
上海青少年学生科学评价体系研究

席居哲¹

(华东师范大学 上海 200062)

【摘要】: 上海应加快推进青少年学生科学评价体系的建构、完善与定制;抓紧做好“评价共同体”的建立工作;发挥高考综合试点改革的辐射带动作用,推动新一轮教育综合改革不断走向纵深;做好青少年学生科学评价配套技术的研发;借助互联网+做减法,通过大数据做加法,立体搭建上海青少年学生科学评价体系,服务于上海教育综合改革和上海高考综合试点改革。

【关键词】: 青少年学生 科学评价体系 教育综合改革

【中图分类号】: C913.5.51 **【文献标识码】:** A **【文章编号】:** 1005-1309(2018)08-0086-011

一、青少年学生科学评价体系研究的重要意义

(一)人才培养的检测器和校准器

青少年学生科学评价体系的构建与实施,在实质上,是将教育支柱理念解析化、具体化和操作化的过程。因此,只有确立青少年学生科学评价体系的引领地位,人才培养与选拔的观念才能真正得到落实。在此意义上,青少年学生评价体系的完善与定制,是教育综合改革迈向纵深的标志。青少年学生评价体系的内容与方法将成为一个国家和地区的育人导航仪和选才风向标。青少年学生评价体系的建构与实施,对于课程与教学和人才素质产生全面而深刻的影响。作为教育系统化“顶层设计”的“操作说明书”,青少年学生评价体系为课程与教学改革提供指导和启发,对新课标制订和教学组织形式创新提供工作指南。各级各类教育,尤其是高中及以前学段的教育将以青少年学生评价体系为准,进行教育方法创新与教育行为设计。所以,教育综合改革的系统、持续向纵深推进,离不开青少年学生评价体系的完善与定制。

(二)教育重构的关键保障体系

国内外核心素养框架的提出与确立,开启了主要世界经济体的教育重构。自世界经济合作与发展组织(OECD)发布《为了成功人生和健全社会的核心素养》报告以来,不少世界经济体、国家和地区相继提出自己的核心素养体系框架。这些核心素养体系框架旨在应对时代变化和未来发展,主要解决“现在”教育把“未来”公民培养成什么样的人这样一个根本问题。青少年学生科学评价体系作为教育改革的准绳,是确保教育重构方向与进程的可控性,促成教育价值实现的机制保障。教育改革要做得扎实、收到实效,必须要保证教育始终行进在正确的轨道上,其标志是目的明确、方向正确、路径准确,这些均依赖于教育评价对精确性的追求。而要使评价更具生态效度,那种重课程轻社会,重具体学科轻跨学科,重知识学习轻技能掌握,重智力培育轻非智力因素成长,重竞争轻协作,重终结性评价轻过程性评价,以及主要来源于教师而不是多信源评价的青少年学生传统评价模式,必须要予以革新。

¹基金项目:上海市决策咨询研究重点课题(编号 2017-A-026-B)。

作者简介:席居哲,教育学博士,华东师范大学心理与认知科学学院副教授。

(三) 国家人才战略的迫切需要

中国当前面临着从人力资源大国向人力资源强国的人才战略转型,提升人力资源开发质量是当前教育的重要使命。从世界范围看,随着人力资源越来越成为国家经济社会发展的战略性资源,制订教育规划、推动教育综合改革就成为抢占国际竞争制高点的重要举措。源于核心素养框架在世界多个组织、国家和地区引发的青少年学生科学评价变革也在持续进行中,并具有了鲜明的竞争意味。发展青少年学生科学评价体系,是教育服务人才强国战略的关键点,也是教育综合改革的一个难点。

(四) 教育综合改革纵深推进的破题点

青少年学生科学评价体系是教育综合改革的逻辑节点。学生评价关联着教育目标、课程内容、教学方法、人才培养与选拔、师资培育、教育管理诸层面,是教育综合改革诸环节的集结点。如果缺乏一套操作性较强的青少年学生科学评价体系,前述教育任一环节的改革都会遇到瓶颈。青少年学生科学评价体系是教育综合改革的重要抓手。教育综合改革错综复杂,涉及的层面非常广泛,而青少年学生科学评价正处在所有问题的矛盾焦点上,做好青少年学生科学评价体系的构建、完善与定制工作,是教育综合改革向纵深推进的合适切入点。青少年学生科学评价体系是教育综合改革的工作支点。以青少年学生科学评价体系为支点,围绕评价进行顶层设计和系统部署,教育综合改革杠杆的力矩就可以达到最大化,以评价推动课程改革,通过评价检验教学效果,以评价指标衡量学生发展结果,使教育综合改革始终走在可自行迭代的循证路上。

(五) 探讨上海高考综合改革试点改革工作的着力点

一是系统回顾和梳理已有国内外核心素养体系框架,探寻这些核心素养框架的指标交集,结合中国特别是上海独特的文化背景与教育传统,进行青少年学生科学评价体系的内容重构,从而完成该体系的完善与定制,真正体现教育为时代变化、为未来社会准备的核心素养诉求。二是紧扣当前国际评价发展趋势,从多主体、多方法、多渠道、多时空的视角来设计更具生态效度的青少年学生科学评价体系,使其具有可操作性,结合互联网+和大数据,提出更具随时操作性的青少年学生科学评价发展愿景。三是紧密结合当前上海高考综合改革试点改革的使命与背景,从顶层对青少年学生科学评价体系进行框架设计,从底层对内容指标进行可操作化设计。

