

职教动态与参考

(2021 年第 5 期——“教育评价改革”特辑Ⅱ)

(总第 18 期)

文艺职院 职业教育研究中心编印

2021 年 6 月 28 日

目 录

◎政策解读

中央全面深化改革委员会会议：完善科技成果评价机制..... -1-

◎专家解读

在教育评价改革中落实立德树人的根本任务..... -3-

挖掘过程性数据中的教育价值..... -8-

利用信息技术促进教育教学评价改革创新..... -13-

◎专家论坛

职业教育学习质量评价的内涵界定和困境表征..... -19-

职业教育学业评价方法刍议..... -23-

中央全面深化改革委员会会议：完善科技成果 评价机制

新华社北京 5 月 21 日电 中共中央总书记、国家主席、中央军委主席、中央全面深化改革委员会主任习近平 5 月 21 日下午主持召开中央全面深化改革委员会第十九次会议，审议通过了《关于完善科技成果评价机制的指导意见》。

习近平在主持会议时强调，加快实现科技自立自强，要用好科技成果评价这个指挥棒，遵循科技创新规律，坚持正确的科技成果评价导向，激发科技人员积极性。

会议指出，党的十八大以来，党中央部署推进科技评价体系改革，聚焦“四个面向”的科技成果评价导向逐步确立，激励创新的环境正在形成，带动我国科技创新能力明显提升。同时，分类的科技成果评价体系尚未建立，指标单一化、标准定量化、结果功利化问题还不同程度存在。

会议强调，完善科技成果评价机制，关键要解决好“评什么”、“谁来评”、“怎么评”、“怎么用”的问题。要坚持质量、绩效、贡献为核心的评价导向，健全科技成果分类评价体系，针对基础研究、应用研究、技术开发等不同种类成果形成细化的评价标准，全面准确评价科技成果的科学、技术、经济、社会、文化价值。要加快构建政府、社会组织、企业、投融资机构等共同参

与的多元评价体系，积极发展市场化评价，突出企业创新主体地位，规范第三方评价，充分调动各类评价主体的积极性。要把握科研渐进性和成果阶段性特点，加强中长期评价、后评价和成果回溯，推进国家科技项目成果评价改革，健全重大项目知识产权管理流程，加强科技成果评价的理论和方法研究，引导科技人员潜心研究、探索创新，杜绝科技成果评价中急功近利、盲目跟风现象。要加快推动科技成果转化应用，加快建设高水平技术交易市场，加大金融投资对科技成果转化和产业化的支持，把科技成果转化绩效纳入高校、科研机构、国有企业创新能力评价，细化完善有利于转化的职务科技成果评估政策，鼓励广大科技工作者把论文写在祖国大地上。要改革完善科技成果奖励体系，重在奖励真正作出创造性贡献的科学家和一线科技人员，控制奖励数量，提升奖励质量。

（来源：新华社，2021 年 5 月 21 日）

在教育评价改革中落实立德树人根本任务

冯刚 张智¹

近日，中共中央、国务院印发《深化新时代教育评价改革总体方案》（以下简称《总体方案》），将“坚持立德树人，牢记为党育人、为国育才使命”作为重要原则遵循，将完善立德树人落实机制作为关键目标追求，重点围绕“五类主体”进行政策设计，呈现了在新时代教育评价改革中落实立德树人根本任务的具体路径选择。

一是强化对各级党委政府思想政治教育工作的评价。思想政治工作是一切工作的生命线。《总体方案》要求各级党委认真落实领导责任，建立健全党委统一领导、党政齐抓共管、部门各负其责的教育领导体制，把思想政治工作作为学校各项工作的生命线紧紧抓在手上，贯穿学校教育管理全过程。从教育评价的角度对思想政治工作体制机制的完善健全作了强调。通过对教育领导体制的评价改革，确保思想政治工作的生命线地位，推动树立科学的教育发展理念。《总体方案》要求各级党委和政府建立健全党政主要负责同志深入教育一线调研、为师生上思政课、联

¹ 冯刚.河南大学党委副书记、副校长，张智.河南大学教育考试与评价研究院教授、博士生导师.

系学校和年终述职必述教育工作等制度。这有助于缩减各级党委政府主要负责人与思想政治教育工作一线的距离，既助推其了解思想政治教育的现实状态，拿出更具针对性、时效性的思想政治工作政策举措，又能发挥其在思想政治教育中政治站位高、教育影响大的优势特点，提升思想政治教育工作的有效性和辐射面。

二是推进对各级各类学校教育质量评价标准的制定和实施。

在改革学校评价的政策设计中，落实立德树人根本任务是《总体方案》最根本的目标追求，要求加快完善各级各类学校评价标准，将落实党的全面领导、坚持正确办学方向、加强和改进学校党的建设以及党建带团建队建、做好思想政治工作和意识形态工作、依法治校办学、维护安全稳定作为评价学校及其领导人员、管理人员的重要内容，健全学校内部质量保障制度，坚决克服重智育轻德育、重分数轻素质等片面办学行为，促进学生身心健康、全面发展。在幼儿园层面，制定幼儿园保教质量评估指南；在义务教育学校层面，制定义务教育学校办学质量评价标准；在普通高中层面，制定普通高中办学质量评价标准；在职业学校层面，将德技并修作为重点评价内容；在高等学校层面，改进本科教育教学评估，突出思想政治教育。《总体方案》将立德树人要求融入教育质量标准的指标设计和构建过程，把各级各类学校思想政治教育工作的过程性节点和结果性要点精准转换为相应的教育质量评价标准，通过评价实施，定期公布评估结果、加强监测结果运用等方式倒逼被评价对象持续改进思想政治教育工作，不断

提升立德树人工作质量效果。

三是突出教师师德评价关键内容和考核任职评价改革。高校教师要坚持教育者先受教育。要加强师德师风建设，坚持教书和育人相统一，坚持言传和身教相统一，坚持潜心问道和关注社会相统一，坚持学术自由和学术规范相统一，引导广大教师以德立身、以德立学、以德施教。习近平总书记对所有教师提出了做“四有”好老师的要求，对思政课教师提出了“六要”的要求。教师的思想道德水平和政治素养直接关系到立德树人的工作的质量和效果，教师师德师风正、政治素养高是思想政治教育工作取得实效的前提条件。在教师评价改革中，《总体方案》坚持把师德师风作为第一标准，要求克服重科研轻教学、重教书轻育人等现象，把师德表现作为教师资格定期注册、业绩考核、职称评聘、评优奖励首要要求，强化教师思想政治素质考察，推动师德师风建设常态化、长效化。在考核任职方面，《总体方案》将上思政课作为高校领导班子成员年度述职的重要内容，要求学校党政管理干部选拔任用，原则上应有思政课教师、辅导员或班主任等学生工作经历；高校青年教师晋升高一级职称，至少须有一年担任辅导员、班主任等学生工作经历。

