

PROSPECTS

教育展望

国际比较教育季刊

134
专栏

在不同的情境脉络中 评价学生的成绩

特邀主编

SARAH HOWIE

TJEERD PLOMP



联合国教科文组织国际教育局

Vol. XXXV, no. 2, JUNE 2005



上海教育出版社
SHANGHAI EDUCATIONAL PUBLISHING HOUSE

图书在版编目(CIP)数据

教育展望. 134 / 联合国教科文组织国际教育局编著;
华东师范大学译. —上海: 上海教育出版社, 2007. 1
ISBN 978-7-5444-1135-6

I. 教... II. ①联...②华... III. 教育—世界—丛
刊 IV. G51-55

中国版本图书馆 CIP 数据核字(2006)第 164938 号

教育展望

国际比较教育季刊 总第 134 期

在不同的情境脉络中评价学生的成绩

联合国教科文组织国际教育局

上海世纪出版股份有限公司

出版发行

上海教育出版社

易文网: www.ewen.cc

(上海永福路 123 号 邮编: 200031)

各地新华书店经销 江苏启东市人民印刷有限公司印刷

开本 787×1092 1/16 印张 7.75

2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷

ISBN 978-7-5444-1135-6/G·0936 定价: 30.00 元

(如发生质量问题,读者可向工厂调换)

本刊所载文章的观点及材料,由作者自行负责,不代表联合国教科文组织国际教育局,文章中所用名称及材料的编写方式并不意味着联合国教科文组织国际教育局对于任何国家、领土、城市或地区或其当局的法律地位或对于其边界的划分表示任何意见。

一切信件均应寄:

Editor, Prospects,
International Bureau of Education,
P.O. Box 199,
1211 Geneva 20,
Switzerland.

欲了解国际教育局的计划、活动及出版物,
请查询其因特网主页:

<http://www.ibe.unesco.org>

一切订阅刊物的来信应寄:

Springer,
P.O. Box 322, 3300 All Dordrecht,
The Netherlands

联合国教科文组织(UNESCO)

国际教育局(IBE)2005年 ISSN: 0033-1538

《教育展望》编委会

编委会主任

Cecilia Braslavsky

编委会成员

Cesar Birzea, Norberto Bottani
Mark Bray, Antonio Guerra Caraballo
Lawrence D. Carrington, Elie Jouen
Kenneth King, P. T. M. Marope
Mamadou Ndoeye, Fernando Reimers
Bikas C. Sanyal, Buddy J. Wentworth
Yassen N. Zassoursky, Muju Zhu

助理主编

Carla Borges

助理编辑

Brigitte Deluermoz

中文版编委会

主 编

王建磐

副主编

杜 越 庄辉明 周南照 钟启泉

编 委

(以姓氏笔画为序)

王建磐 王斌华 冯大鸣 左焕琪
任友群 庄辉明 李月娥(Molly Lee)
杜 越 郑太年 范国睿 周南照
郑燕祥 赵中建 钟启泉 祝智庭
程介明 董建红

编辑部主任

任友群

副主任

冯大鸣 郑太年(常务)

编辑部地址

上海市中山北路 3663 号
华东师范大学《教育展望》中文版编辑部

2005 年主题

1. 在不同的情境脉络中评价学生的成绩(1)
2. 在不同的情境脉络中评价学生的成绩(2)
3. 职业技术教育和可持续发展培训
4. 教育政策对话:来自非洲国家的经验

第 134 期

教育展望

国际比较教育季刊

第 35 卷, 2005 年 6 月第 2 期

中文版 2005 年第 2 期(总第 134 期)

目 录

编者按	<i>Pierre Luisoni</i>	1
观点 / 争鸣		
通过快速学习评价监测基本技能的获得: 来自秘鲁的个案研究	<i>Helen Abadzi, Luis Crouch, Marcela Echegaray, Consuelo Pasco & Jessyca Sampe</i>	5
评价学生的成绩		
专栏导言	<i>Sarah Howie & Tjeerd Plomp</i>	21
第三部分: 学生成绩的国际比较研究		
学生特点和家庭背景对伊朗学生数学成绩的影响	<i>Ali Reza Kiamanesh</i>	24
系统层面的评价:	<i>Sarah Howie</i>	36
影响数学成绩的语言及其他背景因素		
学生成绩的国际比较:	<i>Mojca Štraus</i>	47
斯洛文尼亚教育政策制定的指标		
TIMSS-R 在 18 个发展中国家是如何对教育产生 作用的	<i>Warwick B. Elley</i>	57
第四部分: 学生评价标准		
学生评价标准: 学生评价的范式转换	<i>Arlen R. Gullickson</i>	69
动向/案例		
TIMSS, PISA, PIRLS 和世界社会中的低教育成就	<i>Jens Naumann</i>	83
中国教育发展与研究		
高校招生考试制度变革的趋势: 走向多元化	张民选 高耀明	99
我国高等学校的中外合作办学: 政策、实践与展望	董秀华 朱益明	107
引领教师专业成长, 提升学校管理品位, 为学校 持续发展提供有力支撑	倪振民	116