二、核心素养框架的内容分析

(一) OECD 的核心素养框架

OECD 的核心素养框架从社会心理学角度选择核心素养并界定具体内容,提出以实现个人成功生活与发展健全社会为基础,确定了三类核心素养:一是交互使用工具的能力,具体包括交互使用语言、符号和文本的能力;交互使用知识与信息的能力;交互使用技术的能力。二是在异质社会群体中有效互动的能力,具体包括与他人建立良好关系的能力;合作能力;管理与化解冲突的能力。三是自主行动的能力,具体包括适应宏大情境的行动能力;形成并执行个人计划或生活规划的能力;保护及维护权利、利益、限制与需求的能力。OECD 组织统筹的国际学生评估项目(PISA),从与生活相关情境、学生须具备的能力、相关的科学知识和学生的态度4个维度对学生能否参与社会所需的知识与技能进行评估。PISA 与“成人素养与生活技能调查”是 OECD 核心素养框架的两大年龄段(未成年人、成年人)的测量操作化。

(二) 欧盟的核心素养框架

欧盟核心素养框架是欧洲理事会为应对全球化、知识经济和信息时代的挑战,把新基本技能(new basic skills)作为优先策略。始于2001年的“教育与培训2010项目”,旨在到2010年建立起知识社会所需要的欧洲教育和培训新体系,其核心即形成欧洲核心素养框架。2005年,欧盟正式发布《为了终身学习的核心素养:欧洲参考框架》,其目标之一是支持各成员国确保其所

培养的年轻人在基础教育与培训结束时,具备一定水平的核心素养以使他们能够应对生活,并为未来学习和工作打下基础,并且确保这些国家的成年人能够在人生中不断发展和更新自己的核心素养。2006年12月,欧洲议会和欧洲理事会联合批准了这一框架,该框架即成为欧盟各成员国适应信息时代教育的纲领性文件。2006年版的欧洲核心素养框架主要包括8个方面,涉及使用母语交流、使用外语交流、数学素养和基本的科技素养、数字素养、学会学习、社会与公民素养、首创意识与创业精神、文化意识和表达,每一素养又从知识、技能和态度3个维度进行具体描述;强调:基本的语言、文字、数学、信息能力是终身学习的基础,而“学会学习”将支持个体的所有学习活动,批判性思维、创造性、主动性、问题解决、风险评估、决策、情绪管理等也是终身学习者不可或缺的重要素养。

(三)美国的核心素养框架

在关注受教育者未来职业发展方面,不仅有美国学者研发的类似职业兴趣测验(如霍兰德职业兴趣自测)、职业性格测验(如迈尔斯-布里格斯类型指标 MBTI)等心理测验及风靡市场的职业咨询,美国政府也特别注重从系统导向上审视教育与职场的对接问题。美国劳工部在1990年成立职场基本技能秘书委员会(SCANS),目标是识别和确认年轻人在职场中获得成功所必需的技能。1991年,SCANS发表了《职场对学校教育的要求》,列明美国职场基本技能的五大职场基本素养(包括资源管理、人际素养、信息素养、系统化素养和技术素养)。为达到这五大职场基本素养,必须依赖于学生个人的心智素养和个性素养,SCANS提出学生应具备三大基本素养,即基本素养(读、写、算、听、说)、思维素养(批判思考、制定决策、解决问题、想象能力、学会学习、推理能力)和个人特质(责任感、自尊感、社会性、自我管理、正直诚实)。1994年,美国国家素养研究院又启动了“为未来而准备”(EFF)项目。这一项目旨在为成人学习研制核心素养指标体系,主要包括有沟通技能(阅读理解、通过写作表达观点、有效的语言表达、积极地倾听、批判性的观察)、决策技能(解决问题和作决策、计划、运用数学解决问题并与他人沟通)、人际技能(与他人合作、指导他人、提议和影响力、解决冲突和协商)和终身学习技能(对学习负责任、研究性学习、反省和评价、使用信息和通信技术)。2002年,美国成立21世纪技能联盟(P21),主要是探寻那些能让学生在21世纪成功的技能,从而建立了21世纪技能框架体系。该框架体系由两部分构成:一是核心学科与21世纪主题;二是21世纪技能。核心学科包括:英语、阅读或语言艺术、世界语言、艺术、数学、经济学、科学、地理、历史、政府与公民,其目标指向是让学生像学科专家那样地思考。21世纪主题包括全球意识、金融、经济、商业和创业素养,公民素养、健康素养、环境素养。这些21世纪跨学科主题,着眼于培养学生的跨学科意识和运用多学科知识解决复杂问题的能力。21世纪核心技能主要指学生在未来工作和生活中必须掌握的技能、知识与专业智能:“学习与创新技能”(包括创造力与创新、批判思维与问题解决、交流沟通与合作)、“信息、媒体与技术技能”(包括信息素养、媒体素养、信息通讯(ICT)素养)、“生活与职业技能”(包括灵活性与适应性、首创精神与自我导向、社会与跨文化素养、生产性与责任制、领导力与负责心)。

(四)亚洲地区的核心素养框架

新加坡基于21世纪劳动力需求特点,提出了“思考型学校和学习型国家”这一建设愿景,其目标是培养自信的人、主动的学习者、积极的贡献者和热心的国民,即4个理想教育成果。其核心素养框架包括核心价值、社交及情商能力、21世纪技能。核心价值包括尊重、诚实、关爱、有弹性、和谐、负责,在素养框架中起决定作用,决定着培养什么样的社交及情商,再进而决定着21世纪需要培养哪些技能,最终指向政府提出的4个理想教育成果。

