

学前儿童科学学习与学习环境评价

Kimberly Brenneman
Rutgers University
National Institute for Early Education Research

摘要

在美国，尽管对学前科学教育的兴趣并不是一件新鲜事，但是该领域的学习引起了学前教育较为关注以及具有较高科学素养和成就的公民的重新关注。尽管早期科学教育吸引了更多的关注和资金投入，人们也期望通过高质量的教育支持来提高科学和相关领域的入学准备和成绩，但是因为科学教育缺乏合适的学习和课堂教学质量评价的工具，研究和机构的评估就受到了较大的局限。这篇文章报告了几种很有应用前景的工具和方式来评价儿童科学学习的过程和支持这种学习的教学质量的评价。本文讨论学习和知识评价，包括支持和评价日常互动中的科学学习、个性化教学的表现性评价、过程性监测以及课程评价、科学学习的直接评价、以及科学相关技能和品质的评价。同时，本文还讨论了与科学学习有关的课堂质量的测量。

引言

尽管对学前科学教育的兴趣并不是一个什么新的话题(有关科学教育的综述，请参看 Riechard, 1973)，但是该领域的学习引起了学前教育较为关注以及具有较高科学素养和成就的公民的重新关注。在早期教育领域，使用较为普遍和受人尊敬的学前教育课程如创造性课程(http://www.teachingstrategies.com/page/CCPS_Studies.cfm)的开发人员强调科学方面的学习内容，纯科学内容方面的教育课程也

已经产生(Brenneman, Stevenson-Boyd, & Frede, 2009)。全美幼教协会 (NAEYC, n.d) 认为, 应该给儿童提供各种不同的机会和材料来学习科学的核心内容和原则。早期开端计划中有关儿童学习成果的框架中也把科学放在八个入学准备的领域中(U.S. Department of Health and Human Services, 2003)。不管是作为独立的一个学习领域还是作为认知和语言发展的一部分, 美国大多数州都对学前科学学习提出了学习目标 (Scott-Little, Lesko, Martella, & Milburn, 2007; Snow & Van Hemel, 2008)。当前奥巴马政府行政管理机构把科学领域作为学前教育阶段的优先发展内容 (“Barack Obama and Joe Biden's Plan,” n.d.), 国家科学基金会学前教育-高中 (K-12) 探究性研究计划已经开始实施并发布了研究课题申请通知, 并已开始资助科学、技术、工程以及数学教育方案 (STEM) 的研究, 以帮助 3-5 岁儿童和幼儿教师。科技和金融界的领头人物也努力支持美国学前儿童有机会接触科学、技术、工程以及数学(e.g., American Honda Foundation, 2010; Motorola Foundation, n.d.)。PNC 基金会 (2009) 通过它的“成功与科学成长”的计划, 已经在非正式科学教育组织和学前机构中投入资金。波音公司提供资金支持美国公共电视台 (PBS) 的“科学儿童”节目(<http://pbskids.org/sid/>), 诺斯洛普格拉曼 (Northrop Grumam) 公司资助的“窥视大世界”节目 (<http://www.peepandthebigwideworld.com>), 这两个资助的学前教育节目都是有关科学教育的。

这些课程、政策和经费的资助反映了人们的信念: 儿童早期接触科学、技术、工程以及数学概念可以更好的适应未来的生活, 早期经验对入学准备和未来学习起关键作用 (参看 Beering, 2009)。正是因为很多已有的研究对儿童早期学习经验对长期的教育和社会影响的重要性已经达成共识, 尤其是能对那些还未享受到这种服务的人群所产生重要的作用的肯定(Barnett, 2008; Bowman, Donovan, & Burns, 2001; Committee for Economic Development, 2006; Shonkoff & Phillips, 2000), 而且研究表明, 一般的课堂质量考察与儿童的学习结果之间具有中等水平的相关[Burchinal et al., 2008; Melhuish et al., 2008 (EPPE); Haahr, 2005 (PISA)]。根据这些研究的结果, 我们认为, 在早期提供高质量的科学学习经验, 将获得有助于儿童喜欢科学和在科学上取得长期成就的回报, 这一假设是合情合理的 (National Research Council, 2005)。增加能够明确应对这一问题的研究非常迫切, 尤其是最近美国佛罗里达州一项大规模的研究表明, 科学入学准备比其他的

领域要晚，至少对于那些学习很可能有问题的学生是如此(Greenfield, Jirout, et al., 2009)。

尽管对早期科学教育的兴趣和投入都在增加，我们也期望高质量教育支持将会提高科学及其他方面的入学准备和成绩，但评价和研究却因为缺乏合适的工具而有很大的局限性。最近国家研究委员会（NRC）的学者报告了儿童早期的评价(Snow & Van Hemel, 2008)，认为早期教育科学领域中的评价不属报告的范围，因为“很简单它并不是理论、研究或者实践的基础... 尽管科学的重要性不言而喻（p.59）”。目前，早期教育领域还没有能回答与支持和提高科学的教学和学习有直接关系的问题的工具。这些工具和方法将会反映早期科学特殊的视角以及更一般的教育的视角。因此，任何人设计和使用评价工具的首要一点是明确儿童学习中的思维技能和内容的学习目标。然后我们可以提出问题：“我们设计的非正式和正式的学前科学教育方案能有效地达到这些目标吗？会不会有些方案比其它方案更有效？一种典型的具有科学思维的课堂需要什么样的材料和教学的互动？我们如何从有限的教育资源中获得最大的“成功”？我们如何保证给每个儿童提供合适的学习机会，所以他们对科学学习的兴奋、热情、已有的内容知识和推理技能能得到保护和进一步发展”。

