

“双减”政策驱动下的教育评价改革探究

马瑞 冀小婷

摘要:“双减”政策自实施以来成效显著,但还有继续完善的空间,如教师的作业设计能力弱、评价方式不当、教学质量亟待提高等。所以,应建立由大中小学教师和高校师范生组成的“教育-实践共同体”,借助“实践—认知—再实践”的学习路径提升教育实践能力。同时,要以形成性评价为支点,建立形成性评价标准和反馈机制,提高教学质量,构建良好的教育生态。

关键词:“双减”政策;教育生态;教育评价;形成性课堂评价

中图分类号: G40-011

文献标识码: A

2021年7月24日,中共中央办公厅、国务院办公厅印发了《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》(以下简称“双减”)。该文件明确提出,要构建良好的教育生态,促进学生的全面发展与健康成长。具体工作目标是减轻学生过重的作业负担和校外培训负担,同时减轻家庭教育支出和家长的精力负担。随即,各省市相继出台了配套实施方案,并雷厉风行地进行了整顿执行。本文针对该政策执行以来的情况作一探讨。

一、“双减”政策实施的阶段性效果与问题反思

首先,对培训机构实施分类管理;其次,对学科类课外辅导明令禁止、严格监管;再次,对非学科类以提升素质为目的的培训采取鼓励和规范发展的措施;最后,严控课业负担,定期检查作业总量,核查学校是否按时放学、课后作业时间是否控制在规定范围之内。可以看出,其监管力度是有史以来最严厉的一次。

我们必须认识到,“减”的目的并不仅仅是表象上的“减”,减是为了“增”——提高作业管理水平,提高课后服务水平,提高课堂教学水平。只有做到这些,减负才能从量的控制走向质的提升,从治标走向治本。“双减”政策实施几个月来,各方积极行动,成绩斐然。仅从学校层面看,除了作业公示与刚性监控外,还增加了课后服务、成立兴趣小组等项目。部分试点地区在典型推广、平台服务、评价体系等方面共同发力,做了许多有启发意义的尝试。然而,我们也看出,还有许多亟待解决的问题。

(一)学校师资不足,是落实减负措施的重要阻碍

学校处于“严格执行行政命令”和“操作无良策”的夹缝里,正课后的时间变成了钉自习。

收稿日期: 2022-01-15

基金项目: 2021年度天津市教委科研计划专项任务一般项目(外语教育)“双减背景下英语师范生教育实践课程评价改革路径与效果研究”(52WR210026);天津市大中小学“课程思政”研究专项课题“专业视域下课程思政一体化设计模式及实现路径研究”(GX-005)。

作者简介: 马瑞,天津师范大学(天津 300387)教育学部讲师;冀小婷,天津师范大学外国语学院教授。

这种方式除了解决双职工家庭无法提早接孩子的问题外,对提升学习质量意义不大。课后兴趣小组的活动时间和活动项目比较有限,素质教育的初衷难以实现。虽然写作业的时间有所限制,但可能会出现孩子回家后无事可做而玩手机、打游戏的现象。习惯了督促孩子完成作业的家长们,面临孩子回家无事可做的情况,出现了新的焦虑,担心学习成绩会下滑,于是想找课外辅导的“火苗”重新点燃,甚至有些家长已开始四处打探,寻找隐蔽或变相的培训点。可以看出,如何安排减负之后的课后育人活动,是当前在家长方面存在的大问题。

(二)仅靠中小学教师而提升作业设计水平的做法不现实

我们知道,一线教师任务多、负担重,有的教师教育理念和专业知识相对陈旧,无法通过少而精的作业设计明显提高学生的综合素养。作业减量容易,但提质难,这不是靠一纸行政命令在短期内就能彻底解决的。自“双减”政策实施以来,很多省市安排高校教师、科研机构的研究人员和相关专业教师与中小学合作,希望在师资培训方面发挥高校优势,助力“双减”措施的落实。合作的具体方式以讲座、听课和评课为主。但这种合作方式也存在质疑,因为高校教师往往止于教育知识的生产,其研究成果的形式主要是发表论文和出版专著,与教育实践存在明显的脱节现象。听完讲座后,中小学一线教师总有一种讲得很好、但无法应用的感觉。在听评课活动中,也会指出各种问题,但对如何解决这些问题却提不出多少有效的办法。究其原因,许多高校或科研机构的学者并不真正了解一线教师面临的问题与难处,就只好提一些大而空的想法。鉴于此,高校教师和相关研究者,应该更多地深入教育一线,了解和掌握教育实践中存在的现实问题,通过“学术-实践共同体”展现教育研究的社会责任和实践情怀。[1]

