



# 美国小学科学表现性评价案例研究的分析与启示

## ——以“物体的浮与沉”为例

○陈 乾

**摘 要** 表现性评价作为现今科学教育研究的热点,如何有效实施愈发受到关注。通过对美国纽约市州政府开展的四年级学生科学学业评估试题案例《Float and Sink》进行文本描述,并对案例的目标导向、任务设置及问题组织加以分析,提出表现性评价的实施要立足课程标准、构思设计要注重真实情境、表现过程要贴合学生的元认知活动等,以期对我国表现性评价实施有所启发。

**关键词** 科学学业评价 表现性评价 美国小学教学

随着国家新课程改革的不断推进,表现性评价的应用受到人们的广泛关注。美国针对新课程标准开发并构建了科学学业评定系统,而国内目前对于此类评价实施过程中所出现目标偏离课程标准、任务设定与现实情境脱节、问题组织难以有效监测学习成效等问题,研究并分析美国表现性评价案例,对我国中小学教师明确表现性评价的步骤、过程及意义具有一定的研究价值和启示意义。

### 一、表现性评价案例样态描述

本文的评价案例选自纽约市州政府对本市小学四年级学生进行的全州科学评估中的一道含有表现性任务的综合试题<sup>[1]</sup>(Mixed-Item Formats with Performance Tasks)。这种综合试题(除了包括选择题和简答题之外)还含有大量表现性任务类试题,可以通过观察、实验、设计操作、调查和角色扮演等任务类型来协助老师判断学生的探究技能、逻辑思维及语言表达能力等是否习得或是达到什么样的水平。

在“浮与沉”的案例中主要以观察实验作为表现型任务的核心,期望达到的评价目标旨在引导学生在科学知识概念的理解上能够知道不平衡的力会使物体运动的方向和速度发生改变;熟悉科学探究的一般过程及其应用,尤其关注学生提出和明确

问题、构建解释和设计解决方案等科学实践的能力表现;由于案例并未与真实生活情境相脱节,在情感态度和价值观方面可以培养学生真实的科学态度,如尊重事实和坚持证据,有预见性地应用科学知识等角度进行综合评价。

在设置表现性任务时,详细地阐明题目设定的背景和条件,明确地告诉学生需要做什么。实验中提供的主要器材分别是一个铁制圆形容器(船)、两个分别重 10 克和 4 克的圆孔垫片(货物)和一个蜡烛,以此模拟小船载物的情境进行实验探究。分别提出以下几个问题:

问题 1:如果将一个小垫片(货物)放在船内的边缘处,而不在中心位置。当你把船放进水里后会发生什么?在方框内画出你认为会发生什么,在横线上写出对此现象预测的解释。

问题 2:按照问题 1 的描述,将小垫片放置在船内边缘处,然后把船放在水中,观察发生了什么。在方框内画出你观察的景象,在横线上写出对该现象发生的解释。

问题 3:当船装载不同的垫片时,你还想知道它会发生什么其他现象吗?写下你的问题。

问题 4:研究你提出的问题,并设计一个实验来解决这个问题,把你发现的结论画下来或写下来。

问题 5:想想你能从刚刚做过的实验中学到什么。在下列表格的叙述中,如果可以从刚才的实验



中得到答案就在“已学习”一栏中标记;如果不能,就在“未学习”中标记。(分别在表格中标注√或×)

已学习	未学习	表述
		当货物放置在船的边缘时,船会发生翻转并下沉。
		船的行驶需要一个电机驱动。
		船自身越重,它越容易沉入水中。
		由金属制成的船装载铁制品后仍然可以漂浮。
		圆形的船比长条形的船更容易漂浮。

这项任务主要是为了评价学生在提出、明确问题和计划、实施调查中的科学实践表现(更具体地说,指仔细精确地观察并描述科学现象,严密周全地推理并探究实验结果),在“力与运动”的物质科学知识领域内,指向学生在真实环境中如何理解物体的上浮与下沉。测试时监考员大声地把实验要求读给学生,并演示如何将垫片放在船上,以及将船放入水中。整个任务中,对于前三个问题任务,学生在监考员的指示下基本可以顺利完成;而最后两个问题任务尤为值得关注:学生需要根据自己所提出的问题自主设计一个实验来探究并予以回答;学生通过上述一系列表现性任务自我反思判断可以学到的有效信息。无论是前者那种易于操作的任务类型,还是后者这种开放式问题或建构式问答都需要州政府组织安排相应的专家评委制订一套可用来衡量学生在完成任务过程中的表现及其成果的标准。不仅涵括对学生学习行为反应及其成果特征的指标描述,还应该具备相应的水平等级判断。以问题1(表1)和问题4(表2)的评价量表为例:

表1 《Float and Sink》表现型任务问题1的评价量表

记分:3
绘图/回答反映了以下想法:船是浮动的,但倾斜到一边。盘在船的 内侧边缘上的位置使船漂浮不均匀。
记分:2
绘图/回答反映了以下概念:船是浮动的,但倾斜到一边。没有理由 解释它为什么倾斜。
记分:1
绘图/回答表明船舶漂浮,但没有认识到,偏心放置的重量导致船舶 浮动不均匀。
记分:0
指示船舶沉没或其他答案/图纸的绘图/答案。

从上述案例的表现性评价量表中不难发现,评价目标并不仅仅聚焦于学生是否可以对物体漂浮或下沉原因的概念解释;也没有着重考查学生观察和描述现象的一般技能。而是通过问题情境将目标具体化,学生必须认识到当重物放置在船的非中心位置时,它们会使船以一定的倾斜角度产生不同的

表2 《Float and Sink》表现型任务问题4的评价量表

记分:2

答案需涵盖以下标准:

- (1)与所提的问题紧密相关且设计的实验可以为问题提供答案。
- (2)将自己的观察以绘图和文本的形式详细呈现出来(例如,船向左倾斜,负载下降,并迅速沉没)。

记分:1

答案需涵盖以下标准:

- (1)与所提的问题有点有关,其实验至少是针对提出的问题所设计的。
- (2)绘图和文本的表达内容是可以理解的,但不完整或不详细(例如,船倾斜)。

记分:0

与所提问题完全无关的其他答案。

浮沉现象,甚至希望学生可以由此发现重物在中心的偏移位置不同,对小船在水中的倾斜方式也造成不同程度的影响。本项任务的研究重点是引导学生对浮沉知识能加以整合并对关键信息进行观察和描述。评分标准是根据学生能否对科学现象的重要特征加以把握,以及如何将相互关联的特征有机地结合起来,需要学生仔细地观察现象,描述和比较;根据所提出的问题、疑惑选择并使用合适的工具、仪器和材料进行调查、分析和实验,学生最后再审视反思学习结果,以便达到与预期监控目标相适应的标准。

## 二、表现性评价案例特点分析

笔者结合本文案例根据美国课程标准的期盼要求、任务情境的真实程度以及对于学生反思活动的监测把控这几个方面加以分析。

### (一)评价目标聚焦课程标准

美国在2011年制定的《K-12科学教育框架:实践、跨学科概念、学科核心思想》<sup>[2]</sup>(以下简称《框架》)中明确指出三个关键词“科学和工程实践”“跨学科概念”“学科核心思想”,旨在构建相互整合的三维科学目标,关注学生在K-12不同学段中的学习情况。2013年又出台了《新一代科学教育课程标准》<sup>[3]</sup>(《Next Generation Science Standards》,以下简称“NGSS”)其中用表现期望(课程目标)来阐述学生的学业成就,是对《框架》的详尽表述和补充。

具体来看,该表现型评价案例涉及三年级物质科学领域的第二个核心概念的第一个子概念:力与运动。具体表述为:不平衡的力会对物体的运动方向或速度造成影响,通过观察物体的和受力情况和运动状态可以预测其未来的运动趋势<sup>[4]</sup>。NGSS中在这块关于科学和工程实践部分,要求学生能够区分



科学的(可检测的)和非科学的(不可检测的)问题。根据若干变量之间的因果关系及先验知识预测合理的现象及实验结果,来描述力是改变物体运动状态的原因。3至5年级的实验设计可以建立在学前阶段到2年级的经验基础上,首先确定一个简单的实验问题,根据实验对象、操作工具、系统开发,包括材料、时间或成本等若干条件的限制,进行观察、测量以生成实验数据并以此作为解释现象或设计实验方案的证据基础,运用科学的思想方法测试两个不同物理模型对同一个实验对象的影响,比较一个问题的多个解决方案,构建所观察的不同变量间的关系解释在解释中找出支持自身观点的证据进行科学辩护。本案例在跨学科概念中的关注重点是对于物体受力模式与运动规律的识别,利用生活中常见材料模拟小车载物的基本物理模型,以探究设计的方式不断测试、解释物体的沉浮与哪些因素有关,帮助学生进一步理解平衡力与非平衡力如何影响系统的稳定与变化。