最后的问题也许是那些每天与儿童打交道的成人最感兴趣的。发展方面的研究告诉我们，儿童入园之前就已经拥有关于物理、化学、心理学以及生物科学领域的粗略知识，他们已经开始推理，并形成后来科学思维的基础(Duschl, Schweingruber, & Shouse, 2007)。幼儿探究周围世界的方法也使我们联想到科学家的探究方法。下面来自克兰南和欧克斯（Callanan and Oakes, 1992, pp. 221-222）的母女对话就提供了强有力的证据：

儿童：为什么爸爸，吉姆(哥哥)和我有蓝色的眼睛，而你是绿色的？
妈妈告诉她，她的蓝眼睛是来自爸爸，告别晚安后，就离开了房间。

5分钟后，孩子叫她的妈妈，说：我喜欢赫尔曼，然后我有了蓝眼睛，爸爸喜欢赫尔曼，然后他有了蓝眼睛。吉姆喜欢赫尔曼，他也有了蓝眼睛。如果你喜欢赫尔曼，你也会有蓝眼睛。

妈妈告诉女儿说，不是喜欢赫尔曼就会有蓝眼睛的。这时她才意识到儿童不能理解和解释为什么上帝给了妈妈绿眼睛，这一颜色不能改变。

儿童：你能不能尝试喜欢赫尔曼？这样我们可以看一下你的眼睛是否会变成蓝色。

在这个短暂的互动中，儿童使用了多种探究技能，包括对眼睛颜色和电视节目喜好的*进行观察和观察结果的描述*，*比较*它们，*询问*眼睛颜色的来源，*反思*妈妈跟她解释的为什么会有这些差异的原因（也许这种解释比较隐晦），她不能理解这个解释，对看到的差异的原因*创造了自己的解释*，然后，*设计了一个测试*来发现自己的解释是否正确。这个例子似乎有些特别，但是我们这些与孩子打交道的人会发现自已遇到过相似的故事，这反映出儿童早期思维的能力，我们对此有责任来支持、庆贺以及对这些能力提出挑战。早期教育领域还需要有力的研究证据来表明，学前期高质量的科学学习经验能对学业成就、科学阅读和专业成就带来长期的积极影响。在拥有这些证据之前，我们需要义不容辞地给儿童提供他们喜欢的全面的学习经验，激发他们天生的好奇心、渴求知识的欲望以及对科学问题深入的兴趣。

学前科学评价的现状

最近国家研究委员会在一篇报告*儿童早期评价：为什么，做什么，和怎么做* (Snow & Van Hemel, 2008)中把评价定义为“收集信息以做出明智的教学决策” (p.27)。教师和政策制定者希望为儿童早期科学教学做出明智的决策，但他们的能做的事情很有限，因为目前教师和科研人员可以使用的，信效度较好的评价工具很少 (另参看 Brassard & Boehm, 2007)。不管是评价科学学习的哪一方面都没有好的工具。也就是说，不管是课堂上的教师希望通过评价儿童个体的学习和技能来为学生提供个性化的指导教学，还是研究人员通过一个儿童样本来评价课程的有效性或者课程方案，或者是研究人员或管理人员通过观察来测量课堂的科学学习的环境，目前几乎鲜有综合性的考察工具。接下来，我们会提供更多的有关科学评价的具体细节，以及研究团队在这方面的前沿领域的进展。

本文的开始是对学前课堂中教师对儿童的观察和互动时所作的日常评价做一个简单的介绍，然后描述更结构化的表现性评价，教师们运用这些工具来评价儿童科学知识（以及其它入学准备内容）的建构过程。接下来是有关机构评价的讨论，重点放在一个新的标准性测量上，它可以在大规模的研究中使用，可用来评价大样本儿童在科学的入学准备方面的情况，并能对机构的长处和短处提供具体信息。最后，对学前课堂中支持儿童科学学习的质量的评价工具进行了回顾。介于课堂的整体质量和儿童入学准备结果之间的关系，我们可以推测高质量的科学学习课堂也与积极的学习结果有密切关系。当然，至于这个假设是否正确，它是一个需要实证研究来回答的问题，如果没有有效的心理测量工具来考察学习的结果和课堂教育的质量是无法回答这种问题的。

学习和知识评价

在日常互动中支持和评价科学学习

在日常生活中，学前教师需要观察儿童，并使用一种支持儿童在认知、社会，身体和情感领域发展的方式来和他们进行交流。成人通过观察并与儿童互动获得信息，然后通过活动、讨论、材料以及问题来进行回应，鼓励儿童探索并学到更多有关他们周围世界的知识。完成这个挑战需要教师理解儿童发展和在不同领域中的学习的顺序。就科学学科来说，这需要教师对儿童已有的科学各方面的基础性知识的种类，他们已经掌握的推理技巧，以及这些技能的不足了如指掌。这还意味着，为了支持儿童沿着科学学习的道路和轨迹前进，教师需要知道学习和发展的过程(Duschl et al., 2007; Gelman, Brenneman, Macdonald, & Román, 2009)。总之，在对发展的一般性理解以及个体对科学学习的需要和兴趣的理解的基础上，教师不仅仅知道教什么，还要知道怎么教。不幸的是，很多学前教育工作者坦言，自己在这个领域对科学知识的了解以及支持儿童学习的能力都很令人担心(Greenfield, Jirout, et al., 2009)。这些担忧并不奇怪，因为不管是在课堂上还是在实习中，早期教育的教师培训很少会重视科学教育(文献综述请参看 Brenneman et al., 2009)。因此，那些想帮助儿童学习科学的教师必须通过填补自己的知识空白，付出额外的时间来准备教学(Worth & Grollman, 2003)。