我国香港特别行政区教育局在《学会学习——课程发展路向》中,将课程架构分为学习领域、共通能力、态度与价值观3个相互关联的部分。其逻辑关系是:共通能力(9项)贯穿在学习领域(8项)中,态度与价值观则渗透在各学习领域。态度和价值观单列,可以看出香港地区21世纪核心素养框架对此的重视。价值观影响态度和信念,是形成态度和信念的基础;价值观分成核心价值和辅助价值,前者是普适性的价值观,后者在运作上有助于维持核心价值观。在此核心素养框架中,分为个人和社会两个方面,分别提出个人核心价值(含生命神圣、真理、美的诉求等11项)和个人辅助价值(含自尊、自省、自律等14项),社会核心价值(含平等、善良、仁慈等21项)和社会辅助价值(含多元化、正当的法律程序、民主等)。此外,态度包括了乐于参与、批判性、首创精神等18项。单从项数上看,就足以看出香港特别行政区政府对价值观的重视程度。

我国从2001年开始实施基础教育课程改革,“情感态度价值观”和“知识与技能”“过程与方法”共同构成三维目标观,体现在学科课程中的各个环节。2014年开始的新一轮课程改革中,落实立德树人被当作课程改革的根本任务。教育部以文件的形式明确要求把培育和践行社会主义核心价值观融入国民教育的全过程,倡导富强、民主、文明、和谐、自由、平等、公正、法治、爱国、敬业、诚信和友善。教育部组织研究学生发展的核心素养体系,明确学生应具备的适应终身发展和社会发展需要的必备品格和关键能力,2016年推出“中国学生发展核心素养”。在核心素养中的社会参与大类下的“责任担当”,具体指学生在处理与社会、国家、国际等关系方面所形成的情感态度、价值观和行为习惯,具体包括社会责任、国家认同和国际理解等基本要点,涉及崇尚自由平等,维护社会公平正义,明辨是非,具有规则和法治意识,具有国家意识,了解国情历史,认同国民身份,能自觉捍卫国家主权、尊严和利益,具有文化自信,具有热爱党、拥护党的意识和行动,理解、接受并践行社会主义核心价值观,具有中国特色社会主义共同理想,有为中华民族伟大复兴梦而不懈奋斗的信念和行动等要点。

(五) 内容趋势特点及关注点分析

国内外核心素养框架内容有以下趋势特点:一是战略性。诸核心素养框架是各组织、国家和地区主动为未来做教育准备,具有明显的战略性。二是综合性。表现为科学素养与人文素养兼顾,个人成长与职业发展兼顾,时代变化应对与未来社会准备兼顾,个人、国家和世界兼顾,价值观与知识技能兼顾,课堂学习与社会实践并重,把育人和选才相结合等。三是系统性。这些核心素养框架在应对时代变化和未来发展双重驱动下,要求对面向未来的受教育者做更系统的发展规划。往往是在共识推动下和政府主导下,多方专家参与下,进行系统设计的过程,反映了包括学习科学、教育科学、心理科学等一系列科学前沿。四是共通性。人类立足于现在为未来作准备,在世界对信息时代人的发展目标这方面,有着许多共同的追求,这构成了核心素养框架的交集。这个交集兼有认知性素养和非认知性素养,体现了信息时代和知识社会的新要求,这些交集素养至少包含有协作、交往、创造性和批判性思维等。五是差异性。在共通性之外,还有体现着不同国家和地区的文化特色的差异性。比较典型的,美国的核心素养框架散发着浓郁的职场气息,而俄罗斯的核心素养框架则包括文化休闲、日常生活的内容。六是全龄段。大多数代表性核心素养框架不只关注青少年学生,而是将视域扩展到从幼儿至老年的全年龄段,从终生视角考虑核心素养框架的建构。

从国内外多个核心素养框架的成果看,如果依据素养的综合性特征,把情感、态度和价值观范围的相关内容有机地融合到相关条目中,并对内涵相近、层级相当的条目予以合并,可得到两个范畴(领域素养和通用素养)共18项核心素养。其中领域素养包括基础领域素养6项(语言素养、数学素养、科技素养、人文与社会素养、艺术素养、运动与健康素养),新兴领域素养3项(信息素养、环境素养、财商素养);通用素养包括高阶认知3项(批判性思维、创造性与问题解决、学会学习与终身学习),个人成长2项(自我认识与自我调控、人生规划与幸福生活),社会性发展4项(沟通与合作、领导力、跨文化与国际理解、公民责任与社会参与)。从各项核心素养出现频次看,沟通与合作、信息素养、创造性与问题解决、自我认识与自我调控、批判性思维、学会学习与终身学习、公民责任与社会参与等7大素养为各国组织和经济体高度重视。

三、科学素养评估框架分析

(一) 美国 NAEP2009 科学评估框架

美国教育进展评估(National Assessment of Educational Progress, NAEP)是美国全国学生学业成就评估体系,其通过测量学业成就趋势来衡量全国教育目标达到与否,对国家基础教育质量监控产生至关重要的作用。2006年NAEP推出新的“NAEP2009科学评估框架”,根据广泛接受的国家科学教育标准和评估水平进行设计,同时也考虑到地方的课程标准。评估框架旨在提供科学素养评估标准,打造国际竞争人才,使美国年轻人在科技方面保持国际竞争力。科学内容和科学实践是NAEP2009科学评估的两大内容。科学内容涉及物理、生物、地球与空间科学3个领域;科学实践则分为识别科学原理、应用科学原理、实施科学探究、进行科学设计4个维度。