获悉联合国教科文组织国际教育局主席 Cecilia Braslavsky 逝世,我非常悲痛。在过去几个月中,她以惊人的毅力和勇气忍受着疾病的折磨,这值得我们付出所有的尊敬。Cecilia 是一位杰出的、坚定的教育家,全身心地投入到她的国家(阿根廷)、她的地区(拉丁美洲)所面临的——实际上是全世界面临的——教育问题的解决。作为一位理论家、实践者、研究者、顾问、教师和管理者,Cecilia 以卓越的精力、智力和热情在教育领域中工作着。她对教育和教育的重要性充满激情,并将这一点与她惯常的热情融合在一起。她的全家深深地怀念她,我对他们致以我最深切的哀悼。全球无数的朋友和同仁都沉痛地怀念她。她的逝世,让联合国教科文组织和教育界失去了一位我们时代最重要的教育家。这是一个巨大的损失。

Koichiro Matsuura

联合国教科文组织总干事

按照惯例,这篇编者按本应由 Cecilia Braslavsky 撰写。直到她生命的最后一周,她还相信她有力量做这件事。但是在 6 月 1 日早晨,在以超乎寻常的勇气同可怕的癌症进行了七个月的艰辛斗争之后,她永远地离开了。她的逝去,使《教育展望》的读者无缘读到她最后的讯息。

Cecilia Braslavsky 将她全部的生命贡献给全球教育的推进。尽管在这条路上充满了重重障碍和种种考验,甚至是严峻的考验,但她还是成功地将国际教育局建设成为一个充满生机的、日益强大的机构。不幸的是,在她可以开始收获源自她卓越贡献和无比勇气的成果的时候,她却离开了我们。我们将记住她非凡的个性、高水平的专业地位和无瑕的品格。她愿以精力、勇气和希望奋斗到生命的最后一刻,这一切让我们深深钦佩。为表示我们的敬意,2005 年 12 月份的《教育展望》将全部献给她,纪念她的专业生活和智力贡献,纪念她在全球推进优质教育中作出的勇敢的、富有远见的奋斗和表现出的无瑕品格。

为了纪念这位伟大的女教育家,我们将继续她的工作、她的事业,永不放弃。报答她为教育作出的贡献的最好方法,就是将我们的努力融入她深深信仰的事业:有质量的全民教育(quality education for all)。

Cecilia Braslavsky 选择将评价问题作为这两期《教育展望》的主题,因为它是教师和家长关心的重大问题,更是教育决策者关心的重大问题。本期是这个系列的第二期,专栏中的文章分析了多种评价学生成绩的方法,这些方法是测量走向全球优质教育系统进展状况的工具。特别是,上期《教育展望》推动人们探讨如何在国家和国际

评价中考虑质量问题,本期不但遵循了这一思路,还进一步提供了一些线索,展现了国际上认可的评价,即 TIMSS,如何反映在国家政策中,又如何服务于有质量的全民教育的改善。

本期“观点/争鸣”栏目的文章,反映了国际和国家的专家在努力提高贫穷国家的全国评价方法的范围和结果质量方面所取得的成果。鉴于这类国家在对其学生和整个教育系统进行评价时会面临无数困难,就迫切需要找到能低成本地评价更多学生的策略,特别是评价那些生活在贫困条件下的学生。秘鲁的一个实验性测试向我们展示了满足这一需要的一个革新性策略,因而,本部分的作者对此进行了详细的描述。

如前所述,Sarah Howie 和 Tjeerd Plomp 主编的专栏中的文章探讨了对不同国家和地区进行国际比较评价的效果。上一期(第 133 期)的系列文章重点在于国际学生评价项目(PISA)以及南非和东非教育质量监控联盟(SAQMEC),后者在 2004 年召开的第 47 届国际教育大会上被授予夸美纽斯奖章。本期将详细分析国际教育成就评价协会组织的第三次国际数学和科学学习研究(TIMSS)的影响。更具体地说,本专栏将概述这类评价创新能够如何贡献于现在和未来的国家教育系统改革。

尽管分析的视角和层面不同,但所有的专栏作者都指明了国际比较评价工具的重要性,也指出了会遇到的困难,提出了主要的行动原则。这种方法为教育系统提供的投入可以是多方面的,无论是对当地而言还是在结构层面上。