四是完善学生评价方式方法，树立以德为先的评价理念。对学生的评价导向牵引着学生的成长方向。《总体方案》强调树立科学成才观念，要求坚持以德为先、能力为重、全面发展，完善综合素质评价体系，切实引导学生坚定理想信念、厚植爱国主义

情怀、加强品德修养、增长知识见识、培养奋斗精神、增强综合素质。在完善学生评价方式方法上,《总体方案》强调根据学生不同阶段身心特点,科学设计各级各类教育德育目标要求,引导学生养成良好思想道德、心理素质和行为习惯,传承红色基因,增强“四个自信”,立志听党话、跟党走,立志扎根人民、奉献国家。通过信息化等手段,探索学生、家长、教师以及社区等参与评价的有效方式,客观记录学生品行日常表现和突出表现,特别是践行社会主义核心价值观情况,将其作为学生综合素质评价的重要内容。学生不同成长阶段会呈现不同的身心发展特点,根据不同阶段的身心特点设计德育目标要求,探寻多元主体参与评价的制度机制,突出日常行为表现的过程性评价,正是尊重学生成长规律和思想政治工作规律的必然结果。

五是强调务实用人评价,形成正确用人导向。教育的目的是为社会培养有用的人才。用人单位对人才的判断和选择,是人才培养质量的重要评判维度。用人单位的选人用人导向会直接影响人才培养的目标确定和方向选择。用人单位的选人用人标准忽视思想道德和政治素养,既不利于受教育者主动、积极提升思想政治素质,也不利于教育者形塑受教育者的健康人格。《总体方案》强调树立正确用人导向,要求党政机关、事业单位、国有企业带头扭转唯名校、唯学历的用人导向,建立以品德和能力为导向、以岗位需求为目标的人才使用机制。将品德作为首要的用人标准,破除了唯名校、唯学历,忽视思想政治素养的选人用人取向,既

有助于受教育者的积极修为，不断提升素质能力，减少精致利己主义者和高才低德者的出现，也有助于保证用人单位的和谐稳定、事业发展，进而为整个国家社会的繁荣进步提供保障。

教育评价改革是一个系统工程，是实现教育治理体系和治理能力现代化的重要举措。在改革评价中落实好立德树人根本任务，必须遵循教育治理现代化尤其是思想政治教育治理体系和治理能力现代化的理念要求，强调教育评价改革的精准性、系统性、综合性，协同推进针对上述“五类主体”的教育评价改革工作，做到相互呼应、相互配合、相互支撑、相互促进，如此方能真正通过教育评价改革调整，依托教育评价的“指挥棒”效应，充分助力培育担当民族复兴大任的时代新人。

（选自：中国教育报，2021 年 1 月 14 日）

挖掘过程性数据中的教育价值

辛涛²

2020 年 10 月 13 日,中共中央和国务院正式印发《深化新时代教育评价改革总体方案》(以下简称《方案》),《方案》指出,教育评价事关教育发展方向,要全面贯彻党的教育方针,坚持社会主义办学方向,落实立德树人根本任务,遵循教育规律,针对不同主体和不同学段、不同类型教育特点,改进结果评价,强化过程评价,探索增值评价,健全综合评价,着力破除唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾,建立科学的、符合时代要求的教育评价制度和机制。

一、关于教育评价改革的又一重要纲领性文件

2018 年全国教育大会指出要破解“五唯”顽瘴痼疾,扭转不科学的教育评价导向。全国教育大会以来,教育系统采取了一系列措施,包括出台规范校外培训机构、推进中小学生减负等文件、研制高考评价体系等措施来破解基础教育“唯分数”、“唯升学”现象,但这并非一日之功。时至今日,我国教育评价中的“五唯”仍根深蒂固,造成教育的功利化倾向严重。教育的根本任务是培养德智体美劳全面发展的社会主义建设者和接班人。其中,德智体美劳共同指向全面发展的人,五个方面相辅相成、相互促进。根据泰勒的目标取向评价模式,有什么样的教育目标,就会有什

² 辛涛.北京师范大学中国基础教育质量监测协同创新中心常务副主任,国家督学.

么样的评价标准与评价内容。“五唯”导致的教育目标窄化，必然会反映在对学生的评价标准上，我国学生长期以来备受诟病的“高分低能”便是例证。

以分数和升学率为主的评价导向如果不及时改变，必然导致各利益相关主体最大化地追求学生的考试成绩和分数，进一步助长校外培训热潮、加重学生学业负担。PISA 连续几轮测试均表明我国参测省市学生学业负担在国际上处于前列，并由此带来学生生活中情绪不健康等一系列问题。

针对这一顽瘴痼疾，《方案》指出，要“改进结果评价，强化过程评价，探索增值评价，健全综合评价”。这是落实全国教育大会破“五唯”，扭转教育不科学评价导向的一项重大的、实质性的举措。

二、在四大评价中，过程评价居于核心地位

其一，强化过程评价就是要扭转教育评价中只重结果评价的倾向。当然，强调过程评价不是要否定或摒弃结果评价，作为教育评价的一种重要类型，我们认为结果评价是必要和必需的。改进结果评价，是要摒弃结果评价中“唯分数”和“唯升学”的现象，使其发挥应有的作用。当前我国最重要的结果评价就是中高考，近年来我们也一直致力于中高考的改进，如高考实行选考、一年多考制度，将综合素质评价结果纳入招生录取等措施，就是在力图打破“一考定终身”“唯分数”的困局。而且我国多年来的改革经验也表明，过程评价与结果评价并不冲突，过程是结果的前提，结果是从过程中产生的结果，二者相依相存。因此，在

中高考改革中探索和贯穿过程评价的理念是可行与可为的。其二，过程评价从内涵上包含了增值评价。过程评价就是要对学生的整个学习过程和成长过程进行评价，其在某种程度上属于一种个体内差异评价，即将评价对象的过去与现在进行比较，以此评价学生的进步情况，这与增值评价具有内涵上的一致性。而且，若增值评价只关注学生的考试分数和成绩的进步状况，而不管学生在整个学习过程中其他方面的进步情况，则其也会不可避免地走向唯分数评价的藩篱。其三，综合评价更要注重过程评价。当前，综合评价在我国主要表现为要将考试分数和成绩以外的东西纳入进来，目前在基础教育阶段实行最彻底、也最成熟的则是综合素质评价。综合素质评价从其本质上来说是一种典型的过程性评价，其实施过程也是彰显育人价值的过程。但由于其自身还未形成较为成熟的理论与实践操作体系，加之评价结果的“高利害”使用，在实践操作中稍有不慎极有可能将其异化为结果性评价。如何发挥综合素质评价的过程性育人功能，彰显其育人价值是综合素质评价实施中要考虑的重要课题。

