

初中生科学探究能力评价方式的比较研究

CHUZHONGSHENGKEXUEKUTANJUNENISUPINGJIAFANGSHI DE BIXIAOYANJIU

广西哲学社会科学“十一五”课题
《科学探究能力的表现性评价》
(06FJYX005) 阶段性研究成果

罗国忠 著
广西教育出版社

广西哲学社会科学“十一五”课题《科学探究能力的
表现性评价》(06FJYX005)阶段性研究成果

初中生科学探究能力评价方式的比较研究

罗国忠 著

广西教育出版社

图书在版编目(CIP)数据

初中生科学探究能力评价方式的比较研究/罗国忠著.

南宁:广西教育出版社,2007.10

ISBN 978-7-5435-5005-6

I. 初… II. 罗… III. 科学研究—能力—评价—初中
IV. G632.46

中国版本图书馆 CIP 数据核字 (2007) 第 161776 号

初中生科学探究能力评价方式的比较研究

罗国忠 著

☆

广西教育出版社出版

南宁市鲤湾路 8 号 邮政编码:530022

电话:0771—5865797 5852408(邮购)

本社网址 <http://www.gxeph.com>

读者电子信箱 book@gxeph.com

全国新华书店经销 广西地质印刷厂印刷

*

开本 890mm×1240mm 1/32 7.25 印张 200 千字

2007 年 10 月第 1 版 2007 年 10 月第 1 次印刷

ISBN 978-7-5435-5005-6/G·4060 定价:15.00 元

如发现印装质量问题,影响阅读,请与承印厂联系调换

目 录

第一章 绪论	(1)
第一节 问题的提出	(1)
第二节 文献综述	(3)
第二章 本研究评价体系的构建	(81)
第一节 理论依据	(81)
第二节 评价目标	(82)
第三节 评价任务	(83)
第四节 评分标准	(109)
第三章 学生在三种评价方式中的表现水平及特点	(113)
第一节 学生在现场观察中的表现水平及特点	(113)
第二节 学生在工作单中的表现水平及特点	(132)
第三节 学生在纸笔测验中的表现水平及特点	(155)
第四节 整体讨论和结论	(166)
第四章 三种评价方式中的内部比较	(169)
第一节 现场观察中的内部比较	(169)
第二节 工作单中的内部比较	(173)
第三节 纸笔测验中的内部比较	(179)
第四节 整体讨论和结论	(184)

第五章 三种评价方式之间的比较	(188)
第一节 工作单与现场观察的比较	(188)
第二节 纸笔测验与现场观察的比较	(194)
第三节 纸笔测验与工作单的比较	(201)
第四节 整体讨论和结论	(209)
第六章 总的讨论、结论和设想	(214)
第一节 讨论和结论	(214)
第二节 进一步研究的设想	(217)
参考文献	(220)

第一章 绪 论

第一节 问题的提出

在新旧世纪之交，我国发起了新一轮基础教育课程改革，其纲领性文件《基础教育课程改革纲要（试行）》指出：“改变课程实施过于强调接受学习、死记硬背、机械训练的现状，倡导学生主动参与、乐于探究、勤于动手，培养学生搜集和处理信息的能力、获取新知识的能力、分析和解决问题的能力以及交流与合作的能力。”相应的，针对初中物理学习内容实践性强、与学生的社会生活联系紧密的特点，《全日制义务教育物理课程标准（实验稿）》明确提出“注重科学探究，提倡学习方式多样化”等理念（中华人民共和国教育部，2001a）。很显然，培养目标和教学方式发生了变化，教学评价也必须进行相应的改革。因此，《全日制义务教育物理课程标准（实验稿）》要求“重视学生在活动、实验、制作、讨论等方面表现的评价，不赞成以书面考试为唯一的评价方式”（中华人民共和国教育部，2001a）。然而，迄今为止，我国初中物理各类考试还是通过单一的传统纸笔测验来评价学生的科学探究能力。在国内，在可以检索到的有关科学探究能力评价研究的文章中，绝大多数涉及纸笔测验方式，极少涉及目前国际上主流的科学探究能力评价方式——现场观察、工作单，即使有，也仅是一些思辨性的思考、呼吁或介绍，基本没有实证性的研究成果。总的来说，科学探究能力的评价研究在国内仍处于初级阶段，理论和应用研究都十分匮乏，尤其缺乏科学、系统的设计方法与评分系统，这与新课程大力倡导科学探究是极不相称的。