(三)教学质量与作业完成存在直接因果关系

作业是检验教学效果的重要方式,若教学方式陈旧无趣,教学质量不高,作业的效果自然不会很好。很多学校开展作业设计的改革创新,试图增加趣味性和多样性,但形式的变化大于内涵的发掘,作业并没有成为课堂的延展和所学的应用。有的学校把作业搞得很花哨,难度也比较大,但由于教学的铺垫不够,学生完成作业时感觉困难很大,由此催生了更大的课后辅导需求。课上的欠债由课下来还,本身就违背减负目的。所以,作业质量的提升还须从课堂教学上下功夫,两者不可分割。从某种意义上讲,课堂教学水平提高了,教学目标达成了,作业设计得适宜,学生愿意做、做得快,作业负担自然就减轻了。但要注意,负担轻并不意味着质量提高了,还须对作业进行必要的反馈,学生的进步是在不断的作业反馈中取得的。因此,要提高课堂教学水平,就需要掌握课堂形成性评价的理念和方法。

(四)评估标准不改变,就无法做到真正的减负

教学质量的好坏需要科学的评估标准,管理者需要,教师也需要。但传统的绩效问责式评估标准以分数为核心:教师业绩的考核以分值为标准,教师会把这种压力转化到学生身上,对学生的学业表现也唯分数论,以高分形成光环效应,分数高则全部好,分数低则全部差。这种思想在家长身上表现得更为明显,这是造成减负难的又一主要原因。因此,要想办法真正触动和改变家长对评价方式的陈旧认识,发挥家庭与学校的合力作用。

总结“双减”以来的成效和问题,我们看到,国家从行政层面给予有力的政策支持,尽量减少外部因素对提高学校教育质量的负面影响;同时,从外显层面减少学生在学校教育中的课业负担。但这离提高教育教学质量、促进学生全面发展的总目标还有很大距离。从政策的长期性来看,以往减负政策的落实并不理想。据统计,改革开放四十多年来,我国总计出台了35项中小学减负的政策文件,其中专项政策有11项,涉及减负的相关政策24项。[2]减负的执行情况基本上是始于轰轰烈烈,终于不了了之。究其根本原因,是教育质量不高所致。所以,要根治学生作业负担过重的问题,就必须从教学质量抓起,必须将价值塑造、知识传授与能力培养融为一体,构建新时代教育发展的新模式。这是短期内在其他方面的保障尚不能马上见效、而能

使“提质减负”目标落到实处的最直接、最重要的抓手。[3]

二、提升教学质量的方法与策略

从上文可以看出,在“双减”政策的实施过程中,存在作业设计水平不高、评价方式不当、质量监控缺失、人力资源不足、综合治理机制不完善等问题。为此,我们要对这些问题作系统反思,以更有效的手段真正做到减负提质,构建最佳教育生态。

(一)建立“教育-实践共同体”

高质量作业体系的建构是落实“双减”政策的关键,它涉及作业的设计、指导、评价及课后育人活动等,因此,必须有一支专业能力强、富于创新精神的师资队伍作保障。由于短期内解决师资短缺的问题并不现实,我们不妨从发掘现有资源的方法入手,利用高校资源,构建大中小学协同发展的教育“共同体”,其成员由高校专业教师、高校师范生和中小学一线教师组成。当然,积累一定经验后,还可以推广到退休教师、培训机构以及其他行业的爱心人士。

作为潜在的资源库,高校师范生可在高校教师教育专业的教师指导下进行专项训练,广泛、深入地参与中小学的教学实践。培训之前,专业教师要下沉到基础教育一线,和中小学教师一起教研,了解他们是如何进行作业设计和实施的,并要分析总结,找到其中存在的问题。在此基础上进行改进性设计,形成案例,在教研研讨会上通过事例分析帮助一线教师提高理论水平,组织他们进行新单元的作业创新设计。然后,通过独立设计、小组修改、分头实施来检验其可行性。其实践方式要以说课的形式呈现。年级组教师要集体参与,并按评分标准给出反馈意见,说课的结果纳入年终考评系统。

