



# 教育评价

JIAO YU PING JIA

[日] 田中耕治 © 著  
高峡 田辉 项纯 © 译



北京师范大学出版集团  
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP  
北京师范大学出版社

# 教育评价

JIAO YU PING JIA

[日] 田中耕治 © 著  
高峡 田辉 项纯 © 译



北京师范大学出版集团  
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP  
北京师范大学出版社

---

**图书在版编目(CIP)数据**

教育评价 / (日) 田中耕治著; 高峡等译. —北京: 北京师范大学出版社, 2011.1

ISBN 978-7-303-11747-5

I. ①教… II. ①田…②高… III. ①教育评估—研究—日本 IV. ①G40-058.1

中国版本图书馆 CIP 数据核字(2010)第 216023 号

北京市版权局著作权合同登记号: 01-2009-4391

---

营销中心电话 010-58802181 58808006  
北师大出版社高等教育分社网 <http://gaojiao.bnup.com.cn>  
电子信箱 beishida168@126.com

---

出版发行: 北京师范大学出版社 [www.bnup.com.cn](http://www.bnup.com.cn)

北京新街口外大街 19 号

邮政编码: 100875

印刷: 北京中印联印务有限公司

经销: 全国新华书店

开本: 170 mm × 230 mm

印张: 14.25

字数: 198 千字

版次: 2011 年 1 月第 1 版

印次: 2011 年 1 月第 1 次印刷

定价: 28.00 元

---

策划编辑: 郭兴举

责任编辑: 李洪波 曾忆梦

美术编辑: 毛佳

装帧设计: 毛佳

责任校对: 李茵

责任印制: 李啸

**版权所有 侵权必究**

反盗版、侵权举报电话: 010-58800697

北京读者服务部电话: 010-58808104

外埠邮购电话: 010-58808083

本书如有印装质量问题, 请与印制管理部联系调换。

印制管理部电话: 010-58800825

# 译者序

什么是教育评价？心理学、测量学和教育学意义上的评价究竟有着怎样的不同？不同的评价立场、方法和结果的使用会对学生的发展产生怎样的影响？这些看似简单的问题实际并不容易回答，甚至许多多年从事教育研究的人对教育评价的理解仍处于混沌状态中。进入 21 世纪以来，我国的教育改革在如火如荼地进行，作为教育过程中的重要一环，评价所具有的指导、反馈和调节功能为越来越多的人感受和认识。正确地认识、恰当地把握、有效地运用教育评价，促进学生全面、健康地发展和成长，即成为实现教育改革目标中的关键环节。

本书作者田中耕治教授是日本教育评价领域的权威专家，现任职于日本著名学府京都大学教育学院。一直以来，他坚持从教育学和教育方法学研究的立场出发，孜孜不倦地探索教育评价的意义、价值和方法，本书即田中教授经多年积累形成的研究成果。2008 年 7 月日本最权威的出版机构“岩波书店”出版了此书，引起很大反响，受到教育界学者和专家的广泛好评。短短两年内已再版三次，仍供不应求，这在日本教育学术界是很少见的。

在本书中，作者通过对教育评价概念的梳理和阐释，对教育评价功能的比较和分析，对日本教育评价发展历史的回顾与评述，向读者系统展现了日本教育评价发展的历史、现状与问题，使人们清晰地了解近代以来，“教育评

价”思想在日本本土化的过程所经历的曲折、面对的困惑和不懈的探索，教育评价观念的演变和发展对教育实践、考试制度等产生的重要影响。此外，作者还通过丰富的史料、事例，以及对日本课程变迁、学力模式发展的深刻分析，揭示了课程与评价、学力与评价、指导与评价的有机关系。这些都使本书具有很强的可读性，凸显了作为教育方法学者的研究特色，使之既可以成为学者的宝贵研究资料，也可以成为一线教师的实践指南。这也是本书区别于其他教育评价类著作的主要特点。