表3 《Float and Sink》表现性评价案例在  
NGSS 三个维度表现期望的落实

学科核心思想	跨学科共同概念	科学与工程实践
PS2.A 力与运动 指导学生知道不平衡的力会对物体的运动方向或速度造成影响,通过观察物体的和受力情况和运动状态可以预测其未来的运动趋势。	通过科学教育引导把握事物运动的模式和规律,认识到力与运动之间的因果关系,不断提高研究和论证层次;理解系统如何在力的作用下从稳定均衡状态转向不平衡,又怎样达到新的平衡状态。	学会提问题并明确自己到底解决什么问题,并较为准确地解释物体浮与沉的科学现象,设计相应的解决方案进行调查研究,运用证据阐明、辩护自己的观点,完成工程设计和科学探究。

## (二)任务设置面向真实情境

评价案例的标题为 Float and Sink,提及这一词语的时候,儿童脑海中往往会浮现出幼时放在水中的玩具沉浮画面,为给学生营造一个充满童趣的学习环境,也考虑到其已有知识经验的限制,设计者通过给出几样学生熟悉的实验器材,将其作为一个切入口来帮助学生模拟课程情境,通过科学探究的形式,先利用经验预测小船可能发生的情况,再让学生利用材料动手操作去验证实验模型是否与自己的假设相同,在完成整个表现性任务的过程中,学生需要逐步理解科学探究的一般过程,实验设计的原则和设计评价等知识。尤其是问题4中,将学生自身放在一个自主探究的位置之上,他们会呈现出各式各样的设计方案,无论实验设计方案是好是坏,都应该允许让学生亲历实验之后,对自己的方案进行辩护、讨论、测试和评估进一步分析方案设计的可行性和可靠性如何,如发现可能不同质量的

垫片对实验影响未考虑到,或是没有考虑负载在容器偏离中心位置改变等,再主动对方案完成新的改进。因此,案例设计的情境不再是预设的、静态的、封闭的,而是向孩子们展示一个动态变化的真实世界,形成一种开放的过程观,学生的科学态度直接决定其在科学教育中的卷入动机和兴趣程度,而这种态度的培育应当是鼓励他们在一个真实情境中开展交流互动并产生相互影响,将知识领域与生活经验割裂的态度取向是“虚假”的科学态度。

除了情感态度的目标指向,评分规则的要素也是围绕着情境而展开。由上述的评价量表中可知评价的指标不应该是纯粹考查学生识记方面的内容性知识,而是要通过情境中的具体细节增加能够有效探查学生科学概念和科学探究能力发展水平的操作性要素<sup>[5]</sup>。针对学生对于小船浮沉研究问题的选择是否合适,能否完全地回顾背景信息,并利用较为清晰、准确的科学语言解释现象及缘由,包括实验变量的适当控制和器材设备的操作使用等来准确评定学生科学概念和探究技能的学习进程达到哪一阶段,及时反馈概念理解的偏差和能力发展中存在的缺陷和不足。是学业评定的最重要的功能是促进学生发展,借鉴美国开发新的科学学业评定系统的经验,所以,在使用表现性任务的同时时刻关注问题面向的情境是否与评价要素紧密相契合。

## (三)问题设计关注反思体验

学业成就评价为促进学生发展,通常需要改进教师教学,但是学生对于自我发展轨迹的跟踪与追寻同样不可忽视。案例的表现性任务中,最后一个问题即引导学生回顾本次学习中自己的所学所得,明确自己进一步努力的方向。

通过表格中列举的一些陈述,学生会审视并反省已经发生的学习活动,并对学习结果做出解释和判断,如学生对于“当货物放置在船的边缘时,船会发生翻转并下沉。”这句表述进行思考,是否能产生清晰的思路或持有积极的态度并得出正确结论。倘若对整个解释和判断的过程和结果能以合理和完善的形式表征,即完成了自己在这部分事实性知识的认知优化。反之,如对“船自身越重,它越容易沉入水中。”这句表述使学生内心产生怀疑、困惑等情绪,即意识到问题的存在,并开始广泛搜集并分析已有的相关经验,特别是实验探究时得到的有效信息,对它们进行细致思考,认识问题的成因之后,学生为解决面临的问题,会积极转变新的思维模式并采取不同的学习策略,可以是利用概念转变调整原有的经验结构,或者通过科学探究的方式更加直观地接触科学知识的产生过程。案例中设置这样的问题不仅可以提高儿童对于新信息的解决处理能力,