高校教师也应以类似的方式培训师范生,并在课后将作业的设计和实施方案作为项目式作业进行形成性评价,对提交的作业设计和实施情况按标准打分并予以反馈。把作业设计能力作为师范生见习、实习的内容,以作业设计为抓手,倒逼教学设计能力和评价能力的提升。通过这样的实训,学生既可完成本科师范类教育实践课程体系的学分要求,又能充分了解中小学生的教学状况,理论与实际达到有效结合。经过专业训练和实践的师范生,其首岗胜任力会得到大幅提高,与入职后的工作形成无缝对接;对中小学而言,不但部分解决了师资短缺的问题,还输入了新的教学理念,创新了教学方法。

“教育-实践共同体”的建立以需求为导向,在教学模式的探索中实现了基础教育与高校教师教育的双赢,其真正的意义不仅仅是减负本身的事,更是推动了教育生态圈的良好发展。所以,从理论层面讲,这种探究也为高校实践类课程的评价方法带来新课题、新挑战,是教育评价改革研究的重点课题。

(二)形成性评价体系的构建

形成性评价发生在教师、学生或同伴在搜集、解释和使用学生学习信息的过程中,教师可据此作出改善教学的决策。这种评价方式也被称为学习性评价,它与教学紧密相关,其促进学习的作用受到人们越来越多的重视。[4]比如,在《普通高中英语课程标准(2017年版,2020修订)》中,它要求作为课程评价标准的形成性评价“应贯穿教学的全过程,为检测教学目标服务,以发现学生学习中的问题,并提供及时帮助和反馈,促进学生更有效地开展学习”。因此,形成性评价对提升教学质量具有非常重要的作用。

1. 形成性评价概念的具体化

顾永琦和余国兴根据前人的研究成果总结了师生在形成性评价过程中的四步行为:(1)引导学生展示他们对所学内容的理解程度或学习成果;(2)根据学习目标和评价标准对任务表现进行解读;(3)基于上述解读向相关学生进行反馈;(4)师生共同采取有针对性的后续改进措

施。[5] 后来,顾永琦和李加义将其概括为直观的流程图(见图1)。

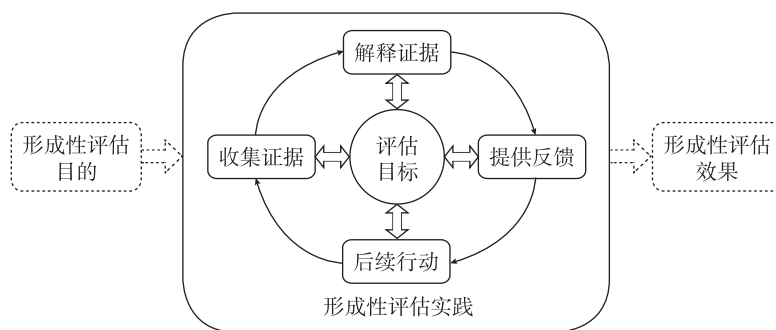


图1 形成性评价概念的具化

在此步骤中,有几个环节很关键。(1)任务的设置。任务是刺激表现行为的媒介,任务设计的质量直接影响成果的产出,所以,任务设计必须依据学习目标而行。(2)评价必须有标准。对学生表现的解读和反馈都须照此执行。学习者和教师都要熟知该标准,共同对照执行。在形成性评价中,反馈可根据量规(rubric)进行,主要是为了让学习者了解自己的学习情况。所以,反馈越详细,对改进的帮助就越大。反馈一般是基于评分标准给出等级分,每个分数档次都有相应的表现描述。(3)反馈之后必须有针对性的改进措施。只有及时、有效地进行改进,才能扬长补短,使所学知识走向内化。

2. 形成性评价的质量监控

要提高教学质量,就必须配备必要的质量监控方法,这样,才能验证其教学方法的合理性。巴克曼和丹博克提出的质量监控流程将课堂教学与评估机制联系起来,为教学质量评估提供了一份较详细的行动方案框架(见图2)。