(二) 国际数学和科学评测趋势(TIMSS)

国际数学和科学评测趋势(The Trends in International Mathematics and Science Study, TIMSS)是由国际教育成就评价协会(the International Association for the Evaluation of Educational Achievement, IEA)发起和组织的国际教育评价研究和评测活动。项目旨在通过评测学生在四年级、八年级段已学到的有关数学和科学的概念、过程和养成的态度,从国际视角评估四年级学生和八年级学生在数学和科学上的学习进展,鉴别学生从四年级到八年级在数学和科学知识 with 技能上的发展,帮助各国在与高水平的学生学习成就相关的关键因素上做国际性的比较(如在课程政策、教学、资源之间的比较)。

(三)高中学生资优特征的评估

众多研究表明,资优不是单一表征:一方面,资优学生通过不同方式表现其资优才华,表现在他们相对异于常人的智力、创造力、工作专注等方面;另一个方面,由于他们敏感性高、企图心强、自主性大,具有高度坚持性和自尊心,在自我认定和获取他人认同上,往往有较大的困难;尤其绝项资优的学生,常有超越凡俗的想法,不容易接受现实,也不易被现实所接受;如果他们不被了解,很容易被视为异类或异端,遭受更多的发展困扰。从诸多资优学生发展与教育的研究文献中不难发现,带着上述两方面特征的资优学生,经常表现出超乎同龄人的卓越能力、独特的个性特征以及相对脆弱的发展动机系统。许多学者对学业资优学生的特征进行大量的探索和总结,Renzulli 的“三环”资优理论强调,能力资优特征评估需要同时关注 3 个维度,即一般能力、特殊能力和任务专注力。Clark(2002)对此进行特征列表,该特征分类和评估内容包括认知/思维特征(能记忆大量信息;高度的理解能力;广泛的兴趣和高度的好奇心;高水平的语言发展和口头表达能力),情绪/情感特征(对于其他人的情感异常敏感;敏锐的幽默感;较高的自我意识;想要变得与众不同;理想主义;正义感强烈),生理/知觉特征(较高的知觉意识;在生理发展和智力发展方面不同步;不太能容忍自己本身的运动技能和对自我的标准之间的差异),直觉特征(对于直觉经验非常开放;在所有努力的领域有显著的创造性),以及社会特征(受到自我实现需求强烈的激励;在对社会问题的概念化和解决上有较强的能力;领导力;对于社会的元需求有兴趣,如公正、真理、美等)。而高创造性资优特征评估包括有智力过程、知识、智力(认知风格)、人格特征、动机特征等方面。

(四)趋势特点及关注点

上述科学素养评估的框架着力评估青少年以下方面:一是认知思维。主要涉及统合统整能力,表现为形成整体观的能力、从多个角度来审视问题、问题解决能力的迁移(即触类旁通)等;批判性思维与创造性思维,包括客观性(实事求是的态度,强调精实纪录,遵循严谨的程序等),批判性(质疑论据的可信度、认识到会有不同诠释视角、对所获信息的求证倾向等)和创造性(喜欢动手创作、喜欢探寻如何殊途同归、联想与头脑风暴使用等)。二是人格倾向。包括经验的开放性,表现为未知性(充满求知欲,富于好奇心、喜欢发问,兴趣广泛、喜欢新点子和新事物,接受变化等),乐于接受挑战(喜欢冒险,乐于接受新知,乐于推断原来的设想,不满足于现状,追求成长和进步等),特别注重内心感受和审美感受性(重视自己的直觉经验,对新的经历持开放态度,对艺术特征非常敏感,注重审美情趣等);模糊的容忍度,表现为对容忍模糊(接受不确定并不因此而焦躁,接受甚至偏好复杂情景事情,不回避不确定的环境),冒险倾向(有冒险特质,乐于接受挑战,富于野心,喜欢追求新鲜刺激等),决策倾向(不盲从权威,不因循传统,独立判断,坚信没有绝对的对与错等)。三是情绪表现。主要包括情绪的动力性,涉及进取心(兴趣广泛、内在动机驱动、跃跃欲试)、自得其乐(充满热情学习、自我决定、积极投入)与成就感(自信、自我掌握感);情绪的坚持性,包括聚精会神(沉迷于任务、享受任务执行与完成过程,心无旁骛等),自我调节(有幽默感会自我解嘲,自制、善于调节自己的情绪,不冲动、审慎决定等);情绪的坚定性(明显的自我意识,乐观主义、富于理想,信守承诺、有责任感等)。四是团队合作倾向。主要涵盖交流与沟通(自我表达,倾听分享,交流学习等),执行与协作(欣赏与尊重他人,宽容与信任,规划能力与分工意识),认同与影响(包括说服力,提出新颖意见,清晰的角色定位等),态度与承诺(极具责任心,善于整合消除成员矛盾,有果断决策力和敏锐洞察力等)。

综上所述,对青少年学生科学评价体系研制有如下启示:一是应将学科、跨学科与非学科的评价有机结合;二是评价不仅要注重结果,更要注重品质以及得到结果的过程;三是既要评价个人面对任务中或在某种情境中的表现,也要评价个人与他人一起面对任务或某种情景中的表现。