对当地而言,系统化的评价活动首先能够识别面临的挑战,其次能够突出多方合作的可能性。而且,在没有可靠的或系统的数据监控教育质量的国家,这种研究能够填补国家留下的空白,如 Ali Reza Kiamanesh 有关伊朗教育系统的文章所示。评价研究对于国家评价项目的开发来说也可以作为一个参照。换言之,国家可以根据自己的最迫切需要强调某一个标准,或者在国际比较研究所测试的诸多技能中增加新的项目。从这一意义上说,Sarah Howie 对于南非评价体系研究的介绍非常能说明问题。最后,Mojca Štraus 撰写的关于斯洛文尼亚的案例努力揭示像 TIMSS 这样的国际测试如何与国家的决策过程相关,直接的影响是如何发生的。

从结构层面看,国际评价的努力可以作为一个工具,用于描绘有质量教育成就现状的广阔图景,并识别走向《达喀尔行动框架》确定的、《千年发展目标》再次号召的国际认可目标的趋势。这些努力通过采用同样的方法和标准来评价多个国家,为比较研究奠定了基础,其中包括这些国家之间的比较研究,也包括不同时期的比较研究。同样,运用这类测试也可以系统研究它们对于被测试国家的影响。这正是 Warwick B. Elley 研究的中心目标。他分析了 TIMSS-R 对 18 个发展中国家的影响,这些国家接受了资助以参与这项评价研究。

不过,对于应当评价何种能力和技能以及如何评价仍有许多疑虑。尽管这一争论不像在开始成为教育体系议程中最优先事项之一时那样新奇,但是,从界定关键概

念,到在最能满足评价需要的方法上达成一致意见,依旧存在困难。在这方面,Arlen R. Gullickson 尽力弄清并描述先前和现在的学生评价标准,同时检视了它们的效能,并提出了改革的需要。

“动向/案例”部分回顾了“观点/争鸣”栏目中提到的关于贫困人口的评价和教育需要问题的争论。这一直是 Cecilia Braslavsky 研究工作的焦点。她和 Jens Naumann 都认同这一观点:大幅增加双边的和多边的发展援助,对于满足贫困人口的现存需求是必不可少的。她也相信,满足发展中国家需求的最佳途径是合作,不仅仅是政府和国际团体间的合作,也包括国家内部不同机构间的合作。

Pierre Luisoni

国际教育局(IBE)代主席

(郑太年 译)

通过快速学习评价监测基本技能的获得： 来自秘鲁的个案研究

Helen Abadzi, Luis Crouch, Marcela Echegaray,
Consuelo Pasco & Jessyca Sampe*

有人曾经对秘鲁中等以下收入的贫困家庭学生做过一项调查,分别要求一年级和二年级学生大声朗读一段短文,然后回答几个简单的理解性问题。结果发现,虽然整个学年即将结束,但一年级只有25%的学生能读出短文中的一两个单词,75%的学生则根本不能阅读;二年级也只有41%的学生能读出个别单词;对三个理解性问题都能正确回答的孩子平均每分钟阅读77个单词,而只能回答一个或两个问题的孩子的阅读速度则慢得多,分别为每分钟15个和41个单词。由此得出的阅读速度与理解力之间的关系与对阅读自动化的记忆研究所得出的结论是一致的,这也再次验证了其他一些观察性研究得出的结论:在很多国家的贫困学生中存在着大量文盲。因此我们认为,世界各国(尤其是采用音标拼写语言的国家)应采取行动,保证所有儿童在读完二年级以后就能顺畅流利地阅读。在这里,我们通过一种操作方便、成本较低的快速阅读检测法,使所收集到的学生表现信息既便于理解,又具有操作性。如果取样合理,操作准确,那么通过这种方法收集基本资料和监控数据不仅当地能够支持,而且在世界范围内也具有可比性。

背 景

自从实行全民教育计划以来,大量的教育投资和一些国家教育制度的声誉都受到了挑战。目前,收入水平低下和中等偏下的国家每年对初等教育的经费投入大约

* 原文语言:英语

Helen Abadzi, Luis Crouch, Marcela Echegaray, Consuelo Pasco, Jessyca Sampe (秘鲁)

Helen Abadzi, 世界银行高级评估官员; Luis Crouch, 三角国际研究所 (Research Triangle Institute, RTI) 国际发展集团副主席; Marcela Echegaray & Jessyca Sampe, 秘鲁 Tarea 的研究心理学家; Consuelo Pasco 也是 Tarea 的一名教育研究者。

