选拔和培养问题,直接影响到国家教育的健康持续发展。其中,评价的问题更为复杂,也更为关键,因为评价是选拔的基础和条件,评价和选拔也会影响和引导培养的过程和目标。党的十八大以来,落实立德树人根本任务、培育社会主义核心价值观、培养学生发展核心素养的政策精神为学生综合素质评价工作指明了方向。因此,在基础教育阶段,评价和发展学生的综合素质成为亟待解决的关键问题。

1 学生综合素质评价系统的产生背景

经典教育评价是以教育目标为依据,运用有效的评价技术和手段,对教育活动的过程和结果进行测定、分析、比较,并给以价值判断的过程。经典教育评价以结果作为价值判断的依据,例如以考试分数作为依据的评价方式虽然并不全面,但是分数确实能够在一定程度上反映学科学习的程度,也相对公平,因此受到认可。但这种评价方式并不适用于综合素质的评价。

早在2002年,《教育部关于积极推进中小学评价与考试制度改革的通知》中指出,要“综合评价学生的发展,充分发挥评价促进发展的功能”^[1]。2010年,《国家中长期教育改革和发展规划纲要

[作者简介] 王殿军(1960—),男,清华大学附属中学校长,教授;
鞠慧(1981—),女,清华大学附属中学永丰学校行政管理中心主任;
孟卫东(1957—),男,清华大学附属中学副校长,特级教师。

(2010—2020年)》发布,明确提出要全面提高学生综合素质。这一时期,围绕学生综合素质评价取得了一定的成绩,也为未来综合素质评价工作的开展奠定了基础,积累了经验。但囿于种种原因,原有的学生综合素质评价,在过程、方式上存在一定的局限性,限制了对评价结果的有效分析和利用,影响了评价的可信度与参考价值。其主要原因有:

一是在理论层面,我国的教育评价理论研究起步相对比较晚,还缺少真正具有实践指导意义的综合素质评价理论。就目前的成果而言,综合素质评价多指向结果利用,却忽视了评价过程对学生终身发展以及学校办学质量提升的作用,这就使得对日常性、发展性的综合素质评价关注较少。相对而言,现有的理论阐述多,深入实践少,对现状、问题的揭示多,策略的构建多呈点状分布,缺少系统性、整体性的思考。同时,各方对学生综合素质评价内涵认识不一,未取得共识性理解。

二是在实践层面,在全国范围内还缺少能大规模推广的、更具操作性的解决方案。首先,现有评价过程,往往重结果、轻过程,重奖惩、轻激励,很难真正体现学生综合素质评价的过程性、发展性功能。其次,采用等级方式进行评价,会使评价过程中很多丰富、真实、多样的评价信息淹没在数据中,使人们难以真正了解和触摸到学生极富生命力的成长和发展过程。再次,在具体的学生综合素质评价中,“简化程序、集中突击、匆忙收场”的工作方式普遍存在。最后,学生综合素质评价结果未能在高校招生中得到有效使用,这在一定程度上直接制约了学生综合素质评价在高中的深入推进。

三是原有的综合素质评价方式难以满足我国教育考试招生制度改革政策真正落地的需要。2014年,国务院发布的《关于深化考试招生制度改革的实施意见》中提出“改革招生录取机制,探索基于统一高考和高中学业水平考试成绩、参考综合素质评价的多元录取机制”^[2],教育部发布的《关于加强和改进普通高中学生综合素质评价的意见》指出

“综合素质评价是对学生全面发展状况的观察、记录、分析,是发现和培育学生良好个性的重要手段,是深入推进素质教育的一项重要制度”^[3]。目前,综合素质评价已经在浙江、上海试点作为高考多元录取的重要参考,其他省市即将启动。原有的学生综合素质评价方式难以满足高校在较短的录取时间内高效使用的需求,与高校招生基本处于“软挂钩”的形式主义境地,录取的公平性、科学性难以保证。在新的历史时期,尤其是信息技术飞速发展的背景下,只有充分利用信息技术,学生综合素质评价系统才能真正对接政策实施的“最后一公里”。