纵观国际科学教育改革,科学探究能力的评价已发展到多种方式并存的格局。从20世纪60年代的科学课程改革开始,国际科学教育界所采用的主流评价方式包括现场观察,以及现场观察的替代方式——工作单、纸笔测验等,非主流方式包括小样本的口语报告等。尽管现场观察被公认为最有效的科学探究能力评价方式,但其高成本却制约了其大规模使用,因此更多的大规模评价项目主要采用较低成本的工作单。尽管纸笔测验的有效性受到质疑,但这种评价方式成本最低,因此很多评价项目在采用工作单的同时,也都采用纸笔测验。

但是,工作单和纸笔测验能在多大程度上替代现场观察?在哪些方面能替代现场观察?纸笔测验能在多大程度上替代工作单?在哪些方面能替代工作单?这些问题仍然缺乏研究。关于工作单对现场观察的替代性,一些研究者往往从经验出发,把这种可替代性当成不言而喻、不证自明的,只有少数研究者如 Woolnough et al. 和 Shavelson et al. 在1990年初期做了实证研究,得出工作单能替代现场观察的结论(Toh & Woolnough, 1990; Woolnough & Toh, 1990; Shavelson, Baxter, Pine, 1991; Shavelson & Baxter, 1992)。但是, Woolnough et al. 的研究样本太小,其结论的普遍性值得怀疑; Shavelson et al. 的研究只涉及小学生不太复杂的探究能力,其结论也未必具有普遍性。关于纸笔测验对现场观察的替代性,不少研究者得出了纸笔测验难以替代现场观察的结论,但一些研究者(Shavelson, Baxter, Pine, 1991; Shavelson & Baxter, 1992)的现场观察与纸笔测验都采用相同的探究内容,而且先后进行,学生的熟悉效应不可避免,因此其结论不可靠;而一些研究者(Tamir, 1974; Tamir, Doran, Chye, 1992)的现场观察侧重于实验技能而非探究能力。关于纸笔测验对工作单的替代性,一些研究者(Tamir & Doran, 1992a)所采用的纸笔测验基本上是选择性反应的选择题而不是建构性反应的开放题,而且这些工作单内容涉及的探究要素不多,探究水平也不高。由此可见,关于工作单、纸笔测验对现场观察的替代性,以及纸笔测验对工作单的替代性问题,已有的研究还不足以给出令人信服的结论。另外,有研究者(Lawrenz, Huffman, Welch, 2001)发现,不同学习水平的学生在纸笔测验和

工作单中的表现不同,高水平学生在这两种评价方式中表现差异较小,而中低水平学生在这两种评价方式中表现差异较大。这表明,任何一种评价方式都不具有天然的公平性,不谨慎地使用某种唯一的评价方式,可能会使一些学生处于有利地位,而另一些学生处于不利地位(Darling-Hammond, 1993; 罗国忠, 2006a, 2006b, 2007a)。由此看来,进一步弄清学生现场观察、工作单、纸笔测验中的表现差异也很有意义。

因此,本研究将进一步对纸笔测验、工作单、现场观察这三种评价方式进行比较研究,具体研究的问题是:①工作单对现场观察的替代性,以及工作单中的探究任务、探究要素对现场观察的替代性。②纸笔测验对现场观察的替代性,以及纸笔测验中的探究要素对现场观察的替代性。③纸笔测验对工作单的替代性,以及纸笔测验中的探究要素对工作单的替代性。进一步对纸笔测验、工作单、现场观察进行比较研究,可补充和完善已有的国际研究,对我国中学生科学探究能力评价体系的建立也具有重要的理论和实践意义。首先,工作单成本较低,比较适合大规模评价采用。如果本研究所构建的工作单能替代现场观察,将为科学探究能力评价提供多样化选择,有助于改变国内科学探究能力评价方式单一化的状况;如果不能替代现场观察,那么找出其存在的问题,对进一步优化工作单也有积极意义。其次,纸笔测验成本最低,是我国现阶段大规模评价的现实选择之一。弄清现行我国物理中考的科学探究试题对现场观察或工作单的替代程度,以及可改进或难以改进之处,对我国初中物理各类考试中的科学探究命题有重要的参考价值。