学前教育领域通过提供综合的和深入的职前和职后幼儿教师培训可以为儿童以及教师提供更好的服务。有关一般教师对待科学教育的态度和信念以及学前

教师对待科学教育的态度和信念的研究将会帮助我们采用聚焦的方式更好的面对这一挑战，以及加强本领域未来对学前教师的科学和教育学内容知识的研究。早期科学评价体系的关键特征是开发能够有助于更好地审视教师的知识和思考过程的工具，这样那些教师培训机构能够更好地帮助教师掌握在学前课堂中评价和支持儿童的科学学习的技能。

目前有个工具《学前教师对科学的态度和行为（P-TABS）》，是由迈阿密大学的研究人员开发的一个全新的和有效的测量幼儿教师对科学的态度和信念的工具。P-TABS 可以用来获得更清晰的关于教师对科学的观点，以及评价教师的在职培训对教师观点的影响(Maier,Greenfield, & Bulotsky-Shearer, 2011)。教育发展中心已经开发和检验了教师教学内容知识的测量工具，作为科学教学和环境量表的一部分内容（STERS，下面会有更多的描述）。这些《科学教师表现性任务》已经被用来考察教师在参加了一个集中的教师在职培训课程以后他们的科学知识方面所产生的积极的变化(Clark-Chiarelli, Gropen, Chalufour, Hoisington, Fuccillo, & Thieu, 2011)。

用于个性化教学、进步监测和课程评价的表现性评价

一种特定的专业性支持可以是儿童观察和评价框架以及对*教师在课堂中如何使用*的培训。教育测试服务中心的（ETS, n. d.）《PATHWISE 早期科学学习的理解》给教师提供了一个评价框架和策略来系统地收集和使用儿童的行为、语言和作品指导教学。PATHWISE 的作者提议“早期科学教育评价的第一个目的是帮助教师观察、记录和反思儿童对自然世界的探索”（p.1）。从这种观点来看，评价在鉴别儿童的优点和缺点方面还是次要的，更重要的是给教师作为一个观察者和儿童知识建构过程的解释者提供支持，从而能更好地支持这个过程(Chittenden & Jones, 1999)。与早期科学评价相似的一种取向同时也是建构主义课堂的重要组成部分（Edmiaston, 2002）。在这一理论导向下，评价有两个目的：记录和阐释儿童的知识 and 推理，同时评价课堂活动和教学是如何支持和阻碍学习的。

在这两种描述的方式中(Chittenden & Jones, 1999; ETS, n.d., Edmiaston, 2002)评价的过程包括在较长的一段时间内通过多种途径收集数据，找到儿童在日常课堂活动中进行科学学习的事实。这些途径包括行动、谈话，以及儿童自己制作的和在合作性小组中完成的作品。每个学生的文件夹包括教师描述的持续性行为和

对话，以及儿童自己的作品（绘画、概念网络、科学日志、雕塑、模型等等），它提供评价儿童理解能力的事实(还可参看 Gelman et al., 2009; Worth & Grollman, 2003)。对这一信息进行解释后用来指导教学，并支持新的学习。当教师在实践中练习使用这些评价程序，他们在对儿童的科学思维和行为进行观察时会变得更加熟练，也能更好地支持学前儿童的科学以及其它领域的学习和发展。

对轶事和儿童科学学习的记录也可以用于综合性的过程-监测工具，这样的工具能够丰富包括科学在内的关键性学习和发展领域，但又不局限于此。评价诸如作品样本系统(Dichtelmiller, Jablon, Marsden, & Meisels, 2001)，儿童观察记录(High/Scope Educational Research Foundation, 2003)，以及早期学习系统(Riley-Ayers, Stevenson-Garcia, Frede, & Brenneman, in press) 为追踪学生的科学学习和其它学习领域的进步提供了一种结构，教师使用档案袋来帮助他们写报告。教师使用伽利略系统(Galileo System)(Bergan, Burnham, Feld, & Bergan, 2009)，在这种系统中，他们根据观察儿童在三种不同的情境中展示出来的技能或知识判断儿童是否已经掌握了关键的入学准备的技能。同样，在完成等级评价的过程中，他们也能够从收集儿童科学学习的信息中受惠。这种评价必然给教师带来了收集数据方面的压力，但这种数据收集的目的旨在提供个别学生在科学、数学、语言、识字、社会性情感技能以及运动技能等方面的信息。儿童需要在特定领域获得更进一步的支持或者挑战，教师这样能够根据学生的需要提供更好的教学。这些评价的结果可以用来为教师和学校提供评价个体在某一时间点上的学习情况的信息，并在一段时间里继续跟踪，然后当各种信息汇总以后可以评价课程方案的目标是否已经达到。

科学学习的直接评价

为了机构评价的目的来对学习进行直接评价时，有时会利用发展心理学和教育文献中已有的比较成熟的考察任务，或者是对这些任务进行改编。范和同事(Van Egeren, Watson, & Morris, 2008)创建了一个儿童的结果性评价量表来评价“开端计划”的科学教育方案。测查方法包括事实评价、生物知识、假设性评价，以及来自发展心理学文献的心理理论(事实评价可以参看 Sodian, Zaitchik, & Carey, 1991, 生物性知识参看 Hatano & Inagaki, 1994,; 假设性评价和心理理论参看 Ruffman, Perner, Olson, & Doherty, 1993)。

用于评价《学前科学方案途径》的工具包括在发展性研究中使用相似的任务，如用于理解儿童知识的来源，或者理解他们创设一个非正式实验性检验的测查工具(参看 Gelman et al., 2009)。