中日两国有着相近的传统文化和教育制度，也同样重视考试和评价。近年来课程改革的不断深入，更是将教育评价推向了亟待改革的关键阶段。为此，我们将本书介绍给关注和探索教育评价的同行，希望通过本书让更多的人理解，同为考试文化浓重的日本是如何将舶来理论去伪存真、融会贯通，创设适于本国的评价制度与体系的，也期待这样的引进对推动我国教育评价的改革能有所借鉴和启示。

本书的翻译工作由高峡、田辉和项纯合作完成。其中，田辉翻译前言、序章和第一部，项纯翻译第二部，高峡翻译第三部、结语和卷末资料，并负责全书的翻译统稿。本书的引进和翻译，得到了田中耕治教授的悉心指教，也得到了北京师范大学出版社和日本岩波书店的鼎力支持，在此，我们深表谢意。

译者

2010年10月

# 前言

当你听到“教育评价”这个词的时候，会产生什么样的感觉？这是一份以教师和学生为对象进行的调查问卷中的一道题。这项调查要求被调查者用“评价就像……一样，因为它……，所以让人感到……”这样一种规定句式来回答“什么是教育评价”的问题。由此得到了一些很有意思的调查结果，有些回答甚至不亚于相声段子中的“包袱”。之所以采用这种调查方法，是因为用比喻的方式提问，比直截了当地询问“你如何看待教育评价”更能够让被调查者根据自身的生活感受，对问卷中提出的事物产生自由联想，以便获得更接近人们对于教育评价真实感受的回答。那么，你将如何回答这个问题呢？

调查结果虽然在意料之中，但仍使人感到人们对“教育评价”的感觉相当压抑（负面）。调查结果中出现了很多类似于“（教育评价）像天堂与地狱（因为有喜有忧）”、“像拔牙、过吊桥（因为心里没底）”的回答。其实这样的回答还算比较好的，甚至还有“像烧红的烙铁，因为会受到伤害，必须抛弃同情与怜悯”这一类相当负面的回答，让人隐约感到回答者对教育评价痛苦的体验。“教育评价”让人联想到“考试”和“成绩通知单”，宛如听到“如果没有考试和成绩通知单，上学该是一件多么令人愉快的事情啊”这样的叹息声。加之，近年来接连不断地进行“教师评价”、“学校评价”、“（学校的）自我评价”、“学力调查”，我们似

乎突然闯入了一个“评价的时代”。

以这样的文字作为开头，或许会让人预感到这本以“教育评价”为名的读本，可能会涉及很沉重的话题。当然，我们不能回避人们对“教育评价”所抱有的晦涩、沉重的感觉，但仔细想来，正是这种压抑和沉重，表明“教育评价”对教育具有相当重要的作用。如此这样，我们应当在梦想没有“考试”、没有“评价”的学校和社会之前，先从各个角度对恶魔般的“教育评价”进行仔细观察，弄清其本来面目，进而探索通过“教育评价”为教育改革服务的渠道，这才是至关重要的。

事实上，关于“教育评价”的理论与实践，我们已经迎来了一个重大转折的时代。简而言之，“教育评价”从原来的对学生进行“估价”、是让学生对学习和发展产生厌恶感的工具，变成保障学生高水平学力、促进学生参加教育实践的有效措施。在本书中，我们将首先在序章部分介绍和明确“教育评价”的新动向，接着在第一部中，回顾“教育评价”的观念转变及其艰难的历程，同时进一步明确“教育评价”的基本理念。

在第二部中，我们将主要探讨伴随着教育评价观念的转换，对“教育评价”的功能提出了哪些新的要求，以及怎样的评价方式能够满足这种要求。特别是通过介绍最近引起强烈关注的“表现性评价”、“档案袋评价”等方法，展现这些方法所具有的丰富的教育构想力。另一方面，在“教育评价”的变化过程中，不仅有教育实践的影响，还有制度方面的影响。以往曾经有很多教育改革都受到了“教育评价”制度的阻碍。因此，作为新的“教育评价”理念，就必须关注如何进行“教育评价”制度的改革。最后在第三部中，我们将通过解读日本“教育评价”的历史，从历史的角度对第一部、第二部展开的内容加以梳理。