放眼全球,许多国家都在学生综合素质评价领域进行了多年探索,目前也处于转型期。在综合素质评价方式、综合素质评价结果的使用、综合素质评价过程的监督机制等方面,美国、英国、加拿大、新加坡、韩国等国的方案各有特色。在综合素质评价方式上,大致可以分为三种类型:一是通过学年考试或平时课程成绩等形式组成的科目考试评价;二是采用表现性评价的方式评价学生某些非学术能力;三是通过学生提交的证明材料对学生综合素质进行评价。在综合素质评价结果的使用上,分为两种类型:一是直接将综合素质评价的成绩作为大学入学成绩的一部分,如韩国、新加坡;二是综合素质评价成绩仅作为高校招生的参考依据,如加拿大的高中报告单。在综合素质评价过程的监督机制上,也大致有两种不同的倾向:一种倾向是加强评分的客观性,另一种倾向是依赖良好的社会诚信文化。但令人遗憾的是,目前还没有哪一种方案能够在全球范围内进行推广。究其原因:第一,由于国情和各国面临的教育问题迥然不同,不能直接照搬复制;第二,已经开展综合素质评价的“先行者”本身也在进行着“破而后立”的新尝试。以美国为例,其学生综合素质评价的相关研究开始较早,理论研究深入,实践成果也较为丰富,为美国高等教育选拔人才发挥了积极的作用。然而,2017年,由近百所美国顶尖私立高中组成的联盟 Mastery Tran-

script Consortium(MTC)提出了一种全新的学生评价体系,与美国高中传统的成绩单相比,它不打分,不评级,只是持续追踪、记录、评估学生8大能力,目前已获美国大学申请系统 Coalition for Access, Affordability and Success(CAAS)的支持^①。这反映了综合素质评价是各国基础教育共同思考和尚处在探索过程中的问题。

综上,质性与量化的问题,统一与个性的问题,真实性、科学性、权威性、诚信等问题,这些问题的复杂程度,使得综合素质评价成为一个世界性的难题。就我国而言,中考、高考改革方案中,虽然明确了学生综合素质发展目标和综合素质评价要求,但仍存在着素养目标难以表述、培养过程难以考量、达成情况难以评价等难题。此外,在综合素质评价实施过程中,各方对如何处理学术能力与非学术能力、全面性评价与个性化评价、定性评价与定量评价、注重过程的评价与注重结果的评价等关系难以取得共识。因此,在综合素质评价的问题上,亟须进行大胆的改革和尝试。

清华大学附属中学(以下简称“清华附中”)秉承“自强不息,厚德载物”的精神,历来重视对学生的全面培养,始终把立德树人放在学校教育工作的首位。经过长期不懈的努力,清华附中切实营造了“全员、全过程、全方位”的育人体系及氛围。十余年来,学校坚持贯彻落实社会主义核心价值观,重视培养学生核心素养,在学生素质教育方面积累了良好的经验,获得了社会的广泛认可。面对社会发展的切实需求、教育改革的迫切呼唤和人才竞争的严峻形势,清华附中责无旁贷,以评价引导和促进学生全面成长、个性发展,努力培养出能够满足党、国家、人民和时代需要的人才,促进中考、高考改革的推进和落实,这是学校投入力量研发“清华附中学生综合素质评价系统”(以下简称“清华附中综评系统”)的初心。

2 学生综合素质评价系统的设计与研发

在深入考察国内外综合素质评价实践现状的基础上,参考古代的“功过格”,借鉴国外现代教育理论,清华附中构建了基于大数据的学生综合素质生成性评价模型,同时借助云计算、大数据等现代信息技术,自主开发完成学生综合素质评价技术平台。该系统通过对上传的海量学生数据进行整理、分析与挖掘,建立了学生综合素质发展模型,通过不断反演推理找出行为背后有价值的信息,实现了评价内容全覆盖、评价引导全过程、评价应用全方位,尤其是评价内容和评价结果可根据使用者的需求进行动态量化设计,能够很好地满足多方对评价体系的需求。该系统的设计、研发主要经历以下三个阶段:

第一阶段,通过聚类和相关性分析,确定“高影响力”活动类型。自2009年起,认真梳理学生在校期间参与的涉及思想品德、人文艺术修养、领导力、公益活动、体育锻炼、团队合作和国际视野等多个方面近100项教育教学活动,采用聚类分析方法和相关性分析,考察不同活动对学生发展的影响程度,搞清楚学生能做什么,应该鼓励学生做什么,进而总结归类了对学生发展有持久深远影响的8个方面的“高影响力”活动。这些活动是:1)社会公益及志愿服务,旨在培养学生志愿服务精神和社会责任感;2)学术志趣及偏好发展,旨在鼓励学生发现个人学术偏好并作出合理规划;3)艺术素养及特长培养,旨在促使学生提升个人艺术修养,丰富自身生活;4)体质健康与体育锻炼,旨在改善和提高学生体能,磨炼意志品质,培养团队精神;5)感动感悟与交流沟通,旨在引导学生正确面对学习、生活和交往中遇到的困难和问题,提高抗挫折能力;6)读书分享与人文思索,旨在引导学生养成爱读书、会读书的习惯;7)社会调查与勤工助学,旨在引导学生走进社会,了解现实生活,更客观地规划自己的未

^① 使用CAAS申请系统的学校包括哈佛大学、耶鲁大学、普林斯顿大学、哥伦比亚大学、斯坦福大学、康奈尔大学、达特茅斯学院、杜克大学、密歇根大学等80余所美国著名高校。

来;8)阶段小结与个人反思,旨在引导学生在总结和反思中不断提升自己。

第二阶段,结合相关教育理论,构建学生综合素质生成性评价模型。借鉴加德纳的多元智能理论、斯腾伯格的成功智力和伯恩斯坦的“表意性秩序”等教育学理论,构建了学生综合素质生成性评价模型,采取自下而上的方式对学生各个方面进行观察、记录、分析,建立了包含9个模块46个维度的行为记录内容,全面记录学生校内外的成长轨迹,从中挖掘学生的发展潜能和倾向,为每一个学生个性、自主、健康和全面发展创造条件。该模型打破了传统的采用构建多级指标体系的自上而下的评价模式,主张学生的综合素质可以通过学校教育来发展,可以通过学生行为来体现,从而可以通过学生全面客观的行为记录、过程积累和发展变化来进行评价。

第三阶段,利用大数据分析、开发学生综合素质评价系统。大数据时代来临,信息技术能充分挖掘教育数据的价值,助力学生成长记录,推动学生

综合素质评价的实施。我们借助互联网技术研发了学生综合素质评价系统的实验版,首次实现了云计算、大数据与综合素质评价相结合,充分发挥“评价过程”的影响力。在实验过程中,系统逐步尝试对学生上传的海量数据进行整理、分析与挖掘,初步为学生建立成长模型。

通过以上研发阶段,一套以过程积累的大数据为评价依据,以促进学生发展的“高影响力”活动为评价内容,采用过程性和发展性评价方式,强调多元化评价主体,以实施公示、互相监督保证评价的真实性和客观性,以满足使用者不同需求的“学生综评发展报告单”(见图1)为评价结果的“清华附中综评系统”最终问世。

3 学生综合素质评价系统的内容和特点

根据《教育部关于加强和改进普通高中学生综合素质评价的意见》中对评价内容和评价程序的要求,清华附中综评系统包括9个模块46个维度(见表1)。这9个模块46个维度,并非评价内容的指标

学生综评发展报告单

| | | | |
|----------|--------------|-----------|------|
| 学生信息 | 姓名:刘 | 性别:女 | 学籍号: |
| | 电子邮箱: | | |
| 学段选择 | 学年:2016—2017 | | |
| 生成日期 | 2017年09月27日 | | |
| 打分模板 | | | |
| 综评积分 | 148 | | |
| 素质模块 | 模块总分 | 记录维度 | 维度总分 |
| 模块1-艺术素养 | 72.5 | 才艺奖励 | 67.5 |
| | | 艺术成果展示 | 5 |
| 素质模块 | 模块总分 | 记录维度 | 维度总分 |
| 模块2-学业水平 | 67.5 | 学业奖励 | 57.5 |
| | | 创新成果 | 10 |
| 素质模块 | 模块总分 | 记录维度 | 维度总分 |
| 模块4-诚信道德 | 1 | 社会公益及志愿服务 | 1 |
| 素质模块 | 模块总分 | 记录维度 | 维度总分 |
| 模块5-活动实践 | 2 | 生产劳动 | 2 |