三、过程评价是破解“五唯”困局的核心与关键

强化过程评价，扭转不科学教育评价导向需从以下方面着手。

一是过程评价需要树立正确的教育评价观。行为目标的教育评价观支配和引领我国教育评价数十载，过程评价呼唤教育评价观由目标导向转变为过程与目标并重。与仅关注目标的达成度不同，过程评价将评价的视野转向学生的学习经验领域，凡是有价

值的学习结果都应纳入评价范围，即关注目标以外的学习结果。这种学习结果不仅是分数，是更广义上的学生经验到结果。因此，教育评价要更重视学生的学习过程和体验。除关注预设的教育目标，还要关注非预设的目标；除了关注学生成长过程中显性行为变化，更要关注其隐性素质的变化情况。

二是过程评价需要拓宽和深化评价内容。由于过程评价将评价的关注点从结果导向了整个教学过程，包括教师的教学过程、学生的学习过程以及师生互动过程，由目标的达成度转向了目标之外的实际教育效果，因此需要拓宽评价的内容。师生教学过程中的投入度、参与度、师生关系，学生学习过程中的自主性、创造性、合作能力、问题解决能力以及学生的态度、兴趣、动机等非认知因素都应纳入成为评价的内容。

三是过程评价呼唤与之相适应的评价方法体系。在评价方式上，过程评价应采用量化与质性相结合的评价方法，但要以质性评价为主，包括档案袋评价、表现性评价等在内的有利于展示学生学习与成长过程的质性评价应成为过程评价的主要方式。在评价主体上，由于评价发生在教学过程中，而不再游离于教学与学习之外，这就要求师生共同参与，尤其是学生的深度参与。

四是过程评价需要做好评价结果的反馈与利用。与结果评价不同，过程评价的最主要功能不是比较与鉴别，而在于发展与改进。过程评价是对学习过程进行评价，并在学习过程中实施评价，目的是为了发现问题、改进学习过程，是一个动态的、持续性的评价过程。

四、过程评价需要建立在数字世界之中

利用物联网技术对物理世界中学生学习过程信息进行采集、管理和存储；利用大数据技术挖掘数字世界中存储的学生学习过程、结果的各种数据以及相关因素的数据，建立相应的教育大数据预测、分类和智能评价模型；利用智能决策和可视化技术进行科学分析和决策。进而减轻教师、学校和家庭的评价负担，提高过程性评价的科学性和准确性，确保评价的常态化和规模化。如上所述，我国现行的综合素质评价是一种典型的过程性评价，综合素质评价的实施有利于过程评价的开展，要继续推进综合素质评价落地。因此，要建立“评价—反馈—改进”的运行机制，做好结果的反馈与利用，挖掘过程性数据中的教育价值，发挥评价对学生学习过程、教师教学过程的改进作用。

（选自：中国教育报，2020 年 12 月）

利用信息技术促进教育教学评价改革创新

近日，中共中央、国务院印发《深化新时代教育评价改革总体方案》，充分体现了鲜明的时代特征，明确提出“坚持科学有效，改进结果评价，强化过程评价，探索增值评价，健全综合评价，充分利用信息技术，提高教育评价的科学性、专业性、客观性”。传统教育教学评价存在评价方法单一、评价内容片面、重智育轻德育、重分数轻素质等问题，在以往的技术条件下很多问题难以解决。随着大数据、人工智能、区块链等新兴技术快速融入教育领域，为推动教育教学评价改革创新提供了条件，有助于开展面向教育教学全过程的纵向评价以及包括德智体美劳全要素在内的横向评价，推动评价方式和评价内容的重构，为教育教学评价改革创新提供可行途径。

一、智能时代教育教学评价创新的需求

智能时代教育教学评价改革是时代需求、技术进步、教育发展等多方面共同作用的结果，反映出智能时代对教育教学评价创新的迫切需求。

智能时代人才培养需求的变化要求教学评价机制的创新。智能时代深刻改变了社会各行业对劳动者知识、能力、素质的要求，进而改变了教育行业的人才培养目标，亟待建立以学生为中心的知识为基、能力为重、德育为先的德智体美劳人才培养体系。学生的批判性思考能力、协作沟通能力、解决复杂问题能力、创新

能力以及“人技”互动能力，成为其面向未来智能时代的关键竞争力。针对新时代人才培养要求，必须创新信息化条件下的教与学方式，发展面向每个人、适合每个人的教育体系。而教育评价是新型教育体系的关键一环，必须通过创新教学评价机制，发挥教育评价的指挥棒作用。

新兴技术在促进教育教学评价方式创新上发挥着重要作用。创新教育评价工具，以大数据、人工智能技术为代表的新兴技术为实现评价方式多元化提供了便利。例如，基于大数据分析技术的评价，不但可以全方位、全过程采集教学数据，而且可以获得情感因素、心理倾向、实践能力等非结构化数据；通过跟踪和记录学生的学习过程并适时发起学习干预，可为教师和学生提供动态、实时的评价反馈，有利于及时调整教学进程；而在区块链等技术支持下，可以设计开发面向终身学习的学分银行，有效构建终身教育立交桥等。

深化教育评价改革是克服“五唯”教育导向的关键途径。2018年全国教育大会上，习近平总书记提出坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾的重要指示。在现有人才培养体系下，教学评价内容主要关注的是学生对知识的掌握情况，有可能导致死记硬背、“机械刷题”等现象，对德智体美劳全面发展考查不足。要重构教育教学评价体系，利用信息技术深化教育评价改革，必须在教育评价内容、评价主体、评价方法、评价工具等方面提供具有针对性的解决方案，支持形成性和个性化的教学评价。

二、教育教学评价创新典型趋势和案例

为进一步加强教育教学过程和结果的绩效监测，教育评价创新受到了主要发达国家的广泛关注，涌现出一批典型案例。

强调能力培养，利用学习档案记录成长轨迹。在评价过程中让学生了解到知识偏好、能力缺陷、发展目标等内容，避免以应试成绩为尺度评价教育质量。例如，美国知名私立高中联盟 MTC 采用全新的学生评价体系，评价内容涵盖创造性思维、数理能力等八大类，实现个性化的学生综合评价。

关注学情预测，通过数据分析给予反馈干预。通过监测学生可能存在的学业风险，为这些学生提供个性化的帮助。例如，采集来源于学生信息系统中的基本信息数据和网络学习平台中的学习行为数据，利用定制算法对每一位学习者进行学习预测，以辨别其是否存在学习困难的可能。

注重教学监测，提供基于大数据的自适应学习。以学习者原有的学习储备和个体特征为起点，动态匹配符合“最近发展区”的学习服务。例如，通过对学生在线学习进行跟踪，精准分析学生的认知投入水平和学习目标达成情况，及时推送定制化的学习计划和内容。

三、信息技术推进评价改革的创新途径

信息技术支持的教育教学评价，既要注重全局性的战略布局和顶层设计，也要关注课程与教学层面的信息化学习方式变革，智能时代需要的教育教学评价改革创新途径主要表现在过程评价、增值评价和综合评价等方面。