第二节 文献综述

一、概念界定

1. 科学探究

1900年以前,大多数教育工作者主要把科学看成是可以直接传授给学生的一堆知识。1909年,杜威在美国科学促进协会的一

次报告中，对这种教育行为进行了批判。报告中提到科学教育过于强调知识的积累，而对科学作为一种思维方式和认知态度却没有给予足够的重视。杜威指出：科学不仅是需要学习的一堆知识，而且也是一种学习的过程或方法。数十年后，杜威进一步提出了“做中学”的口号，强调要利用孩子对活动的愿望和需要，通过实际体验来获得知识。加涅则把学生的认知过程和科学过程结合起来，认为科学概念和原理必须通过科学过程才能习得。而施瓦布认为，科学应该被视作一种概念结构，它随时可能因发现新的证据而被修改。这些观点给人以极大的启示，教师应该以探究的方式将科学呈现给学生，学生也应用探究的方式学习科学知识。

从 20 世纪 60 年代起，施瓦布、杜威以及包括布鲁纳和皮亚杰在内的其他人所做的工作，对新开发的课程教材产生了实质性的影响，其指导思想是保证学生动手实践而不是被动接受或者只是阅读科学，强调科学探究过程的学习与对科学知识的掌握具有同等的重要性。如作为过程的科学（SAPA）、生物科学课程研究（BSCS）、科学课程的改进研究（SCIS）、小学科学研究（ESS）、中级科学课程研究（ISCS）以及物理科学研究委员会（PSSC）等所制定的课程，都整合进了如今至少可以部分地归结为科学探究的内容（AAAS, 2000）。到了 20 世纪 90 年代，科学探究成了科学学习的中心和重点，科学探究既是学习的目标又是学习的方式，这种观念对整个国际科学教育界产生了极大影响，我国科学课程的改革也正是在这种背景下产生的。

尽管科学探究在科学教育中的地位越来越重要，但不同研究者对科学探究概念的界定或描述很不同，这与科学探究的复杂性有关，也与存在多种相互竞争的理论有关（Kind, 1996, 1999; Lunsford & Claudia, 2004）。一般来说，探究可分为广义和狭义两种。广义的探究泛指受好奇心驱使，企图弄清事理的心理倾向或活动。如美国学者韦尔奇等人认为：“探究是人类寻求信息和理解的一般过程。从广义上说，探究是一种思维方式。”人们通常所说的追根究底、好奇、好问就是一种广义探究，这是人类的天性，人皆有之。它既指科学家的专门研究，也指一般人的解决问题的活动；既包括成人那种深思熟虑式的思想实验，又包括儿童那种尝试错误

性的摸索或探索。狭义探究专指科学探究或科学研究，它在研究对象和方式上有其特殊性。正是对于后者，不同研究者有不同界定或描述，如美国学者韦尔奇等人认为：“科学探究是一般探究的‘子集’，它的对象是自然界，是在某种信仰和假设的指导下进行的。”另一美国学者彼得森则进一步指出，科学探究要遵循一定的程序，采用一定的方法，他说：“科学探究是一种系统的调查研究活动，其目的在于发现并描述物体和事物之间的关系。其特点是采用有秩序的和可重复的过程，简化调查研究对象的规模和形式，运用逻辑框架作解释和预测。探究的操作活动包括观察、提问、实验、比较、推理、概括、表达、运用及其他活动。”迄今为止，被国际科学教育界公认权威的科学探究概念是美国国家研究理事会所作的界定。美国国家研究理事会在《国家科学教育标准》中指出：“科学探究指的是科学家们用来研究自然界并根据研究所获事实证据做出解释的各种不同途径。”从这个界定可以看出，科学家从事科学探究的途径、方法未必相同，但研究对象都是自然界，都必须获得证据，然后对这些证据进行种种解释。显然，这种描述与科学主义者的观点不一样。科学主义者认为，科学的对象是纯粹的客体，科学的观察和实验是客观的，所得证据和数据是客观的，对证据和数据的分析也是客观的，因此所得结论也是客观的，因而科学具有不容置疑的客观性和真理性。这种绝对真理观的弊端是毋庸置疑的：对于科学本身而言，无疑人为地设置了障碍，阻碍了科学的发展；对于科学的学习者而言，这不利于学习者批判精神、怀疑精神、求真精神、创造意识和创造能力的培养（罗国忠，2006c，2006d，2007b）。按照后现代科学观的看法，具有不同信仰、价值观和知识储备的研究者在面对同样的证据时，可能会做出不同的解释，这些不同解释要互相竞争，竞争获胜的解释便成为科学理论，写进教科书，以供后人学习。甚至在激进后现代科学观那里，科学是一门解释的学问，科学理论的真理性只存在于一定的语境和关系中，只是一种综合性的、具暂时性质的对话；对同一事实不会有中心、唯一的描述，而是存在着非中心的多元描述；不同理论是从不同角度对事物的透视，它们之间是“平权”的，不同的理论自有其价值所在。显然，激进后现代科学观在消解科学真理客观性的同时，走向