E-mail: habadzi@worldbank.org

是 700 亿美元(作者根据 2003 年的数字估算),但是,如果要在世界范围内百分之百地实现全民教育目标,则每年还需要额外投入 70 亿~170 亿美元(UN Millennium Project, 2005)。除了政府投资和社会捐赠以外,父母的投入也在教育经费投入中占了很大份额——在一些国家甚至占到整个教育投入的 30%。就父母自身而言,送孩子上学没有任何利益可得,所有这些投入都只有一个目的,即让世界上的每一个孩子都能通过学习知识、获取信息而掌握基本技能,并改善自己的生活。

在捐赠团体的帮助下,很多国家有时会大幅度提高入学率,利用社会的力量减少留级现象的发生。以往的经验表明,学习成绩和入学率两者之间紧密相连,相辅相成。但如果增加学生毕业的难度,这种关系恐怕就不复存在了(Crouch & Fasih, 2004)。那些在完成家庭作业时得不到父母帮助的孩子,可能只能通过几本教材接受教师的指导,而这些教师的教育水平本来就不高,授课用的还是外语。在这种情况下,很多孩子可能虽然已经接受了初等教育,但仍然是功能性文盲(functionally illiterate)* (World Bank, 2005)。这样做的结果可能就是资源的浪费,令众人大失所望,抵消了入学率提高所取得的成就,也会因为无法实现期望而产生社会紧张感。

显然,学习的结果必须受到密切的监督,如果这个结果不能让人满意,就必须采取一些适当的措施。一些社会公益组织已经通过广发信息和投入资金等措施,帮助一些国家开发了在国际上具有可比性的基于案例研究的成绩测验系统。但是,这种测验系统价格昂贵,费时较多,并且存在方法论上的问题。使用这种测验比较典型的结果是,读者会得到一个确定的百分比,这个数字等于或者高于令人满意的成绩标准(通常确定为测验题目 60%~75%的正确率)。例如有测验表明,1999 年乌干达六年级学生中有 15%能够正确地或较好地运用英语进行读写(World Bank, 2004a, b, c)。决策者经常会面对非常低的数据结果,但对于学生们会什么不会什么,却没有清楚的概念。通常来说,接受这种测验的学生社会经济条件较好,高于全国平均水平(例如也门、几内亚、尼日尔),但是被试的特征可能每年都不相同。如果参加这种测验的贫困学生的比例逐渐增加,得到的分数可能就会大幅度下降,但是也不能清楚地显示测验变化究竟意味着什么。这样一来,许多低收入国家就不能从测验的结果中找出有实用价值的东西。

然而,众多实地观察表明,许多低收入家庭的学生在学校学到的知识少得令人吃惊。例如在说法语的几内亚,二年级结束时只有十分之一的学生能认全整个字母表,给他们 20 个单词,平均每个学生只能读出 4 个(Spratt & Tamba, 2005)。在马里说法语的学校,也得出了相似的结果。这些国家的学生成绩差可能归咎于没有强制穷人说法语的法令以及法语拼写的复杂性。但是,类似的情况也发生在玻利维亚,而西班牙语既是当地使用的语言,也是根据音标拼写的,根本不存在上面两个原因。在玻

* 指那些虽然能认得一词半句,但不能真正读懂和书写,不能应对日常生活中对文字的需要,如写信、看信、填写求职表格。——译者注

利维亚,尽管有许多捐款和大学的帮助,但每个年龄段最终只有三分之一的学生能够学会理解自己所阅读的简单信息。在这个国家,留级率为30%,很多青少年学生为了帮助自己的家庭而辍学,成为文盲(Schiefelbein, 2005)。同样,在巴西东北部地区,许多四年级学生只能结结巴巴地阅读,而上课用的课本他们则几乎不认识(World Bank, 2002)。

阅读可能只是学生在学校中需要掌握的技能之一,但是他们只有通过阅读才能从书本中获取信息,包括数学课本。那些不能在一二年级或者三年级学会阅读的学生,可能永远不能取得跟那些较早学会阅读的孩子一样的成绩。在众多的年级中,早期获得的读写能力至关重要,因为学生在很多时候必须独立完成一些任务。对那些一二年级毕业后就可能辍学的学生来说,早点学会阅读更是至关重要。^①

在很多国家,我们无从知晓有多少学生是功能性文盲以及为什么会这样。人们通常会把责任归咎于教育时间的限制,但是贫穷的人可能还会面对影响认知能力的健康问题,也许正是这些问题使得他们需要更长时间的教育(参见 Berkman et al., 2002)。是否有成本较低的甚至最贫穷的发展中国家都能容易地使用的简单的技能评价方法?下面我们将介绍秘鲁的一种实验性测试(pilot test)。为了便于理解这种方法所包含的测验问题,我们首先对阅读速度和理解能力作简短的介绍。