对开端计划课堂中通过科学和故事进行奇妙探索活动 (MESS) 方案的评价也综合了本领域内所用方法和发展心理学文献 (如心理理论测查任务) 中的方法，评价机构在促进儿童的科学技能和概念性知识方面的有效性，如动物生命周期和防御机制 (与 S. Ellis 有个人之间的交流, August 31, 2010)。还使用了《表达和接受一个单词词汇测试》(EOWVT and ROWVT)来评价儿童语言技能。

用于考察科学开端的积极影响的评价工具。促进儿童语言发展的课程包括知名的皮博迪图片词汇测试第三版(PPVT; Dunn & Dunn, 1997)，已表明该方案的积极影响(French, 2004)。

在每个案例中，研究人员 (而不是教师) 把对学习的评价作为评价科学课程方案和干预有效性的一种方法。但是一直以来，这个领域还没有综合的评价工具来直接有效地测试儿童的科学知识内容。这种工具的缺乏已经阻碍了研究工作以及对学前科学方案和课程的评价。

很多年前，达利尔 (Daryl Greenfield) 和同事们开始开发这个工具。他们先回顾了本州早期科学的学习目标，包含了强调科学教育的学前课程，其目标是创建州和那些课程所强调的内容和过程性技能的蓝图。最初创建的测查题目反映了这些内容和过程性技能。研究通过专家的评定和预研究来选择最终的题目，并进一步保证工具建构的效度。开端计划课堂测试的结果表明，评价可以用于一系列的知识和技能，可以考察儿童在就学期间科学技能的发展，并发现词汇和学习行为成绩之间具有中等水平的正相关(Greenfield, Dominguez, et al., 2009)。

该团队的持续性工作涉及设计和使用了 80 个题目的测试工具来评价儿童早期动手操作科学 (ECHOS) 教师专业发展和课程方案对儿童的科学学习和其它领域的学习的影响。此外，最初的科学评价工具 (卡片版) 将作为设计电子版工具“*聚焦科学*”的一个基础。最终完成的详尽的心理测量评价希望能够用于国内的研究和机构评价(Greenfield, Dominguez, Greenberg, Fuccillo, & Maier, 2011)。这种评价将帮助州、学区，或者其他教育团体了解儿童在即将进入小学学前班 (5-6 岁儿童) 时的科学学习大致处于什么水平，这样可以有助于为这些儿童制定方案、课程和教学改变的决策，最终促进和改善学习。

对科学相关技能和学习倾向性的评价

当然，儿童其它领域的发展影响着科学学习，同时也会受到科学学习的影响。因此，我们可以合理地去审视儿童学习和发展的其它重要领域，看一下相关的技能和知识。例如，社会技能对科学探究会产生影响，因为儿童在学校参与这种探究时必须学会在科学讨论中分享和呈现自己的观点，尊重别人的观点，在小组实验或者探究中与同伴和成人合作。事实上，在州学习标准和学前课程的文献中，格林菲德和同事们(Greenfield, Jirout, et al., 2009)把合作作为 8 个关键探究技能之一。

同样，儿童个体的学习品质——包括主动性、动机、坚持性和好奇心——都会影响自发的探索。国家教育目标委员会（1995）把学习品质作为儿童学习和发展的一个关键领域，在最近一期的国家研究咨询委员会（NRC）发表的早期儿童评价分册中(Snow & Van Hemel, 2008)可以看到，学习品质是详细描述的评价领域之一。读者可以从中了解有关这一领域和评价的综合讨论，而本篇文章的目的之一是对那些尚未进入现有文献中的内容进行介绍。杰米·吉瑞特(Jamie Jirout)和大卫·克莱厄(David Klahr) (2010a, 2010b)在这方面做了研究，他们开发和检验了儿童科学好奇心的测量工具。

吉瑞特开发了一种类似游戏的工具，它把信息收集情境作为评价个体学习者好奇心水平的一种方式，以此控制不确定性和含糊性。目前使用的计算机版本为水下探索游戏！在目前的情境中，儿童可以重复确认已知信息（也就是说，在一定程度上没有或者只有很低的不确定性），在中等水平的不确定情境下探索（如很少有鱼会出现在窗子上），或者在非常不确定情境下探索（任何鱼都可能会出现）。这个游戏提供了有关个别儿童对不确定性程度的喜好的具体信息。也就是说，研究人员可以根据儿童的选择来评价她或他在这个情境中的舒适度，这样正确的答案或多或少是确定的。行为评价与学前学习行为量表之间不同的等级呈现出正相关，包括学习能力动机、注意力/坚持性、学习态度以及量表的总分(Jirout & Klahr, 2010a, 2010b)。吉瑞特设计 3-6 岁儿童好奇心测量工具的动机来自于一个事实，那就是常常把“好奇心”作为入学准备的倾向性方面内容，但是该领域目前既没有统一的定义，也没有有效的心理测量工具。这个工具有助于我们评价教育

机构对支持和促进儿童的好奇心的水平，它能激发和促进儿童的探索行为并引发更好的学习(Jirout & Klahr, 2010b)。

总之，测量儿童个体的科学学习可以采用不同的形式，形式的选择应该以信息收集为目的。教师观察、倾听和提问以更好的评价儿童在日常课堂活动中某一时刻的想法和理解。表现性评价工具根据儿童的持续性行为和在全自发的和半结构情境下完成的工作样本提供有关儿童学习的形成性评价，它有助于教师设计新的学习经验，为学生在科学和其他领域的学习提供更好的支持和挑战。除了对学生个体的发展进步进行监控，评价的信息还可以用于评价某一课程方案与儿童科学学习和发展之间的相关（注意这一信息的效度，或者对方案之间进行比较只限于那些受过正式培训并达到较好信度的教师，而且应该定期检查以保证评价程序的正确，这样信息得以在不同的班级和方案之间进行比较）。