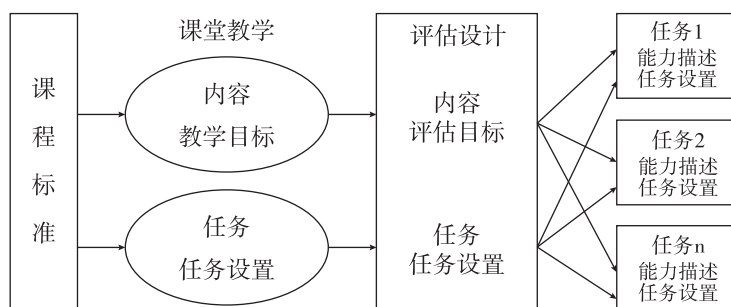


图2 教学质量监控流程

资料来源:根据 Bachman L.& Damböck B. Language Assessment for Classroom Teachers. P140.Oxford : Oxford University Press.2018 制作。

根据该框架,教学质量监控要遵循如下步骤。

第一步,以课程标准为依据,以课本内容为载体,以单元或者章节为单位,确立每一个单元的教学目标,据此设计相应的学习任务和对任务表现的评价方法以检验教学效果。

第二步,基于日常教学行为的教学质量评估,具体分为过程性评估和结果性评估。过程性评估是对照课程标准,核查教学目标和教学任务的合理性。然后对比表现性评价的证据和结果,考查评价结果的合理性。之后,再对比学生的表现和教学目标之间的一致性程度,从而判断教学效果的达成度。结果性评估是对课堂教学的抽样检查。评估内容的设定要结合教学目标,评估任务的设置要参考教学任务,即平常怎么学,评估就怎么做。

第三步,相对于教学环节,评估目标的设定可以是分解性的,也可以是综合性的,它通过具体的任务引导出来。比如,教学中可以将评估目标细化为各种子能力,然后逐项设计任务进行分解式评估,从不同角度搜集表现性证据以体现评估目标的达成度。而综合性的任务设计可以用来考查学生综合运用所学的能力。比如,可以融入课程思政元素,通过项目式作业考查学生运用所学解决实际生活问题的能力。目前,基础教育的作业方式多聚焦于分项能力,比较重视知识的落实,综合能力的评估在教学中体现较少。这正是“教育-实践共同体”合作攻关的重要结合点,它可通过综合性作业检验综合性评估目标的达成度。

上述流程以课程目标为基本依据,体现了从应然到实然的实践步骤。我国教育的根本任务是“立德树人”,课程标准是该宏观任务的具体落实。教育评价对应每个步骤中的内容与方法,它联结起课程标准的各个层级,期望达成内容的效度。这样层层分析,最后设计出的具体评估任务,就是“立德树人”教育目标的具体体现。将抽象、宏大的教育目标具象化,便于靶向性操作,可避免空喊口号、理念与现实脱节的现象发生。这个过程并非单向路径,教学中反映出的问题又可为课程标准的修订资以参考,从而形成不断改进的良性循环系统。

3. 形成性评价的效度验证

任何评价都要有效度验证以证明其合理性。在测试学中,学者经常用这样一个讽喻来说明测评效度低的危害:如果我们测试鱼会不会爬树,那鱼终其一生都只能相信自己是愚蠢的。这里恰如其分地说明了测评中任务设置合理性与测评效度之间的关系。形成性评价的效度验证有传统的效度验证法和基于论证的效度验证法。[4]传统的效度验证法既要确保评估内容的靶向性(curriculum fidelity),以达到教学评的内部一致,同时还要保证效度理论的关联度。

在图2中,从课程目标到课堂教学,再到评估内容,三个层级的目标与活动的相关性、代表性一目了然,这是关于内容效度的验证。理论关联度是对照教学设计的理论“构念(construct)”来检查评估各项活动的相关性,相当于对构念效度的验证。比如,作业目标写的是某特定情境下的交际能力,但这些任务都是以选择题的形式呈现的,其作业任务的理论关联度就不会高。再如,有的英语作业是以趣味背单词的方式呈现的,但其趣味性并没有体现出来,其实就是传统的单元背诵生词表,不过是给这个老作业换了件新“衣裳”而已。没有语境的单词记忆与目前新课标的要求相去甚远。所以,按照效度验证的方法分析作业,可以发现很多质量问题。

4. 形成性评价标准的设定

学习任务或作业设计都应基于真实的任务,这样才能锻炼学生真实情境下解决问题的能力。基于任务的评价常以表现性评价为典型特征,它具备以下六个特征:学生的作答是构建式和产出式,并非识别型作答;任务为开放型,能调动高阶思维能力;任务有意义,体现了真实情景下的参与式;需要各种技能的整合;更注重考查学习者知识和能力的深度。[6](P127)