阅读速度和理解能力的基本要素

阅读包含很强的生物学成分,这是人们最近才充分认识到的。对文本的理解取决于正确的解码及丰富的词汇知识,但也取决于短时记忆(或者叫工作记忆)储存完整信息的能力以及对信息的加工能力(Passolunghi, Cornoldi & de Libero, 1999)。

正如很多教育者所知道的,所谓的工作记忆只能保存7条信息,大约12秒钟(具体内容参见 Reisberg, 2001)。如果这些信息是单词的话,粗略地讲,大概有7个单词可以保存在短时记忆中,并且最多只能加工12秒钟。这意味着要培养功能性读写能力,人们必须能在12秒内读懂大约由7个单词组成的句子,而这大概等同于1秒钟读1个单词,或者是1.5秒读1个单词,也就是1分钟读45~60个单词。^②为了克服工作记忆的12秒钟限制,大脑往往会开发出大块的字母和单词区域,然后迅速进行自动加工。脑成像研究显示,当人们能够流利阅读时,大脑中的一个单词快速识别通道就会被激活(位于左颞枕骨区域)(Shaywitz, 2003, pp. 76-87)。自动化的行为通常具有持续性,所以当人们学会自动化的流利阅读以后,一般就不会再成为文盲。

这样,阅读看起来就成了一件自相矛盾的事情。阅读慢的人必须付出更多的精力和时间去读书。如果他不能在12秒钟内读完一句话,那么他的工作记忆就会超出容量,当读完这句话的最后一个单词时,就会忘记开头的单词。这样就很容易让人失去耐心而最终放弃。那些读书结结巴巴的孩子可能就是功能性文盲。他们可能会苦

苦思索一些句子,但是不能阅读或理解整个课程文本。

为了激活快速识别通道,学习者需要持续练习单词的发音和书写。练习的多少取决于所学语言本身的拼写方法。欧洲的大多数语言拼写都是有规律的,单词的发音和拼写都有相似性。如果孩子们能在相对较好的学校学习发音,那么在一年级末就能十分准确地自动阅读一般的单词了。例如,意大利的儿童在一年级中期或末期几乎就能完全掌握阅读解码技能(Harris & Hatano, 1999, p. 17)。语言转换越复杂,需要的练习也越多,理解阅读内容要求的速度也就越快。一个说葡萄牙语的孩子在读“os mesmos, a gente, bom dia”这句话时,如果至少要部分理解其意思的话,就必须在千分之一秒内把“o”发作“u”,“d”发作“g”,“t”发作“ch”。英语和法语比较复杂,因此需要较多的努力才能进行流利的阅读。理解力有助于培养人们流利阅读的能力,因为在认识的句子中学习单词比单独学习某个单词要快得多(这叫做“单词优先效应”,具体内容参见 Abadzi, 2003a)。

标准和规范

考虑到阅读速度的重要性,人们设计了一些实用阅读能力标准。美国提出的阅读标准是一年级每分钟 30~70 个单词(朗读),二年级每分钟 60~100 个单词(默读),三年级每分钟 90~120 个单词,四年级每分钟 110~140 个单词,五年级每分钟 140~170 个单词,六年级每分钟 160~190 个单词(Barr et al., 2002, p. 76)。智利也有相类似的标准,其教育部为教师网络计划(Teacher-to-teacher Network Programme)确定的目标是,一、二年级学生每分钟读 30~70 个单词。^③智利的非政府组织教育协会(Educando Juntos)提出的目标是,一年级学生每分钟读 34 个单词,二年级学生每分钟读 64 个单词。^④古巴专家 Jose Perez Villar(1996)认为,一年级结束时,一个正常孩子的阅读速度应该达到每分钟 30 个单词。西班牙的一项研究报告(不是标准)则显示,一、二年级学生的平均阅读速度分别为每分钟 50~55 个和 75 个单词(Equipo de Orientacion Educativa de Marbella, 2003)。在美国说西班牙语的低收入者中,如果一、二年级的学生每分钟只能阅读 30~60 个西班牙语单词,就被认为是有所问题的(参见 de la Colina et al., 2001)。这些阅读有所问题的学生的阅读速度大约为一年级每分钟 35 个单词,二年级每分钟 60 个单词。这样一来,在美国确认为阅读西班牙语有困难的标准就类似于拉丁美洲各国提出的正常阅读标准。但是没有人会因此认为,对在美国的西班牙裔儿童的期望要高于对秘鲁中等以下低收入家庭儿童的期望。所以,对待这些标准不能太认真。但是,严格控制阅读是有可能在数周内大幅提高阅读速度和理解能力的(de la Colina et al., 2001)。