了另一端——主观性，是有失偏颇的。但建设性的后现代科学观并不盲目、简单地否认科学的真理性，它认为科学活动是人类对世界永无止境的探索过程，在这个过程中，人类无限逼近但永远也无法达到真理，因此科学认识结果是客观性和主观性的辩证统一，其主观性表现为科学是一种相对真理，科学的认识过程是一个渐进的、不断探究的、日臻完善的过程；其客观性表现为科学的基础是充分的事实依据，而事实依据的获得则随人们认识水平的提高、事实依据获得手段的优化而不断改善，最终以新的事实依据代替旧的事实依据，从而导致新理论诞生。如此看来，美国《国家科学教育标准》中的科学探究界定具有一定的建设性后现代科学观意蕴，是有利于培养学生创新精神的。

美国国家研究理事会在《国家科学教育标准》中，也从学生角度界定了学生的科学探究，“科学探究也指学生构建知识、形成科学观念、领悟科学研究方法的各种活动”（国家研究理事会，1999）。但是，这个界定未免过于宽泛，其中的“各种活动”容易引起歧义。例如，学生在课堂上听课、或自己看书，也是构建知识的一种活动，但这种活动是科学探究吗？类似地，在课堂上听课、或自己看书，也是形成科学观念、领悟科学研究方法的一种活动，但这种活动是科学探究吗？因此，这个界定太宽泛，很容易让人对学生的科学探究产生泛化理解。我国《全日制义务教育物理课程标准》没有直接界定科学探究的概念，而是给出这样的描述，“学生在科学探究中，通过经历与科学工作者进行科学探究时的相似过程，学习物理知识与技能，体验科学探究的乐趣，学习科学家的科学探究方法，领悟科学思想和精神”。这个描述把学生的活动限定在“与科学工作者进行科学探究时的相似过程”，学生的科学探究不再是宽泛的活动，而是类似于科学工作者的探究过程，但科学工作者的探究过程又是怎样的？却没有给出，因此这个描述同样给人留下很大的遐想与猜疑的空间。

实际上，美国《国家科学教育标准》对科学探究的界定中，“多种不同途径”、“各种活动”意味着科学探究的复杂性，很难详细、具体地描述科学探究的过程，因此，很多研究者在大体界定科学探究的同时，还给出科学探究的一些特征，以进一步阐释科学探

究。美国国家研究理事会指出,“教育中的探究与科学中的探究有相似之处”,这些相似之处体现为五个基本特征,包括:①学习者围绕科学性问题展开探究活动;②学习者获取可以帮助他们解释和评价科学性问题的证据;③学习者要根据事实证据形成解释,对科学性问题做出回答;④学习者通过比较其他可能的解释,特别是那些体现出科学性理解的解释,来评价他们自己的解释;⑤学习者要交流和论证他们所提出的解释。这些特征强调学习过程的质疑、收集证据和做出解释等,适用于所有学段的探究教学活动(国家研究理事会科学、教学及技术教育中心,《国家科学教育标准》科学探究附属读物编委会,2004)。而我国《全日制义务教育物理课程标准(实验稿)》也进一步指出,“科学探究的形式是多种多样的,其要素有:提出问题、猜想与假设、制订计划与设计实验、进行实验与收集证据、分析与论证、评估、交流与合作”。尽管两者的探究要素名称不太相同,但没有本质的差异,都表明科学探究是一种过程,有一定的活动程序或阶段。尽管科学有许多门类,科学家都有各自的研究领域,不同科学家研究问题的方式、途径和手段也有所不同,因而不存在统一的研究模式,但无论他们从事哪一门类或哪一领域的研究,从发现问题到解决问题,都要大体上经过这样一些类似的活动过程或阶段:形成问题—建立假设—制定研究方案—检验假设—概括结论。正是上述这些活动过程构成了被称为科学探究的过程,进而也成为判断某种活动是否是科学探究活动的依据。