我期待，以上述内容构成的这一读本，将有助于改变人们对“教育评价”的印象，并且希望该读本中提出的“教育评价”的新理念，能够与教育实践的改进紧密结合在一起。

注：本文开头部分的问卷题目参考了小山真纪的《关于评价与音乐课的印象的讨论》[载《音乐教育学》，第27卷，1997(1)]。





|                            |     |
|----------------------------|-----|
| <b>第二部 教育评价的功能与方法</b>      | 77  |
| <b>第四章 教育实践与教育评价的功能</b>    | 79  |
| 第一节 教育评价的层次、主体、对象          | 79  |
| 第二节 学力评价论的展开               | 90  |
| 第三节 教育评价的功能                | 118 |
| <b>第五章 评价方法的原理与新的评价方法</b>  | 127 |
| 第一节 评价方法的原理                | 127 |
| 第二节 多样化的评价方法及有计划的使用        | 137 |
| 第三节 表现性评价与档案袋评价            | 145 |
| <b>第六章 教育评价的制度及其改进</b>     | 156 |
| 第一节 指导要录与指导的连续性            | 156 |
| 第二节 成绩通知单与公开讨论会的形成         | 161 |
| 第三节 入学考试制度的改革——从“选拔”到“资格”  | 168 |
| <b>第三部 日本教育评价的历史演进</b>     | 177 |
| <b>第七章 近代日本教育评价的历史</b>     | 179 |
| 第一节 “考试”时代                 | 179 |
| 第二节 “考查”时代                 | 182 |
| <b>第八章 第二次世界大战后的教育评价历史</b> | 194 |
| 第一节 第二次世界大战后初期的教育评价改革      | 194 |
| 第二节 能力主义的发展与“相对评价”         | 196 |
| 第三节 批判“相对评价”，创造新的教育评价      | 204 |
| <b>结 语</b>                 | 210 |
| <b>卷末资料</b>                | 212 |

## 序章 学力调查与学力状况

为了能够清楚地阐释当代教育评价研究所面临的问题以及所要表达的社会性关联，首先应该了解在各种调查评价中逐渐浮出水面的当代日本中小学生学习力的实际情况。

在第二次世界大战之后的教育领域里，学校主要培养的“学力”已经由仅有少数精英关心的个别问题，变成了社会各阶层广泛关注的普遍问题。这种关注程度在经济高速成长期一路飙升，从而使“学力”成为一个社会性问题，进而演变成日本第二次世界大战后教育史上具有划时代意义的“学力问题”。各种各样的学力调查，都是以提供“学力问题”的实证数据为目的而实施的。无论是为了给第二次世界大战后第一次“学力低下”问题提供重要数据支持的久保舜一调查(1951)，还是为了解决20世纪70年代初期所谓“垮掉的一代”问题实施的国民教育研究所调查(1975—1976)，都具有很大的影响。<sup>①</sup>

此外，在大家仍然记忆犹新的“学力(下降)论”风波(1999—2003)中，围绕学力调查是否能够证明“学力低下”、学力调查的信度等问题曾发生过

---

① 追溯学力调查的历史，日本大正时期兴起的针对个人的学力测量以及以成年人为对象的“全国壮丁教育普查”(1925—1943)、“日本人读书能力调查”(对象15~64岁，1948)等，是日本学力调查的雏形。然而，有意义的学力调查起源于第二次世界大战之后。战后最初实施的学力调查，是发端于以实施“新教育”引发的“学力低下”问题。针对“不会写信”、“不知道县厅政府所在地”等简单质朴的问题与批评，为了纠清真伪，进行了各种学力调查。其中，最具代表性的是久保舜一的调查、日本教育学会的调查、日本教职员组合的调查等。特别是久保舜一的调查使用了战前学力测量的试题(茗溪会研究部制作的数学试题，1931)，提交了“战后日本儿童的算术学力比战前下降了两年水平”的结果报告，引起哗然。其后具有划时代意义的学力调查，以20世纪70年代逐渐显露出来的“低学力问题”为契机，由全国教育研究所联盟提出报告，认为“有半数以上的孩子跟不上教学进度”，这件事被媒体大肆报道，引发了备受关注的教育质量“下降、降低”的讨论。国民教育研究所组织认可了这份报告的内容，并因此实施了全国规模的学力调查。其结果暴露了“学力差距逐渐扩大”、“随着学年的增高学业成就下降的逆转态势”、“会计算但不明白意思的学力结构”等日本学校教育的严峻问题。