强化过程评价，注重教育评价的动态性和诊断性。传统的教育评价方法以总结性评价方式为主，而通过基于技术的数据采集和分析将评价渗透到教学环节之中，对自然状态下评价对象的真实学习行为轨迹进行跟踪，可以实现基于数据分析证据的差异化和适应性教学，提供对教育数据的全过程采集和教育结果的适时反馈，确保学习效果。

优化增值评价，关注学生努力程度和进步表现。开发多样的评价模型，不以学生考试成绩作为评价的唯一标准，通过大数据、人工智能技术采集学生知识、情感、态度、思维和行为等全过程数据，从发展性角度评估学生的努力程度、学习绩效等，关注学生在原基础上的进步程度，形成纵向比较，激发学生内生动力，引导学生德智体美劳全面发展。

健全综合评价，强化评价主体和手段的多元性。教学评价主体的多元性表现在教师评价、家长评价、同伴评价、自我评价甚至是机器评价的融合发展，推进内部评价与外部评价的整合；教育评价方法的多元性即采用基于智能技术的试题测试、实践操作、面试答辩等不同方法进行综合评价，增强教学评价的客观性、公正性和有效性。

此外，利用智能技术改进结果评价，尤其是优化考试流程，实现从组卷、阅卷到考试管理等各方面的整体提升，也是需要关注的重要内容。

四、技术支持下教育教学评价改革举措

信息技术能够使教学评价手段更加丰富、教学评价过程更加科学、教学评价结果更加准确，探究技术推动下的教育教学评价变革策略，有益于指导基于技术的教学评价创新实践。

首先，健全学生评价综合指标体系，突出品德、能力和素质导向。探究信息技术支持下的教育教学评价适用范围和约束性条件，厘清技术支持下教育教学评价与教学的关系。综合考虑德智体美劳全面发展要求，开展基于数据的规模化测评。由原来结果导向的“单一”评价向过程导向的“多维”评价转变，由主要注重知识传授向更加注重全面发展转变。

其次，创新多元化评价方式，开展基于学习行为的数据分析。利用信息技术跟踪和监测教学的全过程，如课堂考勤、课堂表现、历次考试分数、作业练习分数等。建立学生线下/线上、校内/校外学习和活动的成长档案，全面记录和追踪学生校内外的成长轨迹，强调评价的诊断功能、激励功能、预测功能、调节功能等，发现其潜质和不足，服务学生全面发展和个性成长。

再次，开发智能化关键技术和系统，为教育教学决策提供支持。研发数据采集管理平台，部署物联感知、图像识别、视频采集、平台采集等软硬件采集技术设备，攻克关键技术。支持人技结合的智能评价，加强多源数据之间的互联互通，利用技术实现对区域、学校、学生发展情况的过程监测，加快评价结果的反馈速度，实现以评促教，为管理者的教育决策和教师的教学决策提供数据参考。

最后，推进典型区域重大考试评价制度改革，开展应用试点示范。加大基础评测模型和关键技术研发力度，推进教育评价智能化产品的创新，推动基于数据的过程化评估在教育教学中的应用实践，助力教育科技企业加快智能评测技术研发和成果转化。选取几个典型区域推进重大考试评价制度改革实践，将过程化评价结果纳入重大考试评价范畴，发挥典型示范作用。

（选自：《人民教育》，2021 年第 21 期）

职业教育学习质量评价的内涵界定和困境表征

目前,我国职业教育学生学习质量评价在评价取向、评价内容、评价方法、评价结果等方面存在不足,难以彰显职业教育的本质特征。

一、职业教育学习质量的内涵界定

1. 学习质量的内涵概要。质量产生于生产过程,衡量质量的依据是产品满足顾客需要的程度。因此,学生作为学习主体,其学习质量应该指学习活动满足学生需要的程度。换言之,“学习质量是指学习活动的过程和结果满足学生提升自身素质需要的程度,其通常表现为学生对学习过程的主观感受和学习成果达到预期目标的程度。”

2. 职业教育学习质量评价的内涵诠释。从词源学来看,评价(evaluation)的含义是引出和阐发价值,是对事物做出判断和赋值的过程。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,并给予价值判断的过程”。在职业教育领域,学生学习质量评价是依据一定的评价标准,采用科学的评价方法,对学生学习活动的过程和结果进行综合价值分析与判断的过程。为此,教育部于2019年12月颁布《关于做好扩招后高职教育教学管理工作的指导意见》,旨在“改革学生学业考核评价方式方法,实行多元评价,

把实现高质量就业作为检验人才培养质量的重要标准”，为完善学生学习质量评价体系提供政策保障。

二、职业教育学习质量评价的困境表征

1. 评价取向普教化，忽视主体发展特征。评价取向是实施教育评价活动的理念基础。学生的学习，是个体经过学习后获得知识、技能、能力、情感、态度等的动态发展过程。因此，对学生学习质量的评价，理应是“对其学习过程的连续性观察和评判，而不应仅仅是对其学习的最终结果或者所谓的‘教育产品’的评价”。然而，当前对于职业教育学生学习质量的评价，大多数崇尚结果导向的评价模式，且主要参照普通教育的评价指标，并局限于对学生习得陈述性知识的最终表现进行评价，不仅忽视对学生的技能、情感、态度和价值观的测评，更忽视对学生德、智、体、美、劳等各方面的系统性考量，导致学习过程和学习成果常常出现二元对立的局面。

2. 评价内容学科化，忽视教育类型属性。职业教育的学习过程以基于职业行动逻辑和知行合一的学习机制为典型特征，强调职业性、实践性和应用性，其内容的构成核心理应以程序性知识和策略性知识为主。然而，当前职业教育学习评价仍主要采用传统的纸笔测验方法，并侧重考核学生复述和记忆学科逻辑知识的多寡，而不是考核学生的技能迁移水平。即使在实践操作考试环节，也由于考试内容与职业领域真实工作过程的职业要求相去甚远，不利于培养学生解决实际问题的能力。可见，当前职业教育

学生学习质量评价以理论知识为考核重点，强调学科知识的完整性，主要关注学生的认知发展水平而非专业知识的系统性，严重制约了职业技能的结构化水平和职业技能的迁移水平。

3. 评价方法单一化，制约职业精神培养。长期以来，我国职业教育受“知识本位”“教师中心”等思想的影响，评判学生学习质量的方法非常单一。从评价主体而言，主要依靠专任的授课教师；从评价手段而言，主要采取纸笔测验方法，依据统一的评分标准，以卷面考试成绩的高低来评判学生学习成效。这种评价方法之所以目前依然盛行，原因在于操作简单，评价成本极低。然而，这样测评出的学生成绩并不能有力证明学生在学习过程中的变化，尤其不能反映学生的学习成效与解决生产、生活实际问题的能力之间的关系，更不能培养学生在工作过程中应展现的审辩、谈判、协商、竞争、妥协、民主决策、互相监督和独立思考等职业精神。