2. 科学探究能力

能力指人们顺利地地完成某种活动所必需具备的个性心理特征。能力通常划分为一般能力和特殊能力。一般能力指大多数活动所共同需要的能力,特殊能力是指为完成某项专门活动必不可少的能力。显然,科学探究能力属于特殊能力,是从事科学探究活动所必须的能力。

同样地,由于科学探究的复杂性以及不同研究者的观点差异,人们对科学探究能力也持不同看法,争议的焦点之一是能否把科学探究能力解析为一些要素能力(Lunsford & Claudia, 2004)。一些研究者倾向于认为不同的探究活动存在一些共同的过程要素,因此也倾向于把科学探究能力解析成一系列的要素能力。这类研究者除

使用科学探究能力 (Scientific Inquiry Abilities) 这一称谓外, 还使用其他一些称谓。例如, 一些研究者使用过程技能 (Process Skills) 称谓 (Dillashaw & Okey, 1980; Burns, Okey, Wise, 1985; Germann, 1991; Song & Black, 1991; Doran, Boorman, Hejaily, 1992; Karen & Ostlund, 2004), 一部分研究者使用探究技能 (Inquiry Skills) 称谓 (Tamir, 1974; Gott, 1987; Gott & Duggan, 2002), 还有一些研究者使用实践技能 (Practical Skills) 称谓 (Lock, 1989, 1990, 1992; Buchan & Jenkins, 1992; Doran & Tamir, 1992; Tamir & Doran, 1992a; Tamir & Doran, 1992b; Tamir, Doran, Chye, 1992; Tamir, Doran, Kojima, et al., 1992; Brown & Moore, 1994; Matthews & McKenna, 2005), 等等。尽管这些概念称谓的含义不同, 侧重点也不完全一样, 但有较大交叉。事实上, 一些研究者在交互使用这些概念。在 20 世纪 60 年代的科学课程改革运动中, 美国科学促进会 (American Association for the Advancement of Science, AAAS) 分析了科学家进行科学研究工作的过程及行为, 从中解析出 13 种过程技能 (见表 1-1), 并将这些过程技能融入科学课程 SAPA 及教学活动中 (Kind, 1996)。

表 1-1 13 种过程技能的含义

1. 观察。包括通过感官和凭借仪器设备两方面。要进行研究必须先占有资料, 而要获得可靠的资料就需要观察。不进行观察或缺乏观察技能, 就不能获得有价值的科学事实。
2. 分类。这是将观察到的事物加以整理的过程。分类后有利于作进一步观察。科学家为了使调查研究条理化, 创造了种种分类的方法。
3. 应用数字。指把观察所得的结果用精确的数学关系表达出来。运用数学语言来描述和分析资料, 既省事又能说明问题。
4. 测量。运用尺子、天平、温度计、量杯等有助于获取准确的研究资料, 复杂的测量涉及高度精密的仪器, 需要对仪器的误差加以计量和校正。
5. 应用空间与时间关系。指使观察对象在时空架构中获得定位, 包括对形状、距离、运动和速度的研究及应用等。
6. 交流。就是将观察的结果呈现出来, 方式包括数字、图表、图片或语言文字等。通过交流, 观察结果才得以公开、沟通、检验, 并有利于发现现象之间的关系, 以便进一步预测和推论。

续表

7. 预测。科学远不只是对观察到的现象进行说明,提出聪明的猜想或预言也是科学家工作的重要组成部分。他们常提出“如果……又会怎样?”这样的问题来提出猜想、设计研究思路,使心中怀着某种观念去作观察或进行检验。

8. 推理。这是另一种猜测,是对观察到的现象作解释、思考和逻辑分析。推理能导致某种结果,它有时比观察本身更为有用。科学便是在不断的推理和验证中向前发展的。

9. 下定义。定义可以使表达简洁明了。科学家事先给某个词或某个术语规定明确的含义后,就可以反复使用,以代替啰唆的说明。设想一下,没有基本的科学概念,科学交流将是怎样的混乱和困难。