然而,对中等收入国家学校的实地考察表明,在这些国家的幼儿园、一年级或不设幼儿园的小学前两个年级中,应该有足够的资源培养所有孩子具有流利阅读能力。比较研究显示,在一些捐赠团体中普遍存在一种虽然不是很明显的想法,即认为功能

性读写能力的培养需要4~6个学年(Mingat & Rakotomalala, 2005)。这样,联合国教科文组织在统计一个国家的识字人数时,就是以该国完成四年级教育的人数为基础的。因此,我们所讨论的问题就是政府和捐赠人应该有清晰的期望和概念,更早地培养儿童的功能性读写能力。

秘鲁的一项实验性阅读评价测试

方法

这种实验性测试方法的产生有一定的社会背景。当时,作为世界银行评估项目的一部分,为了找到改善秘鲁学校教学质量的方法,人们进行了一系列的分析和讨论。秘鲁的各项教育指标,不论是普及率还是毕业率,都完成得很好。整个教育系统各级学校的毛入学率都很高。至少在十年内,初等教育的毕业率都在95%以上。秘鲁是“国际学生评价项目”(Programme for International Student Assessment,简称PISA)的成员国之一,其在这个计划中所得的评价分数低于其他拉美成员国家(阿根廷、巴西、智利和墨西哥)。此外,秘鲁四年级学生的数学和语言学业测试的结果通常也是需要考虑的一个因素。

文本单元1 西班牙语阅读文本: Dogo

Había un perrito gordo y peludo llamado Dogo. La familia con quien vivía lo quería mucho. Dogo era un perro obediente, cuidaba la casa, pero no comía toda su comida. Un día salió de paseo con su amo Lucas y se perdió. Lucas se puso triste, pero felizmente Dogo apareció al rato. Lucas lo cargó y lo llevó a su casa.

阅读理解问题:

1. Cómo se llamaba el perro?
 2. Con quién vivía el perro?
 3. El perro comía toda su comida?
-

在这项实验中,我们在学年结束前的两个星期对一年级和二年级学生进行测验。测验题目是从官方的一年级课本中挑选的极为普通的一段话,稍作修改,使之在难易程度和文化上适合贫困儿童的阅读水平(参见文本单元1)。文本的长度控制在60个单词左右,即对一个阅读水平相对较好的人来说,大约一分钟就能读完。^⑤孩子们有可能以前读过这段话,但我们并不把它当作一种干扰因素,因为他们恰好能记住这段话的几率几乎不存在。相反,必须保证这门课是学过的,教材也是曾经用过的。一年级和二年级使用相同的测验文本。由于学校在教学时采用的是本地语言——奇楚亚语(Quechuan)*和Awuajun语,因而测试文本同样采用这两种教学语言。然后,我们要

* 奇楚马兰语系中的一个亚语系,也是该语系中最重要的语言,说这种语言者主要是南美洲安第斯高原各国的印第安人。——译者注

求被试儿童大声朗读所提供的文本,并回答三个简单的理解性问题。

表 1 取样测试学校的特征

县	省 份	地 区	贫困等级 (1 个最贫困的, 5 个比较贫困的)	学校数量和类型
农村林区				
San Martín	Rioja	Awajun	2	3 所标准化公立学校
Loreto	Loreto	Nauta	1	3 所标准化公立学校
农村山区				
Ancash	Yungay	Yanama	1	3 所标准化公立学校
Cusco	Anta	Ancahuasi	2	3 所公立双语学校
城市沿海				
Lima	Lima	Pachacamac	3	3 所标准化公立学校
Piura	Piura	La Arena	2	3 所标准化公立学校
声誉“较好”的公立学校(沿海城市和农村山区)				
Lima	Lima	Puente Piedra	3	1 所声誉“较好”的公立学校
Cusco	Anta	Limatambo	2	1 所声誉“较好”的公立学校
声誉“较好”的私立学校(沿海城市和农村林区)				
Lima	Lima	Puente Piedra	3	1 所声誉“较好”的招收贫困学生的私立学校
San Martín	Rioja	Awajun	2	1 所声誉“较好”的招收贫困学生的私立学校

参加实验的学校并不是随机挑选的,而是从占 40% 的低收入区中挑选 22 所学校,由这 22 所学校代表该国不同地区不同文化的所有学校(沿海地区、山区和丛林地区,城市和郊区)(参见表 1)。要注意排除一些明显的主观偏见,如一些便利条件或是某些特定计划的影响。然而为了节约成本,我们没有选择那些路途较远、一天内不能到达的学校,因此可能遗漏了一些真正困难的学校。为了便于比较,我们还选择了贫困社区中几所“较好”的公立学校和私立学校。

