

国外青少年科学素质测评工具述评*

李秀菊 陈 玲

摘要：科学教育的目标之一就是培养青少年的科学素质。青少年科学素质的测评成为科学教育研究者和各国教育部门关注的重点。本研究对国际上主流的青少年科学素质测评工具的测评对象、测评内容、测评工具的主要形式等进行了综述与评析，并对我国青少年科学素质测评工具开发提出几点思考，以期开发出适合我国青少年的科学素质测评工具提供借鉴和参考。

关键词：青少年；科学素质；测评工具

作者简介：李秀菊 / 中国科普研究所副研究员，博士（北京 100081）

陈 玲 / 中国科普研究所研究员，博士（北京 100081）

培养青少年科学素质 (Scientific Literacy) 是科学教育的重要目标，这一点已经在国际范围内达成共识。^[1] 科学素质测评也一直都是科学教育研究者关注的重点。自 20 世纪 60 年代开始，越来越多的研究者致力于开发评测个人或者群体的科学素质的相关工具。青少年作为国家发展的智力资源，是一个国家科技发展的重要后备人才。今日之青少年，明日之公民，青少年的科学素质水平也决定了一个国家或地区公民的科学素质水平的发展。因此，青少年科学素质的测评研究受到越来越多关注和重视。

本研究综述了国外一些重要组织和主流学者

开发的青少年科学素质测评工具，对这些工具的测评对象、测评内容、测评工具主要形式、取样方法等方面进行分析，以期开发出适合我国青少年特点的科学素质测评工具提供借鉴和参考。

一、本研究选取的青少年科学素质测评工具

本研究摘取了自 20 世纪 60 年代以来，国际上一些学者研制的青少年科学素质测评工具（见表 1）、国际组织研发的青少年科学素质测评工具（见表 2）。国际组织研发的青少年科学素质测评工具主要选择了国际学生评价项目 (PISA)、国际数学与科学趋势研究 (TIMSS) 和美国国家教育进展评估 (NAEP)。

表 1 国外学者研发的科学素质测评工具

开发者	年代	工具内容	题目类型	测评对象	信度	效度
Kimball ^[2]	1967	科学本质的理解程度 Understanding of NOS	29道李克特式 题目 ^①	N=712 大学生★ ^②	分半信度(内部 一致性信度)	内容效度 建构效度 区分效度
Moore & Sutman ^[3]	1970	科学态度和科学兴趣测试 Attitude and interest toward science	60道李克特式 题目	N=67 10年级	再测信度	建构效度

* 本研究系中国科普研究所“青少年科学素质测评体系及案例研究”(编号: 2016LYE030102)项目成果之一。

(续 表)

开发者	年代	工具内容	题目类型	测评对象	信度	效度
Fraser ^[4]	1978	科学态度测试 TOSRA (Test of Science- Related Attitudes)	70道李克特式 题目	N=1337 7-10年级	内部一致性信 度	内容效度 区分效度
Aikenhead & Ryan ^[5]	1992	STS的观点测试 VOSTS (View of Science- Technology-Society)	114多重选择题	N=5000 高中生	内部一致性信 度	
Laugksch & Spargo ^[6]	1996	Test of Basic Scientific Literacy 科学本质、科学内容、科学与 社会关系 NOS, Nature of Technology, Earth/Space Sciences, Physical/Chemical Sciences, Life Sciences, Health Sciences	110道是非题	N=625 大学生★	内部一致性信 度	内容效度 建构效度
Manhart ^[7]	1997	科学内容, 科学与技术, 科学 作为人类的事业以及科学社 会视角 Physical Science, Life Science, Earth/Space Science, Abilities Necessary to Engage in Scientific Inquiry, Nature of Scientific Inquiry, Knowledge, Science as a Human Endeavor, Science and Technology, Societal Perspectives	多重选择题	N=1139 9-12年级	内部一致性信 度	内容效度 建构效度
Glynn & Koballa ^[8]	2006	学习科学的动机 Motivation Toward Learning Science	30道李克特式 题目			
Kathryn Garthwaite et al ^[9]	2013	光和健康两部分内容测试	多项选择, 短问 题, 和开放式问题	N=95 10年级		
Kongji Mun et al ^[10]	2015	内容知识(科学领域的核心 概念)、思维习惯(科学实 践)、特征和价值观、作为人 类事业的科学和元认知与自 我指导 Content Knowledge (Core Ideas of Science), Habits of Mind (Science Practice), Character and Values, Science as Human Endeavor, Metacognition and Self-direction	114个 likert 式 量表	N=3202 7-12年级	内部一致性信 度	内容效度 建构效度