迈阿密大学直接评价的纸质工具和即将出版的计算机版本的*聚焦科学*是一个标准化的测量工具，适合评价教育方案的优缺点以及儿童入小学学前班时所做的准备。

与科学学习相关的课堂质量测量

如果儿童的科学入学准备得到了改善，那么在早期开端计划指南中就不会出现平坦的增长曲线和发展最慢的曲线了(Greenfield, Jirout, et al., 2009)，这样，对学习者的评价就会变得很重要。所以，对作为科学学习者成长和发展的环境的评价同样重要。为了提高成绩，教师和政策制定者需要知道哪些内容和课堂互动能促进更好的学习。环境质量的测量可以从多方面来促进这个努力。教师和行政管理人员使用一个结构性观察工具描述高质量的学习环境的特点，评价他们的方案中与该工具描述的相关指标，或者使用相同的工具来评价其它方案。这些评价工具可以用来明确哪些领域需要提高，或用来指导教师的专业发展。

课堂质量的测量可以使用多个时间点来监测付出的努力和表明方案和环境质量已经得到改善的途径。在这些案例中，结构性观察应由外来观察人员完成，而不是本班的教师。另一种课堂质量的测量工具将包括教师的自我评价（也可以是由导师或带教老师来评价），用来指导和改善他们与儿童的教学互动(Frede, Stevenson-Garcia, & Brenneman, 2010)。最后，如果实施过程能保证心理测量的有

效性和信度的话，课堂质量测量工具可以用于机构的绩效考核目的(Snow & Van Hemel, 2008)。

我们考察儿童成果评价工具的使用情况发现，可用于支持科学学习的课堂质量测量工具较少。一个工作团队回顾了用于评价从学前一直到三年级数学和科学教学支持的工具，得出的结论为这两个领域可用的工具都很少，特别是科学尤其少得可怜(Brenneman et al., in press)。这个发现得到了其他研究人员的肯定(Greenfield, Jirout, et al., 2009; Snow & Van Hemel, 2008)。最近斯诺和赫默尔(Snow & Van Hemel, 2008) 总结认为，目前的课堂环境观察工具对学习环境的评价都只处于一般水平，仅仅只有少部分工具能够用来对与认知和学业技能有关的领域如科学进行较好的评价。本文下面一部分将勾画出目前已有的，尚处于不同完成阶段的评价工具。

幼儿环境等级评价——修订版 (ECERS-R)

《幼儿环境等级评价-修订版》(ECERS-R; Harms, Clifford, & Cryer, 2005)是原先的《幼儿环境等级评价—修订版》(ECERS-E)的进一步延伸，旧版本总体上缺乏对早期阅读、数学、科学和文化多样性的关注(Sylva, Siraj-Blatchford, & Taggart, 2003)。《幼儿环境等级评价—修订版》在测量课堂科学的支持方面比其他已经发表的广泛使用的课堂观察工具更全面。观察评定人员需要评价两个项目：科学内容的呈现以及和教室里是否设置了科学活动区。观察人员还需要在三个科学活动/科学过程性项目（无生命的，有生命的过程以及我们周围的世界，食物准备）中选择其中一个做得最好的项目进行评分。考虑到观察评定人员在班级中观察的时间很有限，不能观察科学活动的整个过程，这种方式也许能提供解决这个问题。但是，有的重要的科学领域仍然因为工具中没有相关内容或由于在观察时另一方面内容呈现的更多而没有得到关注和评价。《幼儿环境等级评价—修订版》的心理测量特征为内在相关信度超过 0.88，加权卡方系数(Weighted Kappa Coefficient) 范围是 0.83 到 0.97，共同效度值也很高(0.87)。该工具的总分平均值与儿童的早期阅读、非言语推理、以及早期数概念之间有显著正相关。科学等级量表本身没有表现出与儿童成绩的显著相关（参见 Halle & Vick, 2007; Sylva et al., 2003）。

《科学教学和环境等级量表 (STERS)》和《学前科学和数学评定量表 (PRISM)》

由教育发展中心 (EDC) (Chalufour, Worth, & Clark-Chiarelli, 2006)和国家早期教育研究所 (NIEER) (Stevenson-Garcia, Brenneman, Frede, & Weber, 2010)两个部门开发的两种工具可用于评价更全面的科学内容、概念和推理能力。教育发展中心的测量工具是《科学教学和环境等级量表》(STERS), 目的是为了测量课堂科学教学质量的变化, 用于评价专业发展干预的有效程度。《科学教学和环境等级量表》使用课堂观察和教师访谈来评定教学人员: (1) 创设有利于探究和学习的物理环境; (2) 提供直接经验促进概念学习, (3) 促进科学式探究的使用, (4) 创设合作的氛围促进探索和理解, (5) 提供深入对话的机会, (6) 积累学生的词汇, (7) 制定深入调查的计划, (8) 评价学生的学习。每个部分都使用 4 个标记 (1=较差, 4=非常好) 来评定内容和互动, 并能在课堂中发现与每一个水平相对应的情况。作者报告了该工具的内部一致性很高 (Cronbach's $\alpha = .96$), 接下来对其心理测量工具的特性的调查还会继续(Clark-Chiarelli, Gropen, Chalufour, Hoisington, Fuccillo, & Thieu, 2011)。