从每所学校的一年级和二年级中分别随机挑选 5 名学生,总共有 245 名学生接受测验。每位学生阅读时都用一个小的便携式数字录音机进行录音,这样就能快速并且完整准确地记录每位学生的大多数重要信息(年龄、性别,是否接受过学前教育)以及阅读速度、语调和需要理解的问题。然后,把这些信息传输到电脑上,便于以后的快速分析和理解统计。这套程序使得在大约两分钟内测试一名学生成为可能,为以后的详细研究提供了第一手资料。对一个操作熟练的评价者来说,测试一所学校中的 10 名学生大约只需要 30 分钟。

结果

一年级和二年级分别只有 25% 和 41% 的学生能阅读文本中一个或一个以上的单词(见表 2)。接受实验的学校,一年级结束时,75% 的学生都不能阅读文本中一个简单的单词。

表 2 阅读结果概况

	参评学生数	至少能读一个单词的学生比例数	平均每分钟的阅读量 *
一年级	109	25%	9
二年级	136	54%	29
共 计	245	41%	20

资料来源：百分比是作者根据被试数量计算的。

一年级每分钟 9 个单词的阅读速度显然远低于多数中等拉美国家每分钟 30 个单词的标准。但这一结果与秘鲁的其他研究结果是一致的。2001 年的全国评估发现,只有 34% 的学生有“充分的”阅读理解能力(参见 Espinosa & Torreblanca, 2003)。因为这次阅读研究实验的对象来自中等收入以下的底层贫困儿童,统计结果低于 34% 也是可以想象的。虽然测试是不同的,但是我们可以参考一二年级每分钟 30 个和 60 个单词的阅读速度。在这个前提下,只有 18% 的被试儿童的表现是“达标的”。这个结果与一些实地观察研究的结论是一致的,即许多农村学校三四年级的学生还不会阅读(Lopez, 2002)。然而,在我们的被试中,有 55% 的学生曾经接受过一定形式的学前教育,他们本应该更早地学习文字(官方统计资料显示,大约有 60% 的儿童接受了学前教育)。

在一二年级获得的阅读技能是非常重要的(见表 2),但是以后的年级如何就不得而知了。功能性识字者是在以后的年级中培养出来的,但并不清楚具体是在什么时候(作者并不希望在较高年级的测验中也得到同样糟糕的结果)。虽然秘鲁在 PISA 中的得分较低,但是这并不意味着很多孩子到 15 岁时还是文盲,就像非洲说法语的那些国家那样。但是,如果孩子们在学校的前几年都是文盲,那么他们以后到高年级时从课本中学习知识的能力就会受到限制。没有足够快的阅读速度,就不可能读懂数学问题,更不能解决问题,他们的数学知识也就可能因此受到限制。这种读写能力学习系统的低效性造成一种恶性循环,既浪费了资源,也浪费了孩子们的时间。

根据回答三个理解性问题的正确率来计算,理解水平与阅读速度的相关系数为 0.82(见表 3)。^⑥那些阅读速度较慢的学生几乎不能正确回答问题,而那些阅读速度超过每分钟 50 个单词的学生能正确回答 86% 的问题(见表 4)。三个问题都能够正确

* 不能阅读学生的阅读速度为零。

回答的学生的平均阅读速度为每分钟 77 个单词,而只能回答一个或者两个问题的学生的阅读速度分别为每分钟 15 个和 41 个单词。尽管这只是一种相关性研究,但研究的结果是符合记忆研究预测的,即理解力的培养需要有足够快的阅读速度(假定学生有一定的词汇量知识)。这样,阅读的速度就决定了理解力的高低,至少在低年级是这样的。

表 3 学校中阅读因素之间的相关性

能够阅读的学生的百分比与阅读速度	0.86
阅读速度与理解力	0.82
能够阅读的学生的百分比与理解力	0.69

资料来源:作者根据调查数据计算得出。

表 4 阅读速度与理解力之间的关系

速 度	一二年级学生数	理解力 (正确答案的平均百分比)
慢(每分钟 4~24 个单词)	40	26%
一般(每分钟 25~49 个单词)	39	53%
快(每分钟 50 个及以上单词)	43	86%