4. 评价结果静态化，难以提升行动绩效。当前往往以一系列分数和各项指标得分等量化（或质性）数据的方式体现职业教育学生学习的最终成果，难以真实反映学生的职业行动绩效。一方面，由于评价的题目往往是学校教师自编的测验题目，难以反映职业工作过程，其评定职业行动绩效的信度和效度值得反思；另一方面，评价的内容难以凸显职业行动的难易程度，要么“仅仅顾及知识的同化和顺应规律，要么梯度太大，超出了学生的最近发展区，导致学习者产生严重的习得性无助情况，丧失了学习信

心”。在这样的评价导向下，学生所学的知识越来越偏离生产实践，操作技能训练不足，迁移能力更是不尽人意，难以满足职业岗位的用工需求。

（节选自《教育与职业》，2021 年第 8 期）

职业教育学业评价方法刍议

赵志群³

摘 要：职业教育要满足企业需求和促进学生发展的双重目的，其学业评价更为复杂。当前职业教育学业评价存在课程评价脱离实践、技能考核缺乏规范、职教高考科学性不足等问题，社会发展拓宽了技能的内涵，具有职教特色的学业评价呈现出技能考试正向职业能力测评方向发展，评价方法多元化成为趋势等特征，未来更多会采用事实性标准与个性化标准相结合，量化评价方法占主导地位，质的方法成为重要补充。

关键词：1+X 证书制度；学业评价；技能考试；职业能力测评；评价标准；职业教育

一、学业评价方法：一个被忽视的重要领域

科学的评价对职业教育的健康发展具有重要意义。中共中央、国务院 2020 年印发的《深化新时代教育评价改革总体方案》（以下简称《改革方案》），提出建立富有时代特征、彰显中国特色、体现世界水平的教育评价体系的奋斗目标。针对学业评价，《改革方案》提出完善“学生学业要求”“过程性考核与结果性考核

³ 赵志群.北京师范大学,博士生导师,主要研究方向为职业教育课程与教学、质量保障体系建设与能力测评.

有机结合的学业考评制度”以及“高等职业教育‘文化素质+职业技能’考试招生办法”等多方面的任务，这对职业院校学业评价工作提出了更高的要求。

学业评价是指以教育教学目标为依据，运用恰当、有效的工具，系统收集学生在学习过程中认知、行为变化的信息和证据，根据一定标准对学生的知识、技能和能力水平做出客观衡量和科学判定的过程。合理的学业评价要体现党和国家的教育方针，尊重和体现学生的个体差异，激发学生努力实现自身的价值。普通教育学业评价相对简单，通常根据课程标准规定的学习目标，采用相应测量工具进行，如标准参照性测验和成就测验等。职业教育既要符合劳动力市场的功利性要求，又要实现促进学生生涯发展的教育性目的，其学业评价更为复杂和多样，除了课程考试外，还有技能竞（大）赛、对口升学考试和职业能力测评等多种形式。与普通教育相比，职业教育学业评价结果的影响也更大，它直接影响着学生的升学和就业、影响着企业员工招聘以及职业院校的发展。目前职业教育的学业评价在方法层面尚存不少问题，主要表现在：

（一）传统学历教育的课程评价因脱离实践而受诟病

院校内的科目考试缺乏可比性，一些地区通过专业测评建立教学质量监控系统，也遇到方法的科学性和经费可及性等困难，即便是发达地区也是如此。如，广州教育局曾在所属职业学校针对开设最多的 10 门专业基础课和 2 门专业核心技能课进行监测，

监测科目也仅覆盖全部课程的 18%，但仍难以全面反映学生的综合学业成绩，后未能坚持下来。

（二）X 技能等级证书缺乏规范的技能考核方法

对已发布的 X 职业技能等级标准进行文本分析发现，目前社会评价组织对技能评价方法的科学性尚不敏感，采用的评价方法有些强调操作技能考试（如特殊焊接技术），有些只提知识考核（如失智老年人照护），有些尽管关注理论和实践相结合但依然采用笔试形式，实操考试也有笔试、网络考试等多种方式，一些证书甚至没有提出明确的考核评价方法。

（三）职教高考科学性不足，“指挥棒”作用有待优化

在各地开展的职教高考中，技能考试多针对特定的操作活动，关注技能的熟练度和可测量性，对该项技能对整体职业能力的甄别和促进作用以及未来工作中的普适性则少有系统考量，且标准含糊，打分随意；专业课考试命题缺乏行业和考试专家参与，难以保证考试效度和区分度。由于缺乏有效的技能测试，很多职教高考的内容主体仍然是文化课和专业理论，这既无法反映行业企业的用人需求，也背离了职业教育宗旨，造成职业院校对专业技能教学的不重视。

（四）技能大赛也存在评价效度问题

调查发现，有的技能大赛获奖毕业生职业发展后劲不足，其综合素质与没有参加竞赛的其他同学相比甚至处于劣势。这说明，如果评价方式有缺陷，技能大赛也无法实现促进人才培养质量提

高的初衷。个别省份开展的技能抽查的科学性问题就更多了。

鉴于职业教育学业评价的难度，很多地区把文化课统考作为职业院校（特别是中职）学业评价的主要方式，这势必造成职业教育“去职业化”趋向，在此美国提供了前车之鉴。2002年，布什政府推出《不让一个孩子掉队法案》，试图解决美国中小学教育质量低下的问题。该法案的实施，提高了学生尤其是处境不利学生的阅读和数学成绩，但是大大削弱了职业课程，很多学校放弃了非考试科目课程，其中就包括职教课程，导致美国中等职业教育事实上的“消失”。目前，强调文化课统考的《不让一个孩子掉队法案》已经被《每一个学生成功法案》取代，提出在基础教育阶段重新发展多样化的课程（包括生涯与技术教育），这为美国重振中等职业教育提供了可能。

综上所述可以发现，在职业教育质量保障体系建设过程中，开展科学、规范且实施便捷的学业评价的难度长期被低估，这也凸显了《改革方案》的及时性和重要性。职业教育学业评价必须反映技术技能人才的发展性目标和职业规范的双重要求，这对评价方法提出了巨大挑战。本文探讨具有职业教育特色的学业评价的方法性特征，以期引发进一步讨论并为具体评价方案设计提供理论基础。

二、职业教育的经典学业评价模型：柯氏评价

职业教育是赋予学习者从事某种职业性工作所需能力的教育，其学业评价必须考虑教学内容与职业活动之间的紧密联系。

学习成功与否，最终只能在真实的职业工作实践中衡量，需要结合实际工作绩效表现。按照判据、表达方式等评价要素的要求，职业教育学业评价的对象不仅包括学习收获，也涉及学习者在企业工作中的绩效水平，在此，经典的“柯氏评价理论”具有重要的意义。