一般表现性评价会使用“量规”作为评价工具。“量规”可以考查学生的学习过程、进展和最终学习结果。同时,“量规”可定义最佳的任务产出,教师可通过考查要素来体现其价值取向。目前,常见的“量规”只关注“表现”要素,以“表现”论优劣,过度重视学生的认知能力,这与全面发展的要求相悖。对“全面、健康发展”的衡量,不仅要看学生的认知能力,其“努力程度”也应纳入“量规”之中。“努力程度”的表现值可由全班同学共同讨论得出,如此,还能起到反思、对比和引领的作用。“努力程度”也可按等级打分。比如,期末的项目化作业,每个学生都是两个分——表现与态度方面的评定。它不强调相互间的攀比,只对照标准落实与反思即可。比如,有的学生虽然认知表现欠佳,但学习态度认真,也可得到高分;相反,很聪明的孩子如果学习不认真,则其“努力程度”项的得分会被扣除一些。“量规”评价改进了传统的只报告分数反馈的方法,其意义在于帮助学习者更多地了解自己的强项和弱项,明确努力方向,属于成长型反馈。

基于以上形成性评价的原理,我们可以设计各种“任务”用以评价。在对任务的内容效度

和构念效度进行核查的基础上,采集学习者的表现信息,然后参照“量规”对其表现进行评价与反馈。学习者自己或同伴同样可以按照该流程进行自评或互评。该方法可用于中小学教师的教学质量、作业设计质量、师范生作业设计的培训或实习试讲等方面的评估。

需要特别强调的是,形成性评价是基于标准的评价,它注重表现的描述和反馈,目的是为了改进和提高,不能用于以分数论高低的竞争场域。

三、结 语

自“双减”政策推行以来,我们已经取得了很大的成绩,但要使政策具有可持续性,必须各方面齐抓共管。但无论怎样,教学质量的提升应该放在首位,这是教育的生命线,是学校教育主体地位的体现。形成性评价体系和学习共同体的建立,可使大中小学教师和高校师范生建立起协同发力的新机制,以提高新时代学校教育质量。质量监控不是为了传统意义上的绩效问责,而是用于教师教学的阶段性测评,以此发现学生学习中存在的问题。师生配合改进教学,达成以评促教、以评促学的目标。总之,“教师唯有在学习和实践中不断提升自己的理论修养和教学实践水平,才能在教学中真正践行唯物史观,培养学生的唯物史观学科素养”[7]。

参考文献:

- [1] 石中英. 研究真问题,真研究问题! 教育研究要回归教育实践需求! [N]. 中国教育报, 2021-12-09.
- [2] 王毓珣,刘健. 改革开放四十年中小学减负政策变迁及走向分析[J]. 教育理论与实践, 2018(31).
- [3] 范涌峰. “后减负时代”基础教育高质量发展的生态重构[J]. 四川师范大学学报(社会科学版), 2021(6).
- [4] 顾永琦,李加义. 形成性评估的效度验证方法[J]. 外语教育研究前沿, 2020(4).
- [5] Gu P Y, Yu G. Researching Classroom-based Assessment for Formative Purposes [J]. *Chinese Journal of Applied Linguistics*, 2020(2).
- [6] 道格拉斯,阿贝维克拉玛. 语言测评:原理与课堂实践(第二版)[M]. 北京:清华大学出版社, 2012.
- [7] 张聪. 党史教育:思想政治课的重要使命[J]. 天津师范大学学报(基础教育版), 2021(2).

The Innovation of Educational Assessment Under the Context of “Double Reduction” Policy

MA Rui JI Xiaoting

Abstract: Since the “double reduction” policy was implemented, remarkable progress has been achieved, even though some problems still exist for further improvement in terms of homework design, assessment development, and teaching quality control. It is suggested that “Community of Educational Practice” consisting of college teachers, primary and secondary school teachers as well as pre-service teachers be established. All participants’ co-work makes it possible to improve education quality in the course of “practice—cognition—re-practice” with the aid of proper formative assessment. Criteria for formative assessment and system for feedback should be developed to improve teaching quality so as to build the good educational ecology.

Key Words: “double reduction” policy; educational ecology; educational assessment; formative assessment

【特约编辑:丁兰舒 责任编辑:申心刚 王韶芳】