四、国际学生评价技术分析

(一)档案袋评价法

这是一种以美国教育学家杜威的“学生为中心”“教学民主化”的教育思想为理论基础,以肇始于 1980 年代评价改革思想为思想基础的一种发展性评价方法。该方法通过学生本人、教师、家长、同伴甚至社会人士,将反映学生学业表现、心理发展等的材料进行分门别类、有计划有选择性地收集,通过对这些材料的分析和解释,反映学生在生活、学习和心理上的发展、成就及进步状况。因受杜威“从做中学”“教育即生活”教育理论的影响,档案袋评价法强调学生在评价中的自我参与性和主观能动性。通过档案袋评价法,可显示学习者的成长进程,收集的是学习者全程的多元信息,便于对学习者的进行立体、全面的了解。美国学者格莱德勒依据档案袋的不同功能,将其分为理想型、展示型、文件型、评价型以及课堂型等 5 类,见表 1。

表 1 格莱德勒档案袋分类系统说明

类型	内容/功能
理想型	通过一段时间的成长,帮助学习者成为自己学习历史的思索者和非正式的评价者
文件型	用于记录学生的成长、进步与成就,内容包括学生作品,同伴反映,教师对其在德、能、勤、绩等方面的评价、观察、考查意见和表现性测试结果
展示型	用于向家长和其他人展示学生本人选择出的最好、最喜欢的作品、成果及获得的荣誉证书、获奖证书等
课堂型	是一种总结性文件,用于向家长、管理者报告学生情况,包括:根据教育教学目标描述所有学生取得的成绩及表现情况的总结,教师的年度课程、教学计划及修订说明,教师给每个学生的德智体等方面的观察、评语
评价型	根据预定的标准,对学生所取得的成绩及表现情况进行标准化的评价,用于向公众和国家机关、企事业单位作报告

资料来源:杨洁、岳大鹏:《档案袋评价法与国际学生评价项目的对比与思考》,江苏教育学院学报(社会科学),2012 年

(二)评价中心技术

评价中心(Assessment Center)起源于 1929 年德国心理学家建立的一套用于挑选军官的多项评价过程,与传统的以心理测验面试等方法相比,评价中心技术具有强调工作的情景模拟性、强调受测被试的实际工作能力以及多种测评方法综合应用等鲜明特点。情景模拟是评价中心技术的核心思想,主要测评技术有公文筐测验、无领导小组讨论、管理游戏和角色扮演等,不同的技术具有不同的适用范围,因此宜根据不同的测评指标及项目应使用不同的测评技术,见表 2。

表 2 评价中心主要测评技术对比分析

测评技术	可调查的能力	复杂程度	使用频率(%)
管理游戏	战略规划能力、团队协作能力、领导能力等	非常复杂	25
公文筐测验	判断能力、计划组织能力、管理控制能力、授权能力、决策能力等	很复杂	81
角色扮演	人际关系能力、言语沟通能力、说服能力、应变能力等	很复杂	未调查

无领导小组讨论	言语表达能力、人际影响力、组织协调能力等	很复杂	59
个人演讲	说服力、语言表达能力、应变能力、沟通能力、战略思维等	较复杂	46
案例分析	业务知识、理解能力、文字表达能力、态度、价值观等	较复杂	73
事实判断	信息搜索能力、把握事实能力、决策技能等	简单	38
模拟面谈	口头交流能力、人际关系能力、问题解决能力等	简单	47

(三)评价技术分析

从档案袋评价法和评价中心技术不难看出,青少年学生科学评价体系研究除了设定指标外,还须对指标进行评价。一是多信源、多主体的评价更为稳健;二是主观评价与客观量化相统一,使评价更立体;三是把过程性评价和终结性评价相结合,使评价更具描述性;四是多种方法并用,根据不同的指标、侧面动用不同的方法,使评价更具生态效度;五是数据采集的代表性、汇报整理的概括性、资料记录的丰富性,使评价更具简洁性。

五、上海普通高中学生综合素质评价的探讨

(一)问题与分析

1. 上海普通高中学生综合素质评价分析。

一是科技素养着力较弱。虽然在创新精神与实践能力模块包含了创新思维、调查研究能力、动手操作能力和实践体验经历等相关指标,但在科技素养作为基本的对于科学技术的理解能力,至少还应包括科技知识、科技能力、科学方法、科技意识和科技品质等层面。二是新兴领域素养体现不足。对国际上日显重要的信息素养、环境素养、财商素养等方面,鲜见其踪影。而这些新兴领域素养既体现了时代变革的特点,也反映着科技高度发达的21世纪社会公民素养的走向。三是社会性发展素养彰显不够。国际诸代表性核心素养框架多数均重视沟通与合作,领导力、跨文化与国际理解等,在目前上海普通高中学生综合素质评价系统中,缺乏明确评价依据。四是个人成长素养相对受到忽视。虽然诸如自我认识、人生规划与幸福生活方面的素养在国外核心素养体系中的关注度也有待提升,但从趋势研判,上海应对个人成长素养予以“前瞻性”收纳。