10. 形成假设。假设是对所要研究的问题提出暂时性的回答,借这种回答以考察观察所得的结果,以便能直接地、清楚地判断观察结果是否支持原有的理论。事实上,假设是一种更高级的猜测,即系统的、有理论依据的猜测。

11. 解释数据。数据本身并没有什么意义,重要的是从数据中所得到的启示。解释数据的目的在于进一步产生推理、预测及假设,进而寻求所研究问题的答案。

12. 控制变量。控制变量的技能在科学研究中至关重要,它涉及实验能否成功,能否真正揭示事物的因果关系。它包括两个方面:一是抓住要观察的对象,将其突显出来,以便集中观察;二是减少或根本排除干扰因素,以便发现因果关系。

13. 实验。实验事实上是人为的一种观察,所以有人说观察是被动的实验,实验是主动的观察。在17世纪伽利略发明这一方法之前,科学进展缓慢,举步维艰,随着实验法的产生和完善,科学进步日新月异,故有“科学是实验的科学”之说。也可以说,一个科学家的实验能力的大小,是其研究水平的一个重要标志。

AAAS认为,前8种为基本技能(Basic Skills),后5种为综合技能(Integrated Skills)。学生掌握了基本技能,就能逐渐掌握综合技能,直至从事复杂的科学探究。由于SAPA课程在国际科学教育界影响很大,其所包含的过程技能教学和评价的针对性和操作性也很强,因此过程技能的分类受不少研究者和科学教师的青睐(Kind, 1996)。但是,另一些研究者对科学探究能力的解析性持不同看法,他们更强调科学探究能力的缄默性。Woolnough & Toh (1990)认为,科学探究是一个不太整齐的过程,它未必有一种方

法,可能有多种方法,这与研究者的灵感、直觉、兴趣、世界观等主体特征有很大关系。而 Hodson (1992, 1993) 则进一步认为,真实情境中的科学探究是一个整体的 (Holistic)、流动的 (Fluid) 活动,探究中的决策和行动依赖于情境,而探究的每一步都会以某种方式改变情境,下一步的决策和行动将建立在新情境中。这类研究者强调科学探究活动本身的复杂性和综合性,科学探究能力是一种综合性很强的综合能力,既难以分解成各种能力,也很难诉诸文字,由此科学探究能力被描述成“工匠技能”或“缄默知识”(Millar & Driver, 1987)。这种“工匠技能”或“缄默知识”不可明确地教,只能通过体验而内化,因而提高科学探究能力的有效途径就是在真实情境中做科学探究,并由一个有经验的行家里手提供现场支持、批评和建议 (Woolnough & Toh, 1990)。这种观点也产生了一定影响,但它对科学探究能力的描绘过于模糊,人们难以把握其明确的特征和过程,教学和评价缺乏可操作性,既难以在大班额中组织有效的训练,也难以描述学生的进步,更难以描述评价的内容。在科学探究教学的历史上,就有因教学和评价目标模糊,导致教学难开展、难管理,教师缺乏教学热情、学生成绩差等问题 (Lunsford & Claudia, 2004)。

Toh & Woolnough (1990) 认为,科学探究能力既有解析性的一面,也有缄默性的一面。为此,他们还研究了显性知识和缄默知识在科学探究中的作用,显性知识对应于教师的明确指导,缄默知识对应于学生的自主摸索。他们发现,仅仅通过教师的指导或学生的自主摸索,对改善学生探究表现的作用不大,但两者结合则大大改善了学生的探究表现。该研究结果不仅意味着科学探究能力既有解析性也有缄默性,而且意味着它们的教学途径不同,科学探究的显性知识可以明确地教给学生,但其缄默性知识需要学生自主摸索动手才能体会,两者结合能更好地掌握科学探究能力,即不仅能把科学探究概念化(或陈述性知识),也能把科学探究内化(或程序性知识)。

正如后现代科学观所言:“对同一事实不会有中心、唯一的描述,而是存在着非中心的多元描述,这些描述都是从不同角度对事物的透视,它们之间是‘平权’的,都有其价值所在。”虽然此话