图1 学生综评发展报告单

表1 学生综合素质评价系统

| 诚信道德 | 学业水平 | 身心健康 | 艺术素养 | 组织协调能力 | 活动实践 | 个人成长 | 集体奖励 | 其他 |
|-----------|---------|--------------|--------|----------|--------|-----------|--------|------|
| 道德奖励 | 学业成绩百分制 | 《国家学生体质健康标准》 | 才艺奖励 | 班内任职 | 活动实践奖励 | 学术志趣及偏好发展 | 班集体奖励 | 好人好事 |
| 失信扣分 | 学业成绩五级制 | 身体机能 | 艺术成果展示 | 校团委学生会任职 | 党团活动 | 艺术素养及特长培养 | 社团集体奖励 | …… |
| 纪律处分 | 学业成绩二级制 | 运动机能 | | 学校社团任职 | 社团活动 | 体质健康与体育锻炼 | | |
| 违法犯罪 | 作业表现 | 体育奖励 | | 社会工作 | 生产劳动 | 感动感悟与交流沟通 | | |
| 社会公益及志愿服务 | 课堂表现 | | | | 勤工俭学 | 读书分享与人文思索 | | |
| 班级值日 | 课堂考勤 | | | | 军训 | 阶段小结与个人反思 | | |
| 课程班值日 | 学业奖励 | | | | 参观学习 | | | |
| 文明礼仪 | 会考成绩 | | | | 社会调查 | | | |
| 集会表现 | 创新成果 | | | | | | | |

体系,而是对学生行为进行记录的内容。这样的系统设计不仅方便学生分门别类地对活动进行记录,而且方便平台系统对行为大数据进行分析,可以直观呈现学生、学校、区域的教育发展情况。另外,9个模块46个维度的系统设计是基于清华附中教育教学实际,各地可按照教育部的要求,因地制宜进行动态化设计调整,这可以很好体现学生综合素质评价“记录在中学,评价在大学”的材料使用要求。

清华附中综评系统在评价程序设计上,贯彻落实了教育部要求的“坚持公正性,严格规范评价程序,强化有效监督,确保评价过程公开透明”精神,在写实记录环节,要求师生客观、写实记录学生成长过程中可考察、可比较、可分析的突出表现;在整理遴选环节,要求实时整理遴选有意义的活动上传;在公示审核环节,记录内容在全校范围实施公示,接受全校教师、同学、家长的监督;在形成档案环节,电子版档案记录了学生不同模块、维度的表现,系统后台对学生上传的活动等级分数、评级分数等数据进行自动处理,可得出学生不同模块、维度的分项积分和总积分,学校管理员可在工具箱中

导出“学生综评发展报告单”;在材料使用环节,平台系统可以根据使用者的不同需求对结果进行动态量化设计,这样的统计功能不仅能帮助教师及时对学生进行教育教学指导,而且能更好地服务于学校教学,服务于区域教育管理,服务于高校招生。

在监督机制上,该系统可以由社会各界全员监督。在记录提交环节,要提供佐证材料,并遵守诚信承诺。在平台中设有“同学圈公示窗口与质疑功能”,教师或其他同学可发表评论或质疑。如果其他同学的质疑被认定为有效,记录造假的学生将在诚信道德模块的“失信扣分”记录维度中被扣除5分;反之如果质疑被认定为无效,提出无效质疑的学生将在“失信扣分”记录维度中被扣除1分。

在综评分数模板的使用上,本系统具有动态量化功能,各使用地区或学校可根据实际情况,确定每个素质模块和记录维度的基本分数、等级分数、地域等级分数、教师评级分数等赋分标准,由系统自动计算学生综合素质评价积分。譬如,根据北京市某高中“诚信道德”素质模块3个记录维度的赋分标准(见表2),某学生获北京市级道德奖励一等奖,则该项活动记录得分 $S=1 \times 3 \times 2.5=7.5$ 分。