国家早期教育研究所开发的《学前科学和数学评定工具》(PRISM) 是一个综合的、包含 16 个项目的工具, 用于测量支持数学和科学发展的课堂教学内容和教学互动的情况。科学考察项目聚焦支持生物和非生物科学的探索, 鼓励科学相关的读、写和表征, 支持对科学概念的考察和讨论、观察、预测、比较和对比, 鼓励用日记、图表和其他表征形式来记录科学信息的内容和教学互动。另外, 测量和分类项目内容跨越科学和数学学科两个领域。该工具计划进行全面有效的研究并继续进行因素结构方面的探索。初步的分析表明内部一致性还可以 (Cronbach's $\alpha = .78$), 与 ECERS-R 的共同效度值为中等水平($R = .41$) (Brenneman, Stevenson-Garcia, Frede, & Jung, 2011)。

未来的研究方向

正如文章序言所说, 目前教育政策制定人员、美国联邦和州政府、工业行政高管、课程开发人员以及研究人员都对学前教育科学、技术、工程和数学的学习表现出强烈的兴趣。很好地利用这种兴趣并把它转化为明确的教育政策和实践建

议，需要有能为幼儿提供积极的学习经验的教学环境和互动方面的有力的研究证据。

这种证据应该通过有研究并有实证检验的工具来获得，以保证这些证据的有效性和可信度，并能够与儿童的学习结果有联系。根据我和我的同事们对这种评价工具现有情况的探究，我们很明显感到在这个领域急需这种工具。

运用一种综合方式来考察科学学习课堂支持的工具应能当作一种客观的测量工具来使用，能使用相同的尺度来比较课堂、课程和方案，并允许我们使用严格的方式来进行评价，以及回答在序言中提出的问题：我们设计的正式和非正式的学前科学教育方案能有效地达到我们为儿童制订的学习目标吗？其中有的方案比其它方案更有效吗？在科学教育取向的课堂上有哪些典型的材料和教学互动？我们如何在有限的教育资源的情况下获得最大的“成功”？我们如何保证每个儿童都有适宜的学习机会来建立和延续他们的惊喜、热情、内容知识以及推理技能，并在科学活动的尝试中表现出来？

结论

在学前教育领域中正确使用评价引起了很多人的关注，他们担心有些评价会对幼儿产生负面的影响。他们担心在测试中，儿童可能会体验到自己不太完美、疑惑的感觉，或感到有压力或者厌倦。这里所描述的评价包括在典型的学前教育机构一日活动中，利用作为儿童日常生活经验一部分的工作样本、对话和活动。其他评价也许还需要利用儿童在一日生活中的时间，但这些活动通常设计成游戏，而且儿童会感到很有趣。幼儿评价还有其它方面的担忧，如数据来自学前儿童，他们的表现可能比大年龄儿童的个别差异更大，而他们也不知道表现良好的“重要性”，这种数据却要用来指导与机构和学校教育的有效性有关的，具有高利益相关的决定。评价的内容越是一般，工具只用于它们原来所设计的目的就越重要。对这一问题的回顾可以参看近来的文献(Brassard & Boehm, 2007; Snow & Van Hemel, 2008)。

本文绝对不会忽略这些重要的担忧，因为对学前科学教育的评价并不是可有可无的。但是问题不是我们是否需要评价科学学习，而是我们如何做才能解决教师、管理人员、研究人员以及政策制定人员提出的问题；所有这些人人都把这种评

价看成是有用的工具；而且评价工具应尽可能地与被评价的儿童的生活无缝对接。

尽管早期科学这个领域目前缺乏好的评价工具，但在评价工具的数量和幼教专业人员支持幼儿科学学习方面仍取得了进步。但这方面的进步并没有反映在现有的文献中。因此，这篇文章的目的之一是对目前早期科学评价工具的状况展开讨论，希望在我们整理出来的评价工具清单的基础上大家不断添砖加瓦。总之，研究科学学习的人和研究幼儿科学教育的人应该协力为评价工具的建设构造蓝图。如果我们齐心协力支持今天的学前儿童学习科学，明天的他们或将成为学生、公民、以及科学、技术、工程和数学界的专业人士。

致谢

本文原先是为 2010 年在北爱荷华大学举行的科学、技术、工程和数学早期教育和发展 (SEED) 大会准备，该大会承蒙 Betty Zan 女士组织。非常感谢 Betty, Ingrid Chalufour, Nancy Clark-Chiarelli, Daryl Greenfield, Michele Maier 以及 Jamie Jirout 分享了他们的研究细节。谢谢 Ellen Frede 和匿名审稿人对文章初稿的点评。他们的慷慨和建议提高了论文的质量；同时声明，本文的任何错误责任在我自己。

【编辑注：其他关于 SEED 会议的论文，参考 *ECRP* 杂志的本期之外的信息。】

参考文献

American Honda Foundation. (2010). Retrieved April 3, 2010, from <http://corporate.honda.com/america/philanthropy.aspx?id=ahf>

Barack Obama and Joe Biden's plan for lifetime success through education. (n.d.). Retrieved March 22, 2011, from <http://www.barackobama.com/pdf/issues/PreK-12EducationFactSheet.pdf>

Barnett, W. Steven. (2008). *Preschool education and its lasting effects: Research and policy implications*. Boulder, CO, & Tempe, AZ: Education and Public Interest Center & Education Policy Research Unit. Retrieved December 11, 2008, from <http://epicpolicy.org/publication/preschool-education> Editor's note: This url has changed:<http://nepc.colorado.edu/publication/preschool-education>

Beerling, Steven C. (2009). *Actions to improve science, technology, engineering, and mathematics (STEM) education for all American students* [Letter from the National Science Board STEM education outlining recommendations for the President-Elect

Obama administration]. Retrieved January 1, 2009, from http://www.nsf.gov/nsb/publications/2009/01_10_stem_rec_obama.pdf

Bergan, John Richard; Burnham, Christine Guerrero; Feld, Jason K.; & Bergan, John Robert. (2009). *The Galileo Pre-K online system for the electronic management of learning*. Tucson, AZ: Assessment Technology. Retrieved March 22, 2011, from <http://www.ati-online.com/galileopreschool/PreWelcomeTechManual.html>

Bowman, Barbara T.; Donovan, M. Suzanne; & Burns, M. Susan (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Brassard, Marla R., & Boehm, Ann E. (2007). *Preschool assessment: Principles and practices*. New York: Guilford Press.