未免偏激,但有助于辩证看待科学探究能力的解析性和缄默性。其实,解析性和缄默性位于一个维度的两端,在这两端之间并非绝对的解析性或缄默性,只不过是有所侧重罢了。对于教育中的科学探究,片面强调解析性或缄默性都无助于科学探究的教学。片面强调科学探究能力的解析性,将导致科学探究能力的机械训练;反之,片面强调科学探究能力的缄默性,将导致科学探究神秘化,失去了教学意义。理性的做法应是两者辩证结合,遵循学生的认知发展规律,在不同教学阶段有所侧重,由简单到复杂,循序渐进地发展。在初级阶段训练某些要素,然后不断整合,到高级阶段鼓励灵活运用,发挥创造性。这个道理与篮球运动类似,打篮球是个整体运动,不同运动员在比赛中的动作流程、风格等未必一样,但这并不能否认篮球运动有些基本的动作要素,因此体育教师常常在教学中让学生练习这些分解的基本动作,如运球、三大步上篮、投篮、扣篮等基本动作,待学生把这些基本动作学会后,再整合,直至熟练运用,随机应变、瞬息决定用哪个动作,这是打篮球的解析性和缄默性的辩证统一。

因此,美国《国家科学教育标准》强调科学探究有“多种不同途径”和“各种活动”,我国《全日制义务教育物理课程标准》也强调“科学探究的形式是多种多样的”,两者是一致的,都表明了科学探究能力的复杂性和缄默性。尽管如此,美国《国家科学教育标准》还是把科学探究能力解析成一些要素能力,表1-2为美国国家研究理事会等所划分的5~8年级的要素能力。

表1-2 美国《国家科学教育标准》中5~8年级的探究能力

判定可以通过科学实验或调查研究能回答的问题。
设计和进行科学实验或调查实验。
利用适当的工具和技术收集、分析和解释数据。
运用事实证据进行描述、解释、预测和构建模型。
运用批判性和逻辑性的思考建立证据与解释之间的关系。
认证及分析另外的解释和预言。
交流科学过程和解释。
把数学运用在科学探究的各个方面。

从表 1-2 中看到, 诸如观察、分类和测量等过程技能似乎消失了, 但实际上, 这些过程技能被整合进了更加广泛的科学探究能力之中, 这些探究能力超越了过程技能的含义。类似地, 我国《全日制义务教育物理课程标准(实验稿)》把科学探究能力解析成七要素: 提出问题、猜想与假设、制订计划与设计实验、进行实验与收集证据、分析与论证、评估、交流与合作。

3. 表现性评价

表现性评价(Performance Assessment)是近十年来教育评价领域中炙手可热的评价方式。表现性评价的权威专家 Stiggins 最初把表现性评价界定为“表现性评价是指为测量学习者运用先前所获得的知识解决新异问题或完成具体任务的能力的一系列尝试, 具体来说就是运用真实的生活或模拟的评价任务来引发最初的反应, 由高水平的评价者按照一定标准对这些反应进行直接的观察、评判, 其形式主要包括建构反应题、书面报告、作文、演说、操作、实验、资料收集、作品展示”(Stiggins, 1995)。这个界定强调任务的真实性、学生的建构性反应、评分者的直接观察等。后来, Stiggins 又把表现性评价界定为“让学生参与一些活动, 要求他们实际表现出某种特定的表现性技能, 或者创建出符合某种特定标准的成果或作品。简言之, 就是我们在学生执行具体的操作时直接观察和评价他们的表现”(Stiggins, 2005)。尽管 Stiggins 前后两个界定的文字表述不同, 但没有实质性改变, 都强调学生的活动以及评分者的直接观察。而美国教育评价技术处则将表现性评价界定为“通过学生自己给出的问题答案和展示的作品来判断学生所获得的知识 and 技能”(Elliott, 1996)。这个定义未免太过宽泛了, 主要强调学生的建构性反应, 没有把表现性评价的其他特征表现出来。按此界定, 学生做除选择题外的简答题和论述题时所作的反应是建构性反应, 似乎也算是表现性评价。但是, 表现性评价的另一权威专家 Wiggins 认为: 不能仅仅因为一项设计是一个建构性反应, 我们就说它是真实的甚至是表现性任务(Wiggins, 2005)。由于表现性评价的复杂性和研究者观点的多元性, 还有其他表现性评价的界定, 但 Stiggins 的界定被更广泛地接受, 因此本研究也主要采用 Stig-