还能够更有针对性地引发学生的反思意识,自觉地以批判的眼光审查表格中的叙述,在参与这种元认知体验过程中不断产生更加富有有效的认知模式和学习策略。

因此,学习者在经历上述阶段的反思体验后,可以看出对于回顾发生过的学习活动是体验过程的开端,确切来说更是一种核心要素,只有学生在类似表格中这些科学叙述的视觉冲击下,才能自觉调动之前的记忆,去判别该问题是否会让自己产生疑惑,如若可以顺利解答或消释当前的迷惑,即检验成功,可以继续对这部分事实性知识总结提高,深化延拓;如若在检验过程中发现自己对这个问题目前不具备解决的能力或者脑海中还未完全建构相应的概念,那么就需要经由上述的方式去探索并提炼方法,如此加以循环往复,即学生的反思体验活动是一个螺旋式上升的进程,前后的阶段往往会前后交错,其间界限并未如此分明,只要最终可以通过实践去验证一些具体问题是否能够得以解决<sup>[6]</sup>,于是儿童在不觉间对于这部分科学知识的整体水平就上升了一个新的层次,从教师评价的视角来看,也是及时跟进形成性评价,通过反馈适切性地调整教学策略,为终结性评价工作做准备的有效方式。


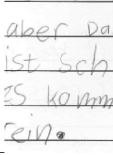


2 Points		Das kleine ding macht auf der einen seite schwer	Translation: The disc makes the ship heavy on one side.
1 Point		es bleibt oben aber das schiffchen ist schreg und es kommt wasser rein	Translation: The ship floats but tilts and water comes in.
1 Point		es kibt	Translation: It turns over.
0 Points		Es geht immer an den rand.	Translation: It constantly moves to the edge.

图1 问题1中学生的任务表现答题情况

### 三、表现性评价案例启示建议

对于解决表现性评价实施在国内科学教育中的部分弊端,既需要挖掘课程标准的相关要求、建构面向真实世界的任务情境,也需要培育学生在表现中的反思学习策略来应对评价中可能出现的挑战。

#### (一)评价应深入挖掘课程标准的内涵

在评价小学生的科学学业时必须以科学课程标准为标尺,才能切实提升科学教育质量,促进科学教育的良性发展。如本文表现性评价案例中的小船倾斜浮沉研究实验,与《NGSS》里“科学和工程实践”部分提出的“发现与明确问题”与“构建科学解释”要求高度契合。2017年,我国颁布了新修订的《义务教育小学科学课程标准》<sup>[7]</sup>。它以培养每一个学生的科学素养为总目标,指出科学课程关注的目标还应强调科学、技术、社会和环境的关系,并根据学生1-6年级不同时期的心理特征和认知规律形成低中高各学段的具体目标;知识领域上除物质科学、生命科学和地球与宇宙科学外增加技术与工程领域,要使学生更加正确地认识自然世界和人工世界的异同点,经历科学探究和工程设计过程,形成对周围世界的整体认识和科学的世界观。同时,在课程实施上,建议采用大概念的课程组织形式,利用统一的科学概念与原理,将不同领域知识与技能加以连接与融通。然而,我国目前的科学学业评价仍然偏重对学科基本知识概念的考查,缺乏对跨学科知识与技能的整合测评;科学探究环节的关注也过于机械死板,学生思维能力的发展受到较大的局限。因此在实施表现性评价时,应当紧紧围绕课程标准,依照上述关于课程目标、课程内容、课程实施等要求完善科学评价框架。

#### (二)评价应着力优化真实情境的建构

一般来说,表现性评价实施情境的真实程度直接反映出学生在知识、技能和情感态度等方面的真实发展状况。在PISA 2015的科学素养测试框架中提及的情境(Context)主要包括科学技术在个人、地区/国家和全球范围内的应用<sup>[8]</sup>,这里并不是评价情境,而是评价学生在特定情境下表现的知识和能力,情境的选择作为表现性评价方式的有效载体和存在基础。但是,学生学习的真实情境在教学评价中有时难以重现的,只能通过情境的模拟,如开展角色扮演活动,操作展演和情景体验等,在一个真实情境中开展互动交流并相互影响,进而在不同学习领域之间建立联系,有利于符合教师期望学生达