Brenneman, Kimberly; Boller, Kimberly; Atkins-Burnett, Sally; Stipek, Deborah; Forry, Nicki; Ertle, Barbina; et al. (in press). Measuring the quality of early childhood math and science curricula and teaching. In Martha Zaslow, Ivelisse Martinez-Beck, Kathryn Tout, & Tamara Halle (Eds.), *Quality measurement in early childhood settings*. Baltimore, MD: Paul H. Brookes.

Brenneman, Kimberly; Stevenson-Boyd, Judi; & Frede, Ellen C. (2009). Math and science in preschool: Policies and practice. *Preschool Policy Brief, 19*. New Brunswick, NJ: National Institute for Early Education Research.

Brenneman, Kimberly; Stevenson-Garcia, Judi; Frede, Ellen; & Jung, Kwanghee. (2011, April). *Assessing instructional quality for science and math: Preschool Rating Instrument for Science and Mathematics (PRISM)*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

Burchinal, Margaret; Howes, Carollee; Pianta, Robert; Bryant, Donna; Early, Diane;

Clifford, Richard; et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*(3), 140-153.

Callanan, Maureen A., & Oakes, Lisa M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7*(2), 213-233.

Chalufour, Ingrid; Worth, Karen; & Clark-Chiarelli, Nancy. (2006). Science teaching environment rating scale (STERS). Newton, MA: Education Development Center.

Chittenden, Edward, & Jones, Jacqueline. (1999). Science assessment in early childhood programs. In *Dialogue on early childhood science, mathematics, and*

technology education. Washington, DC: Project 2061, American Association for the Advancement of Science.

Clark-Chiarelli, Nancy; Gropen, Jess; Chalufour, Ingrid; Hoisington, Cynthia; Fuccillo, Janna; & Thieu, Yen. (2011, April). *Assessing teacher and classroom outcomes from a professional development program in science for preschool teachers*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

Committee for Economic Development. (2006). *The economic promise of investing in high-quality preschool: Using early education to improve economic growth and the fiscal sustainability of states and the nation*. Washington, DC: Author.

Dichtelmiller, Margo L.; Jablon, Judy R.; Marsden, Dorothea B.; & Meisels, Samuel J. (2001). *The work sampling system omnibus guidelines: Preschool through third grade* (4th ed.). New York: Rebus.

Dunn, Lloyd M., & Dunn, Leota M. (1997). *Peabody picture vocabulary test, third edition (PPVT-III)*. Circle Pines, MN: American Guidance Service.

Duschl, Richard A.; Schweingruber, Heidi A.; & Shouse, Andrew W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press; Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education.

Edmiaston, Rebecca. (2002). Assessing and documenting learning in constructivist classrooms. In Rheta DeVries, Betty Zan, Carolyn Hildebrandt, Rebecca Edmiaston, & Christina Sales (Eds.), *Developing constructivist early childhood curriculum: Practical principles and activities* (pp. 53-67). New York: Teachers College Press. Educational Testing Service. (n.d.). *PATHWISE understanding early science learning*. Princeton, NJ: Author.

Frede, Ellen; Stevenson-Garcia, J., & Brenneman, Kimberly. (2010). *Self-evaluation for science and math education (SESAME)*. New Brunswick, NJ: Author.

French, Lucia. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, 19(1), 138-149.

Gelman, Rochel; Brenneman, Kimberly; Macdonald, Gay; & Román, Moisés. (2009). *Preschool pathways to science (PrePS): Facilitating scientific ways of thinking, talking, doing, and understanding*. Baltimore, MD: Paul H. Brookes.

Greenfield, Daryl B.; Dominguez, Ximena; Greenberg, Ariela; Fuccillo, Janna; & Maier, Michelle F. (2011, April). *Assessing science readiness in young low-income*

preschool children. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal.

Greenfield, Daryl B.; Dominguez, Ximena; Fuccillo, Janna M.; Maier, Michelle F.; Greenberg, Ariela C.; & Penfield, Randall. (2009, April). *Development of an IRT-based direct assessment of preschool science*. Paper presented at the biennial meeting of the Society for Research in Child Development, Denver, CO.

Greenfield, Daryl B.; Jirout, Jamie; Dominguez, Ximena; Greenberg, Ariela; Maier,

Michelle; & Fuccillo, Janna. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Education and Development, 20*(2), 238-264.

Haahr, Jens Henrik (with Nielsen, Thomas Kibak; Hansen, Martin Eggert; & Jakobsen, Søren Teglegaard). (2005). *Explaining student performance: Evidence from the international PISA, TIMMS, and PIRLS surveys*. Danish Technological Institute.

Retrieved December 12, 2008, from

<http://www.pisa.oecd.org/dataoecd/5/45/35920726.pdf>

Halle, Tamara, & Vick, Jessica. (2007). *Quality in early childhood care and education settings: A compendium of measures*. Washington, DC: Office of Planning, Research, and Evaluation. Administration for Children and Families, U.S. Department of Health and Human Services.