表2 某高中“诚信道德”素质模块3个记录维度的赋分标准

| 记录维度 | 基本分数 | 等级分数 | 地域登记分数 | 教师评级分数(星级) |
|-----------|------|--|---------------------------------------|---|
| 诚信道德奖励 | 1 | 特等奖:4 一等奖:3 二等奖:2 三等奖:1 其他奖项:1 荣誉称号:1 | 年级:1 校级:1.5 区级:2 北京市级及以上:2.5 | |
| 社会公益及职员服务 | 1 | | | 1星级:1.1 2星级:1.2 3星级:1.3 4星级:1.4 5星级:1.5 |
| 班级值日 | 1 | 优:2 中:1 差:-1 | | |

另外,学生综合素质评价系统可生成丰富多样的图表,见图2至图4。这些功能既能帮助学生关注自身和同伴发展,发掘自身潜能与特长,激励学生自主发展;又能帮助教师进行实时过程评价,有效指导学生发展,关注学生发展趋势,与家长实时沟通;还可以帮助学校实时统计分析各个年级、各个班级乃至各个学生的得分表现,总结学生发展规律,调整教育教学策略,实现科学而高效的学校教育教学管理,最大限度发挥引导激励、指导发展和评价管理的作用。

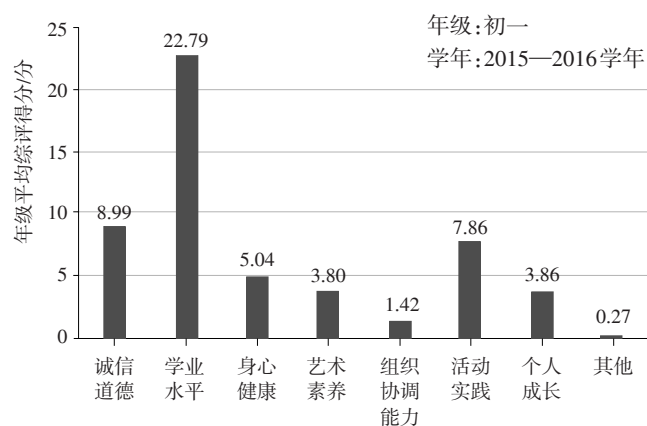


图2 分模块的年级平均综评得分

学生综合素质评价系统,强调用多把尺子衡量学生,对学生各个方面进行观察、记录、分析,注重以学生行为为依据,关注学生成长的过程性与发展性,注重参与主体、评价体系、结果呈现和发展趋向的多元性,注重数据的动态量化和真实性。其根本目的是帮助学生真实地了解自我,明晰自身的优势特长和发展短板,让每一个学生都找到适合自己成长的方式,实现学生发展和学校教育教学管理、各级学校人才选拔的双向共赢,发挥综合素质评价育人导向与评价选拔的双重功能。

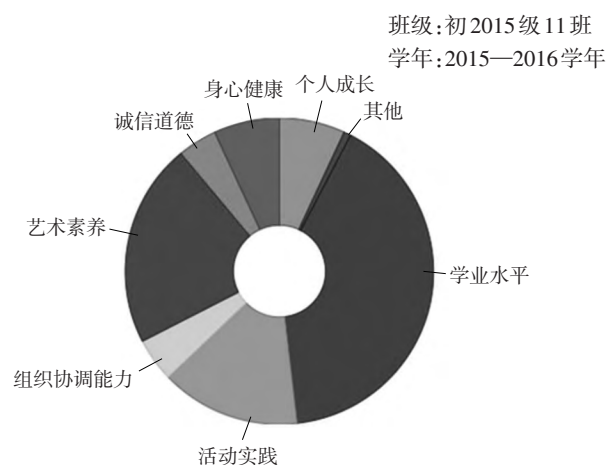


图3 班级素质模块分布图

| 学生姓名 | 学号 | 综评得分 |
|------|----|--------|
| 刘xx | | 300.50 |
| 李xx | | 199.50 |
| 张xx | | 188.00 |
| 闻xx | | 186.50 |
| 严xx | | 185.50 |
| 万xx | | 178.50 |
| 张xx | | 174.00 |
| 鲁xx | | 160.50 |
| 曹xx | | 160.50 |
| 仇xx | | 157.00 |
| 宋xx | | 153.50 |
| 王xx | | 153.00 |