资料来源:作者根据调查数据计算得出。该项计算只包括那些真正能够阅读的学生。

阅读速度与理解力之间的这种关系表明,提高阅读速度就能增强理解能力,但是如果能够重视早期理解力的培养,则可能并不需要过多的练习就能实现流利阅读。当孩子们学会阅读以后,似乎很快就能达到流利阅读的水平(Shaywitz, 2003)。对于那些拼写简单的语言,儿童在 4~6 个月内就能“破解密码”,一年级末就能达到每分钟 45~60 个单词的阅读水平(Seymour et al., 2003)。因此,如果学校肯花几个月的时间让儿童练习阅读,同时也就可能培养他们的理解能力。

尽管取得的平均结果分数较低,但是具有类似社会经济水平的学校之间的区别还是很大的。这再次说明了一些标准的缺失——在这种体制下,只有那些足够幸运的孩子才会遇到负责任的教师,能够在恰当的时间学会阅读。我们通过表格可以看出,三所城市私立学校的学生阅读速度最快,理解力最好(见附录表 1)。相比较来说,城市里的学校更有可能获得较高的分数。农村学校得分较低,但并不都是如此。安第斯地区的三所学校,一年级根本没有一个孩子能够阅读,但是到了二年级,能阅读的儿童的比例上升到 40%~100%,阅读速度也从可怜的每分钟 13 个单词增加到合理的每分钟 60 个单词。在城市公立学校,有一所学校一年级没有一个学生能够阅读,但是在另一所条件相似的学校,100%的学生都能阅读。对亚马逊河地区的两所农村公立学校的评价显示,其中一所学校,一二年级分别有 20%和 80%的学生能够阅读,而在另一所学校,无论哪个年级,甚至没有一个学生能够阅读一个单词,百分比

为零。那些执行自己标准的学校和班级取得的分数也是比较高的,有时他们是在非政府组织或其他机构的帮助下实行的。第11、12类尤其是第7类学校都是很好的例子。国际上也有很多类似的案例,一些学校能超越自己所处的环境[Lockheed & Harris (2004)对牙买加的研究]。一些教师能够比其他教师在学生身上多用180多个小时,积极引导他们(Christian et al., 2000)。

秘鲁没有明确的成绩标准。为了收集更多的信息,弄清楚教师评价自己和学生表现的标准是什么,我们调查了参加实验的学校的54名教师。当被问到这个问题时,71%的教师称自己了解课程结构;几乎所有的教师(94%)都说自己制定了具体的目标,有一些教师对自己是如何制定目标的以及目标是什么,给出了合理的回答。但是,只有41%的教师称自己是根据政府机构或其他教育权威的指示确定自己的目标。大多数人称是根据自己的经验或其他可以信任的同事的经验确定目标的。80%的教师和95%的校长认为课程结构太难,难以执行,或者认为课程结构是不恰当的。73%的教师认为,课程主要是由专家确定的,并没有征求教师的意见。因此,课程指导并不能很好地满足低收入家庭学生的需要。对这些教师来说,他们只能靠自己;只有13%的教师称在自己的要求下得到了帮助,而72%的教师反映,没有特定的机构对他们提出的问题和要求给予帮助(显然,现存的机构并不能很好地发挥自己的作用)。81%的校长认为,职前培训的质量和定位对现实的课堂教学来说是不恰当的。根据这些数据,我们可以很明显地看出,教师培训、对教师的支持以及控制系统并没有帮助教师实现教学目标,教会学生学习。

不同班级表现上的差异和教师回答上的差异表明,我们面临的状况并不是毫无希望的,单独一位教师或一所学校,只要注意技巧,还是能够取得成就的。但是,仍然需要制定明确的标准、管理制度及监督制度,因为在我们的研究中,好结果的出现有一定的随机性。在秘鲁,并不能根据那些没有特殊经济来源、接收贫困儿童、经营最好的学校中的学生在阅读速度和理解水平上取得的成绩制定标准。例如,对城市学校一年级的学生来说,每分钟35个单词可能是一种标准(在我们的案例中,速度最快的三所学校分别为每分钟54个、39个和26个单词,如第7类和第5类学校,学校1和2属于后者)。对二年级来说,阅读速度可能是每分钟70~80个单词(最好的三所学校的速度分别为每分钟96个、77个和67个单词,如第7类学校和第5类学校中的第3所和第6类学校)。这些标准接近我们上面说过的智利的标准,但低于经济合作与发展组织(OECD)国家中产阶级儿童的标准,而且根据这种标准,美国的西班牙裔儿童会被认为是很危险的。

既然农村地区双语学校的得分很低,那么它们可能一开始就应该制定相对较低的标准。很明显,我们只有充分发展并理解那些更贫困地区的双语学校,并根据其情况确定标准,达成共识,才有可能为所有学校确定统一的标准。教育部的在职培训和地区一级或省一级的教育培训(还有大学和教师培训学院的职前培训),都可以促成