2. 上海普通高中学生综合素质评价指标的解构与可操作性。

一是信息管理系统而非真正的评价系统。目前的系统还是一个分门别类的原始记录信息管理系统,尚不是一个可以提供具体化、可视化数据的分析系统。这无疑加大了相关使用方(比如高校招生部门)的工作负荷。由于仅有原始资料可供查看,不仅不同学生的异同很难把握,哪怕对同一学生,不同评价者的评价结果也可能大相径庭。二是缺乏对指标的详细评价解构。指标评价解构是为评价者提供评分依据的参照系,缺乏这样一套体系,无法进行快速而有效的评价,这种“无据可依”的局面导致,让本应成为评价者的(如老师)成为资料审核者;二是本应成为评价结果使用者的(如高校招生人员)成为学生评价者。三是量化程度不够。由于指标缺乏解构和可操作性,目前信息管理系统尚不能提供学生素质指标的量化考核结果,也不能了解一个学生的纵向发展变化情况,更无法提供该生在全班、全校乃至全市的相对位置信息,不能为高考招生部门提供坚实、有效的参照依据。这种情况实际上造成了“仅供参考”,这与研制高中生综合素质评价系统的初衷和上海高考综合试点改革的精神显然不符。

3. 上海普通高中学生综合素质评价方法问题分析。

一是评价信源比较单一。目前学生上传材料,高中老师进行审核公示的做法,未能很好地体现评价主体多元化。主体多元化是获取立体信息的重要渠道,多信源评价是今后学生评价的主要特点之一。二是记录载体不够丰富。学生提交材料以个人报告为主,虽然有与其他机构的信息交换,但记录载体形式不够丰富,较难通过评价让决策者对目标学生形成一个立体、全面而生动的印象。三是终结性评价居多。这种评价方式容易成为课程内容、学习活动、学校管理、教学方法和学生学习的控制器,从而易滑向唯评价论和功利主义,使得最终的结果与评价意欲达到的目的相去甚远。

(二) 解决对策

1. 上海普通高中学生综合素质评价内容的增益与重构。

一是增益体系内容。在现有体系基础上,对模块进行丰富,对指标进行细化。比如,在创新精神与实践能力模块,提升科技素养的具体指标,纳入认知思维倾向(综合统整能力、批判性思维),人格特征倾向(经验开放性、模糊容忍度),情绪表现倾向(情绪的动力性、情绪的坚持性)等内容。同样,也可将沟通与合作、领导力、跨文化与国际理解等社会性发展素养纳入本模块。涉及个人成长素养的内容,可考虑纳入身心健康与艺术素养模块。对新兴领域素养,如信息素养、环境素养和财商素养,可考虑纳入创新精神与实践能力模块。二是重构指标体系。之所以考虑重构指标体系,一方面当前的四模块表述对新指标的吸纳力有限,而当前课堂与课外的二元分离结构,限制了对高中生进行综合素质的全面把握。比如,持续学习力(学会学习和终身学习理念下)、科技人文素养(对科技学习、掌握尤其运用过程中的人文精神)这样的重要素养培养与评价,不能单单靠课外就能完成。实际上,所有素养的培育及其评价基础都离不开对当前教课程组织的创新。同时,宜考虑纳入国际教育、心理领域发展的新趋向。在重构指标体系方面,为贯彻《上海市青少年发展“十三五”规划纲要》,以及“卓越的城市,出彩的青少年”的主题,诸如积极心理学、积极教育领域强调的多种积极心理品质(如抗逆力/心理弹性、好奇心、爱学习、勇气、正直、有活力、爱、善良、公正、宽容和仁慈、谦逊、欣赏、感恩、希望等)也宜纳入青少年学生科学评价体系中。

2. 上海普通高中学生综合素质评价指标的解构与可操作性。

一是编制指标解析指南。对每个指标的详细描述,为评价者提供一个可参考的评价模板。例如,在现有系统中创新精神与实践能力模块下,有“创新思维”这一维度,可先将其分解为项目,再把每一项目分解为表现特征描述(评估观测点),形成一套清晰的指标解析手册,使评价者有一个清晰的参考框架(见表3)。二是设定量化等级。对行为表现程度的分数评定,既依赖于一套严密的数据评价体系,也依赖于一套清晰可借鉴的标准。因此,宜通过李克特(Likert)等级评分(也可选视觉类比,如游码卡尺式计分)形式进行。在评定分数时,同时给出质性评语的位置,便于评价者从量与质两个视角进行。表4列出了可能的一种计分提示表。三是自动列出数据信息。多层化显示信息,可显示出某学生在全市、本区、本校、本年级、本班的位置信息。可视化数据显示,以图的形式直观(比如直方图、拆线图、雷达图等)展现学生综合素质维度得分情况。提供纵向和横向两方面的信息:纵向的显示学生自记录以来的发展变化情况;横向的显示学生与其他同龄人(从全市至本班)相比所处的相对位置。这样便于高考招生人员一目了然知晓某生的基本情况。同时,在数据最底层,将量化分数与原始记录资料链接,便于高考招生人员对目标学生进行更深入的了解,以方便决策。

3. 上海普通高中学生综合素质评价方法改进。

一是务求评价主体多元化。多信息源的评价可为评价对象提供更客观、立体、公正的评价,确保评价的生态效度,这种做法可有效避免方法偏差,也可有效避免弄虚作假的情况。除了突出教师在评价中的主体作用外,家长参与、学生自评和互评都是很重要的信息渠道。表5列出了指标方面的架构样例。二是记录载体多元化。除了文字记录外,图片记录、声音记录、影音记录等也可考虑使用,这样会使得证据更生动、更立体。特别是类似于无首领讨论体现的领导力与沟通协作能力,科技小活动中体现的