表2 国际组织开发青少年科学素质测评工具

测评名称	开发者	测评目的	年代	工具内容	题目类型	测评对象
PISA ^{[11][12]}	国际经合组织(OECD)	对接近完成基础教育的15岁学生进行评估,目的是测试学生在基础教育将近结束时能否掌握参与社会所需要的科学知识与科学技能。	2000年开始,2006年重点科学素质测评,每三年开展一次测评	四维度:背景、科学知识、科学能力和科学态度	简单多项选择题、复杂多项选择题、封闭式自建答案题目 开放式回答	15岁的未成年人
TIMSS ^{[13][14]}	国际教育成就评价协会(IEA)	了解学生对数学和科学课程的掌握情况,对各参与国的教育做一个比较清晰的描述,使参与的国家明白本国教育的优势及弱点。	20世纪60年代开始,1999年开始每四年开展一次测评,每次测评都包括科学和数学两部分	三维度:内容维度、认知维度和综合维度	多项选择题 建构反应试题	4年级和8年级的学生
NAEP ^{[14][15]}	美国教育部开发,ETS(由国会授权,美国教育考试服务中心)组织实施	美国教育进展评价(NAEP)是美国唯一的全国性的、代表性和持续性的评价学生学业成就的项目。NAEP的科学评价依据美国国家科学课程标准,评价的目标是学生知道什么,能做什么	从1969年开始,每隔2年进行一次全国性的测试。	二维度:科学知识和科学能力	选择反应试题 建构反应试题	4年级、8年级和12年级

二、分析结果

1. 评测对象分析

对表1和表2中的青少年科学素质测评工具分析结果表明,由研究者独立开发的科学素质测评工具适应的测评对象全部在初中及以上年级,大部分是针对高中生;并且测评工具适应范围较宽,多可供初高中或者至少供2-3年级学生使用。国际组织研发测评工具则针对固定的年级或者年龄,如PISA针对15岁的学生,NAEP的测评对象是4年级、8年级和12年级,TIMSS的测评对象是4年级和8年级。

本研究认为,针对固定的年级或者年龄的科学素质测评工具需要依据学习进阶理论,按照测评年级或者年龄应该掌握的科学知识和科学能力等设计测评题目,具有一定的难度。一般情况下,需要很多研究者组成研究团队开展相关工作,这也是个体研究者多开发适应年龄段比较宽泛的测评工具的原因之一。

同时,较多的测评工具针对高中生,这与科学学习的年限有一定关系。在一些国家,如新西兰

和中国,高一年级是学生完成科学必修课的最后一年(高二年级和高三年级则多是安排一些科学选修课),对高中生的测试,体现了学生们在学校科学学习必修完成阶段的科学素质状况。

2. 测评工具的主要内容分析

分析表1的测评工具发现:多数测评工具只是对科学素质的1-2个维度进行测量,特别是对科学本质、科学态度和科学内容的测量较多,也有的关注了科学与社会的关系。包含多维指标的测评工具,典型的有以下几个。Laugksch& Spargo(1996)开发的基本科学素质测评问卷(Test of Basic Scientific Literacy)相对全面,考察了三个维度:科学内容(生命科学,物质科学,地球与空间科学,健康科学)、科学本质和科学与社会的关系。韩国学者Kongju Mun(2015)开发的全球科学素质调查问卷(Global Scientific Literacy Questionnaire, GSLQ),包括五个维度:内容知识(科学领域的核心概念)、思维习惯(科学实践)、特征和价值观、作为人类事业的科学和元认知与自我指导。PISA(2015)科学素质测评

中包括四个维度：情境、科学知识(生命科学,物质科学,地球与空间科学)、科学能力和科学态度。TIMSS(2015)的科学素质测评包括三个维度：科学内容、科学认知和科学实践。NAEP(2015)的科学素质测试包括科学知识和科学实践两个维度。

能够看出,在科学内容方面,生命科学、物质科学和地球与空间科学已经成为核心领域,有的测评增加了工程技术领域。科学能力(科学实践)维度主要考察对象完成一个完整的科学探究活动的相关能力。

总的来说,典型的科学素质测评工具中,科学内容、科学能力(科学探究)已经成为核心要素。此外,依据研究目的的不同,增加对科学本质(科学与社会的关系)以及科学态度等维度等。这说明,研究者对于青少年科学素质的维度构成的认识越来越清晰和趋于一致。

3. 测评工具的主要形式

总结表1和表2的科学素质测评工具发现,大部分科学素质测评工具都是纸笔测试。随着计算机技术在大型测试中的应用不断发展,有的科学素质测评也是学生在计算机上完成,但是形式上还是与纸笔测试一致,并没有实作评价。