1959 年，柯克帕特里克（D.L. Kirkpatrick）围绕培训过程中的反应、学习、行为和结果发表了 4 篇文章，确定了职业教育学业评价的四个基本层次。一级评价为学员反应，即课程结束时进行学生满意度调查，对所学知识和技能在将来工作中的有用性进行展望，但这尚未涉及真正的学习结果。二级评价针对学习效果，通过学习过程记录、学习结果测试等方式确定学习者学习结束后在知识、技能和工作经验等方面提高的程度，对课程目标的达成情况进行确认。三级评价在学习结束一段时间后进行，确定学习者行为转变的程度。这些数据的获得比较困难，需要一定外部条件，如真实的工作情境和工作条件等；评价方法也更为复杂，常常需要正式测评和非正式方式（如现场观察）相结合。四级评价针对绩效实现情况，即学习者因参加教育培训后实现的最终结果，强调因学习而产生的绩效提高，如生产服务质量改善和效率的提高，包括精神方面的影响，常见的如对口就业率调查和毕业生发展跟踪调查。四级评价之间的关系如图 1 所示。

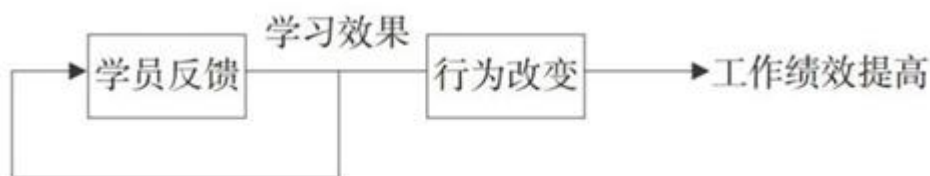


图1 柯氏评价4个级别的关系

三、四级评价需要在企业和劳动市场中了解职业教育对学习者的促进作用，经费和时间花费较大，操作起来也很困难，有时甚至具有科研性质，但是由于体现了职业教育的最高宗旨，因此具备更强的说服力。孙钰林曾以养老服务专业为例，对传统实操考试对学习者的工作绩效的预测效度进行研究。结果发现，职业院校普遍采用的实操考试的成绩（二级评价）并不能预测未来的工作绩效水平（四级评价），其相关系数仅为 0.01。这说明，四级评价中的每一级评价都有必要。从方法上讲，评价结果只有在下一级评价中得到验证才有实际意义，评价方法必须进行科学设计，以确保评价的质量，即所谓信度、效度、区分度和客观性等要求。但是柯氏评价缺乏对教育规律的考虑，不涉及学习者的能力发展和素质提高过程，因此无法独立承担起职业院校学生学业评价的重任。

三、学业评价方式：职业技能考试和职业能力测评

职业院校学生学业评价离不开对学生职业能力或技能水平的评价，其基本方式是技能考试和职业能力测评。

（一）技能考试

不同国家技能考试制度不同，但大多都会采用“知识+实践技能”考核方式。我国的职业技能等级（资格）鉴定考试一般也分为理论和实践考核两部分，采用纸笔测试、现场考核、典型作业、模拟操作等形式。技能主要采用专家观察评分法确定；对职业规范和工作道德等隐性范畴的内容，则通过对工作过程和工作结果的观察进行考评。国外研究发现，目前技能评价中普遍采用的标准化考试的职业效度并不高，因为企业真实的生产服务活动不是简单的操作或线性过程，不能简单按照“对—错”标准或“投入—产出”关系来衡量。

技能大赛是一种社会关注度高和高利害性的、特殊的技能考试，不同大赛在赛项设置、评分标准等方面有所不同。职业院校技能大赛常围绕教学标准及生产技术设置赛项，强调工艺难度，这与世界技能大赛有较大区别；世界技能大赛被认为对专业技术水准和职业素质能够进行更为全面检验，实施成本高昂，但是如果其试题转化应用得好，可以为职业教育的高质量发展提供有效的借鉴。

当今社会，技术发展对传统的技能考试产生了重要影响。人们发现，除了特别简单动手操作行为外，采用主观性较强的观察法对复杂技能评价的结果的信度很低，因此发达国家技能考试出现的一个共同趋势是采用“表现性评价”，即考生“在特定情境中运用获得的知识完成某项任务，以考查其知识技能掌握程度，

或问题解决、交流合作和批判性思考等多种复杂能力的发展状况”。此外，技能的含义也得到了很大拓展，人们更加强调应用知识完成任务和解决问题的能力。《国家职业教育改革实施方案》（以下简称“职教 20 条”）明确指出职业技能等级证书要“反映职业活动和个人职业生涯发展所需要的综合能力”，这说明在“1+X”技能证书制度设计中，“技能”和“能力”已经几乎成为同义词，意味着职业技能正在发展成为更强调认知和综合性的职业能力。此外，传统的技能考试费时费力，特别是在开展大规模的技能测试时，要想达到较高的信度和效度要求，测试费用和人员投入会高到令人无法接受的地步。因此，很多（大规模）职业技能考试需要向成本较低的职业能力测评方向发展。

（二）职业能力测评

职业能力测评旨在评价应用知识技能解决实际工作中的问题的能力，它与职业技能考试的区别在于：考试针对教学（或职业）标准，考察对学习内容的掌握程度，一般有对错之分，反映教学标准的效度；能力测评评价与特定工作相关的认知特征和认知水平，考察实际工作能力，只有能力水平高低之分而没有标准答案，反映职业要求的效度。

职业能力测评的基本功能是将能力发展的实际水平及时反馈给教师和学生，帮助其处理学习过程中发生的问题并进行过程优化，是鼓励先进、鞭策后进和调动学习积极性的手段。当今社会对技术技能人才的全面发展和综合性要求越来越高，技能工作者

必须有能力（参与）设计工作计划，按照经济性、创新性和环境可持续发展等综合性要求对工作的过程和成果进行评价。科学的能力测评要用整体化观念看待职业能力，不但针对工作任务的完成情况，而且还关心学生是否能够对工作过程和工作成果进行反思和改进，应加强对诸如技术敏感性和创新性等高层次实践能力的评价，促进学生的“价值理性”和“事实性评价能力”的发展。

从宏观管理角度看，如果对职业院校学生进行的能力测评能够达到大规模质量监测的信度和效度要求，就可以藉此进行校际甚至区域间的比较，并对不同学校人才培养过程的优劣进行分析和解释，从而为政府的科学决策提供实证基础。在此，需要研究和开发基于先进学习理论（如建构主义和情境学习理论等）的、技能考试与职业能力测评相互补充的、反映综合能力评价要求的、具有高预测效度的评价模型和评价实施方案，并为职业院校和技能评价组织设计和实施高质量的评价提供实操性的工具和实践经验。

四、学业评价标准

个体的学业成就是学习过程的产出或结果，学业评价旨在对学习者的学业成就水平的高低（学习绩效）进行判定和分级，处于某一级别的人必须满足相关标准的质量要求。标准是判断质量优劣和满意程度的门槛性要求，可以是一种标杆、价值取向或一项最佳成就记录，如精确性、准确性或与某一样板（人或事）的相符程度。在教学实践中我们看到，很多教师往往寄希望于制定