创造性、批判性思维以及坚持不懈的坚韧性等,更易通过影音流的记录展现。三是增加过程性评价。在信息技术高度发达的现代社会,过程性评价记录将更为便捷。可开发移动终端数据记录平台,随手拍、随手发、随手评,构建立体的高中生综合素质评价系统。

表 3 创造性思维指标解构示例

维度	项目	表现特征描述
A 创造性思维	A-1 客观性	实事求是的态度,不偏颇采信,持平审视争议; 认识到科学探究并不一定要遵循固定程序,但其中通常包括收集数据、逻辑推论及运用想象来构思假说和解释数据; 强调精实纪录、重复实验证实等,是维持科学可信赖性的基础
	A-2 批判性	质疑论据的可信度、因果关联性、逻辑一致性或推论过程严密性; 认识到科学产生过程虽然严谨,却可能因新现象被发现或新的观察角度的改变而有不同诠释; 对他人的信息或报告提出合理的求证和质疑
	A-3 创造性	有自己的想法,喜欢自己动手,并相信自己可以处理很多事; 注意利用不同办法思考或解决同一问题,并做出相同结果; 运用联想、头脑风暴、概念图等程序形成创意或新想法

表 4 一个 5 级评分制样例

得分	标准
1	未有明显表现
2	表现比较欠缺
3	表现一般
4	表现比较突出
5	表现非常突出

表 5 一个多信源评价架构样例

内容	信源	评价点
诚实守信	教师	不作假,不说谎,不欺骗别人;说到做到,言出必果
诚实守信	家长	不隐瞒自己的真实想法,不掩饰自己的真实感情;兑现承诺
诚实守信	学生	言行一致,不虚伪;一诺千金,重信用

注:这种做法类似编制基于评价主体(评价主体特定性的)的复本测验;有些复本测验在主体间一致性高,比如教师和家长,但学生群体与成人群体的评价着眼点可能有所不同,在设计多信源评价时需要考虑到这个方面(当然不排除 3 个主体间也有许多评价面上的一致,比如“感恩”“心理弹性”等)

4. 大数据:

上海高中生综合素质评价系统的形态走向。海量的云存储与超强的云计算, 综合素养评价的内容多样性、评价信源多元性、记录载体多样性和过程性评价比重渐增, 这些都共同指向一个重要的上海高中生综合素质评价系统的形态: 大数据系统。大数据系统将突破以下限制: 一是与低龄学段的连贯问题; 二是数据记录容量和记录时间瓶颈; 三是数据分析与输出的限制; 四是评价内容、评价主体、评价方法的多元性; 五是依循多分析路径(如某维度、年龄、某时刻等)聚焦的关键事件输出整理瓶颈; 六是全市或某区、某校学生综合素养的现状与发展态势等。

六、做好青少年学生科学评价, 服务于上海高考综合试点改革

(一) 上海高考综合试点改革的成效

1. 推动应试教育改观, 提升育才质量, 利于国家选才。

原来运行的“3+1”模式, 学生的学习只需紧盯4门科目, 这易造成学生对一些科目学习的“策略性忽视”。一方面难以避免唯分数论, 自然地滑向应试教育; 另一方面, 造成学生知识面狭窄, 不适应高校学习和人才培养要求。“3+3”模式拓宽了学生的知识域, 这不仅利于他们在高校的学习和适应, 更利于学生高质量成才, 更好地彰显了“教育面向未来”。

2. 催发学生主动精神, 增进学习意愿, 利于积极成长。

从“3+1”变为“3+3”, 学生可选择思想政治、历史、地理、物理、化学和生命科学中的任意3个科目, 自主选择利于诱发学生的主动精神, 利于保护他们的学习兴趣和激发学习动机。而学习意愿的增强往往与学业质量的提升密切相关, 是积极成长的重要体现。

3. 通过分散考试减负, 降解考试压力, 利于水平发挥。

原来的高考, 所有4门考试科目要在每年6月7日、8日集中完成, 每一科目只有一次考试机会, 这无形中增加了考试压力。新的高考做法则体贴得多: 语文、数学、外语3门科目为6月统一考试, 但外语实行一年两考, 以较高分纳入高考总成绩; 3门选考科目则以学业水平等级考形式进行, 做到随教随考随清。这样的方案利于降解考试压力和考生水平发挥。

4. 汇聚两依据一参考, 综合评价考生, 利于科学选拔。

结合统一高考成绩和自主选择学业水平等级考试成绩, 参考普通高中学生综合素质评价信息, 对学生进行综合评价, 这确保了人才选拔程序的科学与严谨。同时, 这种安排将促使学校教育与唯分数论和应试教育彻底分离。通过综合评价的方式选拔人才, 符合国际趋势, 也是对我国从人力资源大国走向人力资源强国战略的政策回应。

5. 最大限度保护公平, 分数面前人人平等, 给异秉学生以公平机会, 利于“不拘一格降人才”。

上海高考综合试点方案力求在考试、评价和招生诸环节确保更高水平的公平, 打破“一局定胜负”、唯科目考试分数的高考局面, 不仅牢牢守住“分数面前人人平等”的底线, 而且通过程序机制, 给那些有不同特长、不同爱好者以公平的发展机会。