大部分科学素质测评工具的题目都以封闭式的多重选择题、李克特式量表(三点或者五点)、是非对错题和开放式问题为主。其中,对于科学与社会关系和科学态度的测量多采用李克特式量表的方式,主要探查测评对象的观点和态度,每道题目一般只属于一个测评维度。PISA的态度测试采用李克特式量表的方式开展测评。Laugksch&Spargo开发的基本科学素质测试采用是非题的方式,每道题目只属于一个测评维度。PISA和TIMSS测试中科学素质题目以封闭式的多重选择题为主,多采用在一定的背景下,基于科学内容载体,考察科学能力。这是一种相对复杂的命题方法。多数测量工具均采用计分的方式,通过计量测评对象的总得分衡量其科学素质水平。

综上信息,大规模的青少年科学素质测评工具形式仍然以纸笔测试为主(尽管有的在计算机上完成,但是形式未变)。测试题目形式多样,问答题一般需开发评价量规(Rubric)。一般通过计分的方式衡量测评对象科学素质水平的高低。

4. 取样方法

综合本研究选取的测评工具发现,目前,青少

年科学素质测评取样主要有两种类型。

一种是大型调查,如PISA、TIMSS、NAEP等,属于大规模取样,一般是分阶段抽样的方式:第一阶段完成学校阶段的抽样,按照学校的类型、地理位置、学校性质、学校等级和学段等因素设计抽样框;第二阶段在学校抽取学生,一般是随机抽样或者整班抽取。

第二种是典型调查,比较多的研究者自行开发测评工具后,对某一所或者几所学校的学生开展测评,典型调查抽样基本采用随机抽样和方便抽样的方式开展。大规模取样重点关注群体水平,典型调查取样在关注整体的同时也能够关注到特定的个体。

三、关于在我国开展大规模青少年科学素质测评的几点思考

1. 关于测评对象

对青少年学生开展测评,首先要考虑的问题是对几年级的学生开展测评或者是对什么年龄段的学生开展测评。基于对已有工具的研究和分析,测评对象的选择应该从科学学习的阶段入手,建议在以下三个阶段中选择,即完成小学阶段科学学习的学生、完成初中科学学习的学生、完成科学必修课的高中中学生。主要原因如下:第一,在我国,小学科学是综合教学,全国范围内的小学生从小学四年级开始开设科学课,到小学六年级完成了小学阶段的科学学习。这一阶段的科学学习,在科学知识方面综合性强,在学习方式上学校的正规科学学习和校外的非正规科学学习都相对自由和自主,开展相对完整的科学探究的机会较多。第二,我国大部分地区的初中科学教学采用了分科教学方式,因此在初中阶段的科学学习中,学生能够对各个学科体系有初步的认识和理解,初中属于义务教育,能够全面地通过调查学生的科学素质反映这个年龄段我国青少年的科学素质水平。第三,对高中完成科学必修课的高中生开展科学素质测评,这一阶段的学生已经完成了基础教育阶段必修科学课的学习,科学素质水平相对稳定在一定的状态。

此外,基于学习进阶理论,学生在不同的认知阶段,对科学理解深度和广度都不同,因此,研究者在开发测评工具的过程中,应该基于这一理论基础,对于不同认知水平的学生开发不同的测评工具。目前,在我国,学习进阶理论的研究处于

相对初级阶段,开发低年龄段的科学素质测评工具挑战很大。这也是国际上常用的民间学者开发的测评工具也多针对初高中生和大学生的原因之一。开发低年龄段的科学素质测评工具,对施测过程要求也较高。

2. 测评工具的主要内容和形式

测评工具的主要内容与研究者对于科学素质结构的认识密切相关,同时也与科学素质测评的目的有紧密的联系。综合分析国际上针对青少年学生开发的科学素质测评工具科学内容(知识和概念维度)、科学能力(科学探究和过程维度)以及科学态度(对科学的态度和对科学本质的认识维度)是科学素质测评的核心要素。其他要素,比如自我认知等维度可以依据研究需要扩展,也可以包含在上述三维度中。

测评工具的题目形式主要分为四大类,李克特式、多项选择题、是非题和开放式问题。李克特式一般采用三分法或者五分法,通过对不同程度赋予不同分值进行统计计算。是非题和多项选择题都是依据正确答案得分,错误答案不得分的方式统计分数。开放式问题一般采用评价量规或者评价指南的方式,依据答案的正确程度给予不同的分值。一般对于科学态度的测评多采用李克特量表的方式开展。对于科学内容和科学能力的测评则各种方式都有。由于科学能力的复杂性,对于科学能力的测评采用多项选择题和开放式问题的形式居多。开放式问题的评价指南或者评价量规以及同时考察多方面科学内容和科学能力的多项选择题对开发者的挑战较大。

3. 取样方式

青少年学生的科学素质测评的取样多以学校为第一层取样单元,班级为第二层取样单元。在预调查中,一般基于典型性和代表性的原则取样。如果在全国层面上实施调查,则需要以省级行政单位为一级单元,以县(区、市)为二级单元,学校为三级单元。在取样时注意考虑城市与农村、教育发达地区与欠发达地区等因素综合考虑抽样。