绝对“科学合理”的标准，这个目标可能永远也无法实现。莱茵伯格（F. Rheinberg）将学业成就评价标准分为三类，即事实性标准、社会性标准和个性化标准，在此基础上可以界定具体评价指标，其中只有事实性标准属于客观标准。

（一）事实性标准

事实性标准是教学标准（或职业标准）中规定和描述的能力要求，藉此可以对学习结果进行客观评价。事实性标准的优点是：评价结果与教学标准要求的关系清晰可见，可以明确判断是否达到了最低要求，从而权威性地证明达到了某一特定要求（如技能等级）。事实性标准可帮助学生进行目标导向和自我管理式学习，但是无法让老师了解学习难度，对学生的进步也不敏感。另外，工作世界经常很难严格和客观定义什么是优秀、良好或者及格（如很多服务和技术过程无法精确评分）。如果学校对学生的成绩分布有要求（正态分布）时，教师还必须根据社会性标准对考试分数进行修订，这不但造成结构性困难，而且容易引发矛盾，并为教学设计带来困难。

（二）社会性标准

社会性标准是与他人对比时学生表现出来的绩效差异，即学生在一个团体中的相对水平。社会性标准的最大优点是能够识别表现相对较好和较弱的学生，确定能力差异，为开展个性化教学提供基础。其局限性是：可比性只是在特定的团体内有效（班级内部标准）。由于看不到全体成员的共同变化，所确定的成绩有

可能掩盖全体成员的共同努力或相同的退步，掩盖事实性标准的实现程度，以至失去标准本来的理性和可信度。社会性标准无法精确确定考生的分数，这有可能对学习动机造成消极的影响。

（三）个性化标准

个性化标准是与自己的过去进行比较，由此追踪学生的学习和成长过程。个性化标准的优点是：无论是好的还是差的学生都能确定自己的学习状态，自身努力与学习成就间的关系清晰可见，不同阶段的绩效波动可以可视化，这有利于提高学习者的学习动机。采用个性化标准，后进生不会受到优秀同学的打击，而优秀学生也可以与“过去的自己”相比较。但是个性化标准弱化了学生之间确实存在的差异，使学生失去客观评价自己的信息源，不适合评价外部有特定要求的领域（如入学考试和职业资格考试）。

综上所述，每种评价标准都有自己的优势和“盲点”，不可能满足学业评价的所有要求，教师在实践中需要利用多种标准并相互补充，学生也应学会利用各种标准进行自我评价。理想的状态是，个性化标准发挥主导作用，但又不忽略其他标准提供的信息；事实性标准与个性化标准相结合，是未来职业教育学业评价发展的趋势。

五、学业评价的方法

根据不同的背景、需求和目标，学业评价会采用不同的方案，这涉及不同设计理念、数据收集、分析和结果解释的方法。学业评价有多种分类方法，本文不进行分类学研究，只讨论若干重要

问题。

（一）量化测评

19 世纪中叶人们开始进行大规模的学业成就测验，其标志性事件是赖斯（J. Rice）对 33000 名学生进行的拼写能力调查。20 世纪 30 年代，“评价之父”泰勒（R.W. Tyler）采用实验设计模式对教学改革成果进行比较分析，量化评价方法成为学校教育最重要的学业评价方法。职业教育中采用的量化评价有能力倾向测试、学习成就测试、人格测试和能力测评等多种方法。其中，能力倾向测试对个人的潜在能力进行评价，预测个体能力未来发展倾向；学习成就测试对个体的学业成就进行测试，典型的如考夫曼教育成就测验-II（KTEA-II）；人格测试通过特定测试量表（如明尼苏达多相人格调查表、艾森克人格问卷等）对个体的心理特征进行测试，如兴趣、动机、态度、气质和性格等；能力测评方法很多，既有单项能力测评方法，也有综合能力测评方法，如 BCI 方法通过性格问卷对个人能力、自我组织和自我控制能力进行诊断评价，基于等级描述分类的混合式能力评价方法 SOLO（Structure of the Observed Learning Outcome）通过解答问题的表现特征对人的思维发展进行可视化分析。

量化方法可以避免考官个人主观因素对评价结果的影响，对促成目标达成有较好效果，因此在学业评价中发挥着重要作用。但是科学的量化评价模型很难建立，这极大地影响了评价结果的说服力，而且这些方法在过程性评价中的使用效果往往也不理想，

采用互动性更强的质性方法更容易深入理解和解释被观察的现象。

（二）质性评价

人力资源管理最先开始采用基于胜任特征模型的质性评价方法，如“工作胜任力测评法”（Job Assessment Method, JACM）和“行为锚定修正法”（Modified Bars Method）等。这些方法建有能力模型，但是没有量化工具和流程，致使操作灵活性很大甚至过大。德国开发实施了数十种职业能力测评方法，其中有很多是质性方法，如行动导向测评方法、单项技能评价方法和能力总结与汇总方法（Competence balance）等。我国一些职业院校采用类似方法进行形成性评价，如北京市商业学校以职业素养学分和证明、证章评价方式，记录学生在思想品德、技术技能、文化艺术、体育科技等方面取得的成绩，在此基础上给学生颁发“职业素养护照”。苹果公司 A+ 雏鹰计划在其试点院校校企合作学徒培养中建立“学习档案”（Portfolios），有意识地搜集证据，记录和反省自己在某一领域的进步，用可靠的证据记录学生取得的进步并获得成绩。

质性方法的主要缺陷是结果的可比性较差。由于对学生的行为描述、分类、选定能力指标和确定界别的难度大，此类方法对评价人员的专业要求也很高。此外，如果被评价者过于谦虚、夸张或不诚实，评价结果的可信度也会有问题。

（三）混合式评价

职业能力是个体从事一个职业时的心理认知特性,波兰尼“隐性知识”理论被广泛接受,间接否定了按照行为主义理念采用纯量化方法进行能力测量的合理性,量化与质性相结合的混合式方法成为发展方向,国际有影响的职业能力评价方法均据此建立技能模型和测评模型。

21 世纪初开始的 PISA 等大规模教育测评项目引发了全球对职业教育质量评价方式的讨论。讨论取得的一个共识是职业教育质量控制应实现从“投入导向”向“结果导向”的范式转变,而作为学业成就标志的职业能力是职业教育最重要的“结果”。

COMET (Competence Measurement) 职业能力测评是世界上第一个真正实施的大规模职业能力诊断工具,它基于“从初学者到专家”的能力发展逻辑,建立包含“级别”“内容”“行动”三个维度的跨职业能力模型,采用开放性测试题目,确认被测学生的职业能力发展级别,即名义能力、功能性能力、过程性能力和设计能力。教育部督导局在开发《高等职业教育专业评价试点方案》时采用了 COMET 方案,对三个代表性省份物流管理等四个专业的高职学生试验性的能力测评。结果显示,COMET 能力测评可以较好确认不同院校学生的职业能力发展水平,其开放性测试题目不但被企业认可,也广受学生欢迎。