(二) 上海高考综合试点改革亟待解决的问题

1. 在知识与素养的关系上, 仍倚重于知识评价。

从“3+1”到“3+3”，虽然使考生学科基础相对“广博”了，但评价依然主要是在学科知识层面，尚缺乏关于科学人文综合素养的评价。尽管在高校自主招生中，多少会涉及此类评价，但其对青少年发展的教育引导作用远不及他们所处的中小学大。科学人文素养之所以重要，是因为如果撇开科学学科去培育人文素养，有可能会给如何把握和使用科学知识留下一块“心灵飞地”（如会导致科学主义的“功利性”“工具性”偏斜），学生可能成为掌握丰富科学知识和高科技手段的“空心人”。因此，亟须将知识评价与人文素养评价相结合。

2. 在评价形式上，仍多赖于终结性评价。

终结性评价虽有许多优点，但这种评价方式容易成为课程内容、课程组织、学校管理、教学方法和学生学习的控制器，从而滑向唯分数论和应试教育。形成性评价因为重在过程，可以提供学生更丰富的成长信息与优势剖面，其既能作为个体成长的参考，也可作为人才选拔提供脉络性信息。就评价形式而言，“3+3”可视为从终结性评价向形成性评价过渡的起步。但通过机制性设计，逐渐增加形成性评价在人才选拔中的比重，并在教育过程中适当运用诊断性评价（如对学生学业情绪、学习策略等）以调适教与学双方的行为，均是十分必要的。

3. 在评价主体上，学生及家长的主动性有待进一步彰显。

评价多元化是国际趋势，也是更公平的评价方式。评价多元化除了评价内容、评价方法的多元化外，评价主体的多元化也是其重要构件。英国教育改革中推行的三级学生评价体系就颇具启发性。该评价体系包括了国家考试、教师对学生的评价和来自学生的评价。其后者又包括学生的反思性自评和同伴互评。多主体/多信源评价，有利于勾勒出更接近真实的学生综合情况。

4. 综合素质评价体系的内容有待进一步修订和完善。

上海普通高中学生综合素质评价系统主要涵盖品德发展与公民素养，修习课程与学业成绩，身心健康与艺术素养，创新精神与实践能力和学校特色指标等方面，考虑得比较周全，但该综合素质评价体系在诸如科学人文素养、积极心理品质等方面尚较缺乏，而这些却是国际心理、教育领域的新近重要研究和实践议题。该体系仅记录学生在高中3年的数据，其与高中以下学段如何衔接，也是一个值得探讨的问题。

（三）上海高考综合改革下一步的工作着力点

1. 加快进行上海青少年学生科学评价体系的建构、完善与定制。

一是结合上海青少年学生特点和国内外有关青少年综合素养评价研究的新进展，搭建和检验上海青少年学生科学评价体系，使其起到为青少年学生积极成长提供教育依据、为青少年学生发展成才提供参考依据之功效。结合国际青少年心理、教育学科的新进展，充分体现前瞻性，把握让每一位青少年科学、积极、自由、充分发展这一价值目标，进行上海青少年学生科学评价体系的系统研究。二是上海青少年学生科学评价体系宜体现时间上的连贯性，能较好解决当前高中生综合素质评价体系只为高中3年这一问题，做好在学段上的有机连通，使青少年成长情况有更长的轨迹可循。三是结合当前上海青少年学生出现的一些值得关注的倾向特点，特别在生活热情、生命活力、谦逊礼让等方面给予指标权重。四是上海青少年学生科学评价体系至少包括生活热忱（热爱生活、珍惜生命、爱国爱校）、知礼尚德（谦逊礼让、尊重自尊、助人感恩）、认知思维（统整性、灵活性、创造性）、情绪表现（动力性、成就感、稳定性）、人格特征（开放、自尊、容忍）、社会责任（忧患意识、担当性、正义感）这些要点。五是真正执行汇聚评价法，体现多信源评价特点，宜将客观与主观评价相结合，过程性评价和阶段性评价相结合，量与质的评价相结合，他评与自评相结合，最大限度体现评价的多元性。

2. 抓紧做好“评价共同体”的建立工作。

一是对学校管理者、教师、家长、其他相关人员进行教育评价观念的更新至关重要。二是通过培训、网络、及时协助等形式,对评价者进行技术指导。三是营造“包含认真的一票”的教育评价氛围,将评价主体、相关机构有机连在一起,使之成为评价共同体而不是各自为政、按己所好。四是做好评价体系相关研究成果的持续发表,以增强其说服力。五是对评价有异议者进行个别辅导。六是评价的目的不是增加工作量,而是丰富了评价手段,其最重要的功能之一是真正解放学生、科学发现学生和全面认识学生。

3. 发挥高考综合试点改革的辐射带动作用,推动新一轮教育综合改革不断走向纵深。

一是教育管理者、教师的教育理念、教学方法更新,以此驱动教育综合改革的车轮。二是在学校中推行科学人文素养教育、积极教育和全人教育(whole person education),以丰富当前教育行动的价值。三是依托家校联合体和社区,推进家庭教育理念与方法的渐次更新。四是促进学科教育与素养培育的全方位融合。

4. 做好青少年学生科学评价配套技术的研发。

一是开发易于操作、用户交互界面友好的评价移动终端平台,便于评价随时随地进行。二是依托互联网,构建上海青少年学生发展科学评价大数据库。借助互联网+做减法,通过大数据做加法,立体搭建上海青少年学生科学评价体系,服务于上海教育综合改革和上海高考综合试点改革。