图4 学生综评排名图

4 学生综合素质评价系统的使用、反馈与未来

学生综合素质评价系统在2013年完成平台开发之后,首先在清华附中一体化的5所学校试用,后来一些中学先后加入试用行列,为系统的完善和功能提升积累了宝贵的经验。

在试用阶段,该系统也曾遇到是否增加负担、教师使用不积极、学生使用不规范等难题,后通过加强教师、学生培训等办法将上述难题一一破解。师生实时提交过程性记录,这些精力和时间分散在学习的全过程中,不会带来很大的负担。通过开展师生培训,让教师切实感受到系统的使用有助于辅

助日常教学,与常规的学业成绩录入并不重复,真正从内心认同并接受该系统。

目前学生综合素质评价系统已收集学生记录近20万条。北京市教育委员会认为该系统具有服务学校教育教学工作的改善与提高、实时监测教育教学质量、有利于制定科学有效决策、指导区域教育发展的优势,并于2017年在全市所有高中正式推广使用该评价系统。

未来,学生综合素质评价系统将建立起“小学—初中—高中—后高中时代”的应用链,为优秀人才培养发展的长期追踪奠定基础,为学生的学业、升学与生涯规划提供帮助,为中国教育的实证研究积累最完备的大数据。通过对大数据的挖掘开展更深层次的教育应用,激发出更大的教育活力。

参考文献

- [1] 教育部关于积极推进中小学评价与考试制度改革的通知[EB/OL]. (2002-12-18) [2017-11-10]. http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_32/201001/xxgk_78509.html.
- [2] 国务院关于深化考试招生制度改革的实施意见[EB/OL]. (2014-09-03) [2017-11-10]. http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_1778/201409/174543.html.
- [3] 教育部关于加强和改进普通高中学生综合素质评价的意见[EB/OL]. (2014-12-10) [2017-11-10]. <http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/s4559/201412/181667.html>.

Development and Application of the Comprehensive Student Quality Evaluation System Based on Big Data: Innovative Practice of the High School Attached to Tsinghua University

WANG Dianjun, JU Hui, MENG Weidong

(The High School Attached to Tsinghua University, Beijing 100084, China)

Abstract: In exploring the core issues of comprehensive student quality evaluation, the High School Attached to Tsinghua University draws lessons from modern educational theories and, with the help of modern information technologies such as cloud computing and big data, has taken the lead in developing a comprehensive student quality evaluation system. To implement the fundamental task of cultivating morality in students, and following the

(下转第66页)

A Meticulous Design of Balancing the Administrative Authority of the Imperial Examination in Feudal State: On the Shift of the Administration of the Civil Service Examinations

DU Chengxian

(East China Normal University, Shanghai 200062, China)

Abstract: In the 24th year of Kaiyuan in the Tang dynasty (736), a combat broke out between Li Ang, the supervisor of the civil service recruitment examination and Li Quan, one of the candidates. The combat resulted in a significant reform in the administrative authority of the Imperial Examination: the civil service recruitment examinations supervised by the Ministry of Personnel were placed under the control of the vice director of the Ministry of Rites. It's a meticulous design made by the reformer to balance the administrative authority of the Imperial Examination: the match of affairs and power, the decentralization and the balance between the two departments, and the limited exclusive authority. The reform laid the foundation for the administrative authority of the Imperial Examination in the successive feudal dynasties in China.

Keywords: Imperial Examination; examination administration; balance of authority; supervisor of the civil service recruitment examination; Tang dynasty

(责任编辑:陈睿)

(上接第52页)

laws of physical and mental development of students, the system focuses on students' developmental process and behavior, emphasizes the diversity of the participants, evaluation system, results to be presented and development trend, and attaches importance to the authenticity and dynamic quantification of data, with a view to helping students truly understand themselves, be clear about their own strengths and weaknesses and ultimately find their own way to grow. Over the four years of practice, the system, with innovative features such as progression, diversity and measurement, has proved to be a two-way win-win platform for student development, school management and college admissions and has performed a dual function of providing guidance for talent cultivation as well as assessing and selecting students for different types of colleges.

Keywords: college entrance examination reform; comprehensive student quality evaluation; morality cultivation; multi-intelligence theory

(责任编辑:陈睿)