注释:

- ① 表1中李克特式量表是“likert”的中文翻译,一般是考察态度时应用,题目的表现形式为“非常喜欢”、“喜欢”、“不喜欢”、“非常不喜欢”等。多为三分法(三种程度)或者五分法(五种程度)。

- ② 一般的,青少年指6-18岁的未成年人。表1测评对象中带星号的工具虽然测试对象是大学生(College Student),是指作者发表该篇文章时测试对象是大学生,但是实际上工具的适应对象可以下延到高中生,因此也将该测评工作作为分析对象。

参考文献:

- [1] Laugksch, R. C. (2000) Scientific literacy: A conceptual overview. *Science Education*, Volume 84, Issue 1, p. 71-94.
- [2] Kimball, M. E. (1967). Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching*, 5(2), 110-120.
- [3] Moore, R. W., & Sutman, F. X. (1970). The development, field test and validation of an inventory of scientific attitudes. *Journal of Research in Science Teaching*, 7(2), 85-94.
- [4] Fraser, B. L. (1978). Development of a test of science-related attitudes. *Science Education*, 62(4), 509-515.
- [5] Aikenhead, G. S., & Ryan, A. G. (1992). The development of a new instrument: 'Views on Science-Technology-Society' (VOSTS). *Science Education*, 76(5), 477-491.
- [6] Laugksch, R. C., & Spargo, P. E. (1996). Construction of a paper-and-pencil test of basic scientific literacy based on selected literacy goals recommended by the American Association for the Advancement of Science. *Public Understanding of Science*, 5(4), 331-359.
- [7] Manhart, J. J. (1997). Scientific literacy: Factor structure and gender differences (Doctoral thesis). p.12-27. The University of Iowa, Iowa City.
- [8] Glynn, S. M., & Koballa, T. R. (2006). Motivation to learn college science. In J. J. Mintzes & H. L. William (Eds.), *Handbook of college science teaching* (pp. 25-32). Arlington, VA: National Science Teachers Association Press.

(下转第5页)

Policy Innovation with the Support of the Country
----- Comment on The Center for Education Policy Research

Special Correspondent LI Qianya

Editor's note: The Center for Education Policy Research (CEPR) is one of the well-known Research Centers in Harvard University. In the 《2014 Global Go To Think Tanks Report》 and 《2015 Global Go To Think Tanks Report》 released by the University of Pennsylvania, in the top think-tank ranking on the field of education, CEPR was ranked the sixth for two consecutive years. CEPR has not only established a powerful network, but also had two advantages as talent resource and state support. Through many projects, it has given analysis on the data related to schools, teachers and students to solve the difficult problems in education policy, promoting education reform in the United States. Through several approaches like literature method and interview method, our special correspondent makes a comment on various aspects such as the general operation of the Think Tanks, the research focus etc. in order to provide enlightenments for the construction and development of an education think-tank in China.

(上接第 18 页)

- [9] Kathryn Garthwaite, Bev France & Gillian Ward (2014) The Complexity of Scientific Literacy: The development and use of a data analysis matrix, *International Journal of Science Education*, 36:10, 1568-1587.
Kongju Mun, Namsoo Shin, Hyunju Lee, Sung-Won
- [10] Kim, Kyunghye Choi, Sung-Youn Choi & Joseph S. Krajcik (2015) Korean Secondary Students' Perception of Scientific Literacy as Global Citizens: Using Global Scientific Literacy Questionnaire, *International Journal of Science Education*, 37:11, 1739-1766.
- [11] <http://www.oecd.org/edu/pisa-2015-assessment-and-analytical-framework-9789264255425-en.htm> .
- [12] 刘克文, 李川. PISA2015 科学素养测试内容及特点. [J]. *比较教育研究*, 2015, (7) :98-106.
- [13] <http://timss.bc.edu/>.
- [14] 夏小俊, 顾乃景, 柏毅. TIMSS 2015 与 NAEP 2015 科学评估框架的比较研究 [J]. *考试研究*, 2016, (1): 96-101.
- [15] <http://nces.ed.gov/nationsreportcard/>.

A Review of Measuring Tools on Youth's Scientific Literacy

LI Xiuju CHEN Ling

Abstract: Cultivating youth's scientific literacy is one of the most important goals of science education. Therefore, measuring youth's scientific literacy becomes the focus of attention of science education researchers and governments. This paper reviews the measuring subjects, the contents, and the tools of major measuring scales and questionnaires. At the end of this paper, we provide several suggestions for the development of measuring tool suitable for China.

Key words: youth; scientific literacy; measuring tools