成的目标。另一方面,从量表制作的视域来看,关于评分规则的信效度也直接取决于施测情境的真实程度。在使用表现性评价的过程中,首要的是确定与表现性任务相配套的评分标准,其主要来源可以是国家课程标准、实践经验总结、教师专业判断和师生共同商讨<sup>[9]</sup>,对于学生在学习过程中的典型行为反应可以根据特定的表现特征找到合适的维度,依照各指标的权重比例进行分项评分,也可以总结性、概括性地对“产品”整体印象进行评价。一线教师在实践中需要将两种量规评价方式的优劣点加以权衡,紧密围绕具体清晰、操作性强的真实情境,才能打破制约表现性评价在教学实施中的藩篱,为提升科学教育质量添加一条更加科学而严谨、清晰而具体的路径。

### (三)评价应有效培育反思学习的策略

在实施表现性评价时,我们希望能强化学生反思学习的体验感受,着眼于学生的素养发展,但是影响该活动的因素较多,其一,因为学生本身对科学探究活动的反思认识不够,缺乏一定的反思意识,无法调动他们积极的情感因素,良好的学习习惯尚未养成。其二,教师作为学生的模仿对象,大部分老师是否已经掌握反思相关的理论框架及技能方法也有待商榷。其三,家庭及社会教育中对于反思方法的指导力度不强,包括对学生的激励与评价。因此,在保证任务内容多元化的同时,也需要注意对学生反思能力的训练及策略培养<sup>[10]</sup>。帮助学生创设问题情境,利用表格、图像或文字等形式提供激疑环境,科学发现的过程往往都是从质疑开始,从释疑入手,从无疑到有疑,再从解疑到创新。教师适当地引导学生在平时的科学课程学习过程中可以从哪些方面入手进行反思,如在浮与沉的学习过程中,个体采用自主探究式的学习方式还是小组合作式,其各自特点分别如何等;学习中完成的任务是什么,有何特点,我们实现的目标是什么等;完成这些任务有哪些学习策略,该如何选择,如何判断它们是否有效等。最后,应当明确学生的自我效能感对于反思学习的毅力也有较大影响,在教学过程中,帮助儿童正确认识自己的能力、价值及目标,习得这种胜任力后,学生易于对一类知识进行比较归纳,挖掘其中共性的规律,获得一套思路清晰、内容完整、鲜活有效的知识整体,表现性评价的使用也因而重新焕发生机。

表现性评价在评价学生的科学学业成就上还需要考虑施测时间的统筹协调,评分要素的权重如

何分配等问题,但作为对大规模学业评测的一种补充,是极其必要的。由于它旨在引导学生围绕一个真实情境问题,通过收集资料、分析数据、制定方案,测试、交流、评估与优化等环节形成一个完整的解决方案,让儿童更充分地、全方位地展现自己的所学所用所得。

因此,作为教育研究者和一线教师需要思考这些问题,从提升学生科学素养的角度出发,多多借鉴美国课程标准案例中的一些有意义元素,使其成为当前新课程评价改革中所倡导的一种发展性评价方法为后人提供参考。

### 注释:

[1]National Research Council. Developing assessments for the Next Generation Science Standards[M]. National Academies Press, 2014, 38:157-165.

[2]Quinn H, A Framework for K-12 Science Education[M]. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. National Academies Press. 2012:400.

[3]National Research Council. Next generation science standards: For states, by states[M]. Next Generation Science Standards: For States, by States. National Academies Press. 2013:60.

[4]Next generation science standards: For states, by states [M].Next Generation Science Standards: For States, by States. National Academies Press.

[5]徐岩,吴成军.中学生物学科中的表现性评价及其实例[J].课程·教材·教法,2011(8):75-80.

[6]桑志军.反思性学习实践者的内涵、特征及培养[J].教育理论与实践,2012(23):48-50.

[7]中华人民共和国教育部制定.义务教育小学科学课程标准[M].北京:北京师范大学出版社,2017:6,16,59.

[8]OECD (2016), PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy [M].PISA, OECD Publishing, Paris. 2016:23.

[9]张雅君.表现性评价在初中物理科学探究中的应用[D].华中师范大学,2016.

[10]张杰.反思性学习在高中地理教学中的应用研究[D].西北师范大学,2016.

(陈 乾 南京师范大学教育科学学院 江苏 南京 210000)