Harms, Thelma; Clifford, Richard M.; & Cryer, Debby. (2005). *Early childhood environment rating scale* (Rev. ed.). New York: Teachers College Press.

Hatano, Giyoo, & Inagaki, Kayoko. (1994). Young children's naive theory of biology. *Cognition, 50*, 171-188.

HighScope Educational Research Foundation. (2003). *The preschool child observation record (COR)*. Ypsilanti, MI: HighScope Press.

Jirout, Jamie, & Klahr, David. (2010a, June). *Assessing and validating measures of curiosity in preschool children*. Poster presented at the Fifth Annual Institute of Education Sciences Research Conference, National Harbor, MD.

Jirout, Jamie, & Klahr, David. (2010b). *The measurement of curiosity: A review of the literature*. Manuscript in preparation.

Maier, Michelle F.; Greenfield, Daryl B.; & Bulotsky-Shearer, Rebecca J. (2011, April). *Development and initial validation of a preschool teachers' attitudes and beliefs toward science questionnaire*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal

Melhuish, Edward C.; Sylva, Kathy; Sammons, Pam; Siraj-Blatchford, Iram; Taggart, Brenda; Phan, Mai B.; et al. (2008). Preschool influences on mathematics achievement. *Science*, 321(5893), 1161-1162. Retrieved March 24, 2011, from <http://www.sciencemag.org/cgi/content/full/321/5893/1161> Editor's note: This url has changed:<http://www.sciencemag.org/content/321/5893/1161.full>

Motorola Foundation. (n.d.). *Innovation generation grants*. Retrieved March 22, 2011, from <http://responsibility.motorola.com/index.php/society/comminvest/education/igg/>
National Association for the Education of Young Children. (n.d.). *Standard 2: NAEYC accreditation criteria for curriculum*. Retrieved March 22, 2011, from <https://oldweb.naeyc.org/academy/standards/standard2/>

National Education Goals Panel. (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Washington, DC: Author.

National Research Council. (2005). *Mathematical and scientific development in early childhood: A workshop summary*. Washington, DC: National Academies Press.

PNC Foundation. (2009). *PNC Grow Up Great collaborates with 14 science centers to boost science skills of preschoolers*. Downloaded April 3, 2010, from http://www.pncgrowupgreat.com/media/downloads/PNC_GrowUpGreat_SciencePress_Release.pdf

Riechard, Donald E. (1973). A decade of preschool science: Promises, problems, and perspectives. *Science Education*, 57(4), 437-451.

Riley-Ayers, Shannon; Stevenson-Garcia, Judi; Frede, Ellen; & Brenneman, Kimberly. (in press). *Early learning scale*. Anaheim, CA: Lakeshore Learning.

Ruffman, Ted; Perner, Josef; Olson, David R.; & Doherty, Martin. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis-evidence relation. *Child Development*, 64(6), 1617-1636.

Scott-Little, Catherine; Lesko, Jim; Martella, Jana; & Milburn, Penny. (2007). Early learning standards: Results from a national survey to document trends in state-level policies and practices. *Early Childhood Research & Practice*, 9(1). Retrieved March 22, 2011, from <http://ecrp.uiuc.edu/v9n1/little.html>

Shonkoff, Jack P., & Phillips, Deborah A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Snow, Catherine E., & Van Hemel, Susan B. (Eds.). (2008). *Early childhood assessment: Why, what, and how*. Washington, DC: National Academies Press.

Sodian, Beate; Zaitchik, Deborah; & Carey, Susan. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, 62(4), 753-766.

Stevenson-Garcia, Judi; Brenneman, Kimberly; Frede, Ellen; & Weber, Marcie. (2010). *Preschool rating instrument for science and mathematics (PRISM)*. New Brunswick, NJ: National Institute for Early Education Research.

Sylva, Kathy; Siraj-Blatchford, Iram; & Taggart, Brenda. (2003). *Assessing quality in the early years: Early childhood environmental rating scale extension (ECERS-E)*. Stoke-on-Trent, UK: Trentham Books.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families/Head Start Bureau. (2003). *The Head Start path to positive child outcomes*. Washington, DC: Author. Retrieved December 12, 2008, from <http://www.ecechicago.org/resources/pdfs/headstart01.pdf>

Van Egeren, Laurie A.; Watson, Dyane P.; & Morris, Bradley J. (2008, June). *Head start on science: The impact of an early childhood science curriculum*. Paper presented at the Ninth National Head Start Research Conference, Washington, DC.

Worth, Karen, & Grollman, Sharon. (2003). *Worms, shadows, and whirlpools: Science in the early childhood classroom*. Portsmouth, NH: Heinemann.

作者信息

Kimberly Brenneman 研究学前课堂中早期科学和数学学习和支持。她在国家早期教育研究院（NIEER）的工作包括教学质量评价的设计和效度检验，以及编写促进教师的科学和数学教学的专业发展培训教材。Dr. Brenneman，是《学前科学的途径（PrePS）：提供科学性的方式来思考、谈论、做和理解》一书的作者。她同时还作为美国公共电视频道（PBS）“科学儿童”节目的教育咨询员，该节目旨在促进探索、发现以及儿童的科学准备。

联系地址

Kimberly Brenneman, Ph.D.
National Institute for Early Education Research
120 Albany St.

Suite 500
New Brunswick, NJ 08901
Telephone: 732-932-4350, ext. 239
Fax: 732-932-4360
Email: kbrenneman@nieer.org

译者：澳大利亚纽卡斯尔大学教育学院 刘婷
审校：中国华东师范大学学前教育系 周欣