COMET 能力测评不仅适合宏观层面和中观层面的学生学业评价,对不同院校的教学质量进行比较,也适合微观层面的教学评价,对教师的教学设计提供指导。COMET 还被工业与信息化部

以及世界技能大赛和“一带一路”技能大赛等不同的技能评价工作采纳。在设计能力测评实施方案需要特别考虑的是，哪些评价手段既能获得大量深入和有效的数据，而不会产生过多的经费和时间投入。在此需要解决的关键问题，一是关于产出和投入的因果关系分析，二是具体评价工具的开发，包括测试题目类型的研究等。

（四）诊断性评价

“诊断”的概念起源于医学，本义是通过特定的工具检查病人的病症，并有针对性地开具处方。布鲁姆（B. Bloom）最先在教育领域提出“诊断性评价”概念，指在学习开始之前对学习者现有水平进行评价，强调预防性的“事前行为”。诊断性评价是内部评价方法，其最大特点是采用“协商模式”，即相关人员共同协商裁定和解决相关问题，从而提高评价结果使用的可能性。诊断不是简单的现状描述，而是通过获得足够深入的信息，揭示学习中存在的问题及其成因，并通过制度化的措施形成持续改进的机制。

进入 21 世纪，诊断成为引领学校建构反思意识、发现问题、提高效能、促进自主发展的新型教育技术，特别是教育部通过促进建立职业院校“教学诊改”机制的行政措施，使这一思想得到了普及。值得注意的是，诊断不仅仅是收集数据，更重要的是使利益相关者都参与到诊断过程中，共同发现问题、寻找解决策略并实施改进。学业成就的诊断性评价是采用量化或质性工具，根

据职业教育利益相关者的共同诉求（“会诊”）对学生在学习取得的成果和存在的问题（“病情”）进行多维度分析，确认和总结原因（确诊“病因”），并提出解决策略（“开处方”），从而提高学习者的学习能力，以取得更大成功（“健康”）。

六、总结与展望

鉴于职业教育学业评价的复杂性，简单一种或几种评价方法都不可能完全满足所有评价目的的要求，评价方法的多元化成为必然的发展趋势。纵观人类历史，教育评价理念和方法论发展经历了“科学主义”“人文主义”和“后现代主义”三个阶段，这对职业教育学业评价方法发展也有重要的启发。

19 世纪科学技术发展取得了辉煌的成就，其强调证据、关注逻辑和量化设计的认知范式影响着评价方法的发展。基于决定论思想的“科学主义”教育评价通过“科学”手段和量化方法收集信息，形成对评价对象的认识并由此了解事物的“本来面貌”，此类评价强调量化指标和统一标准，注重外部评价，但是对情感态度和创造力等主观因素的评价有很大局限。

20 世纪中叶，秉承多元化观点的人文主义者发展了质性评价方法。据此，评价不局限于针对纯粹的客观事实，也关注意义、价值和情感等主观因素，管理者的授权、团体凝聚力和合作思想同样重要。斯塔弗而比姆（D. Stufflebeam）建立的 CIPP 评价模式就是这种思想的集中反映，即关注评价的背景（Context）、投入（Input）、过程（Process）和产出（Product），考虑主观经验在

评价中的作用，强调描述性的定性分析。

进入 21 世纪，后现代主义思想注意到科学主义和人文主义评价观的优势和不足，开始弱化量化分析而重视知识建构、应用条件以及由此带来的差异，由此提高了质性方法的价值，典型代表是第四代评价、参与式评价和授权评价理论。第四代评价关注社会的多元化现实和多元利益主体的需求；参与式评价立足行动研究，强调促进公民参与评价；授权评价则进一步发展了参与式评价，为利益相关者提供操作性工具，通过讨论过程的量化和可视化实现质性与量化评价的融合，即混合式评价。按照库克（T.D. Cook）的后实证主义多元化评价理论，高质量的评价应提倡利益相关者的多元参与、对相互矛盾的数据进行分析和解释，强调动态建模，并关注不同情境和时间段的评价结论的可推广性。

在现代职业教育中，学生学业成就的获得是个性化过程，学业评价要为学生提供差异化和多样化的启发，这需要改变传统的认知导向、仅强调产出的评价方式。按照后现代主义评价理念和现代学习理论，学业成就在很大程度上是“人为的设计”的，并没有一定要遵循的固定标准和特定答案，而更多要通过与参与者的协商和对话来确定。职业教育学生的学业成就有多种表现形式，可通过工作过程和产品展示，通过再现、重组、迁移和解决问题展现，也可表现为创造性、社会性、认知性、生产性和行动导向的工作绩效。学业评价过程应当嵌入在工作和学习过程中，这既有对个体的评价，也有对集体的评价，需要参与者相互信任和团

结合作。学业成就具有价值取向，学业评价需要参与者的理解和反思；学业评价还要确定学生的发展潜力和需要解决的问题，这需要通过机构化和系统化的支持，帮助学生解决自己所特有的问题。

目前我国职业教育学业评价的基础仍然是共享价值观，价值多元化还没有受到足够的重视。随着现代职业教育和终身学习体系的建立，以及受教育者的多样化（如高职扩招）发展，未来应当承认和尊重不同的取向和要求，职业教育学业评价也会出现从外部评价转向内部诊断，从一元控制走向多元治理。尽管量化评价方法仍然会占据主导地位，但质的方法会成为重要的补充。

总之，应关注传统考试存在的理念和技术问题，如果缺乏现代教育理念和教育测量理论的指导，不关心测量的技术指标，学业评价就无法支撑高质量职业教育体系的发展。我们需要加强相关研究和实践探索，如关注综合职业能力发展，这意味着不仅要评价动手操作技能，而且还需关注认知和精神运动领域的内容。按照情境学习理论，专业实践能力只能在特定的工作情境中被调查和获知，需要采用案例分析或民族志方法；指导实践的知识只能在特定情境下才能被辨识，对专业能力的评价需要在真实的工作情境中、以完成典型工作任务的考核方式进行，高技能人才评价需要更为复杂的考核方法。这里最大的困难是开发情境性考试题目，即有助于确立工作能力的标准化案例性任务，它既要反映重要的企业实践价值，又要能藉此评判被试者职业能力的高低。

例如，广州市将世界技能大赛标准转化为技能人才标准，通过产教融合方式开发高新技术领域的新赛项，不但满足了高技能人才培养和评价要求的“实用性目标”，而且有效促进了学生创新能力的发展，践行了职业教育立德树人，培养具有职业素养和领域特长的、德智体美劳全面发展的社会主义接班人和劳动者的要求。

（选自《中国职业技术教育》，2021 年第 10 期）