激烈的争论。然而，在关于“学力低下”的大讨论中，人们关注的焦点是此次调查结果与以往学力调查以及国际学生学业成就相比是否下降的问题。<sup>①</sup>而得出下降结论的“学力质量和构造”，究竟以什么评价方法来认定，其中的问题实质是什么，这些问题都还没有充分说明。

这种只强调“学力水平”的观点使得教育上的解决策略更加短路，会导致诸如“增加高考压力”、“机械灌输式教育”等主张的抬头。不言而喻，这样的解决方法不仅不能打开现有局面，而且只能导致“学力问题”的进一步恶化，历史早就证明了这一点。因此，我们要提出一种使人们能够冷静而灵活地解读学力调查的分析视角，并且以这种视角来探究当代日本学生学习能力的真实状况。

## 一、学力调查与分析视角

首先，我们将“学力调查”定义如下：

普通的学力调查是指，以就读于小学到高中（近些年往往包括大学）的学生群体为对象，对他们的推定学业成就实际情况进行有组织、有体系、有系统的把握，提供实证性数据的调查。调查结果用于调整和改善对学生的学力产生某种影响的社会环境（地域、经济、文化等条件）和教育环境（与学校、课程有关的各种条件）。

然而，实施学力调查，首先要通过确定调查科目和问题结构的学力观以及学科教育研究和教育评价目标，具体确定调查日期、调查期限、调查对象的选择范围、抽样标准以及结果分析方法等实施细则，以确保调查的有效性和精确度。同时，实施学力调查的主体机构的性质、参与成员的选定、直接调查者和评分者的熟练程度等，都直接影响着调查的成败。

此外，在进行国际学力调查时，要考虑到各国的文化、教育条件的差异，必须选取可比性较高的调查项目，还要考虑语言翻译的实际情况。

---

<sup>①</sup> 关于“学力低下论”（1999—2003），本书参考了日本“中央公论”编辑部中井浩一编的《辩论·学力崩溃》（中央公论新书，2001），与市川伸一的《学力低下论》（千曲新书，2002）。

当然,要实现能够满足该定义所有条件的学力调查是比较困难的,其中 PISA (Programme for International Student Assessment)、TIMSS (Trends in International Mathematics and Science Study)这样的国际学力调查,以及日本国立教育政策研究所教育课程中心进行的全国调查(以下称国研调查)等,都自觉依照上述学力调查的各项条件,值得研究与参考。<sup>①</sup>

那么,究竟哪些条件决定学力调查的分析视角呢?如前所述,仅以“学力水平”进行分析,未免太单一,为了对学力调查进行多层次、多角度的分析,至少需要“学力水平”、“学力差距”、“学力的内容和结构”和“对学习的愿望”四个角度。

---

① OECD(经济合作和发展组织)实施的 PISA 调查,每三年一次。已经实施过的调查有“PISA2000”、“PISA2003”、“PISA2006”。该调查以各国义务教育阶段的 15 岁学生为对象,主要考查各国学生对于将来生活必备的知识与技能的掌握程度。这项调查对各国的教育政策有很大影响(参见原田信之编:《扎实的学力与多样化学力——各国教育改革现状与学力模式》,新希望书房,2007)。而且,“素养”(literacy)原本是“识字能力”的同义词,但从 20 世纪 80 年代开始,扩展为普通交流能力,现在也用于表示某一特定领域的知识、能力(计算机素养、媒体素养)等。由于受到 PISA 调查的影响,近来一般都将其理解为学力中的“灵活应用能力”。

另一方面,由 IEA(国际教育成就评价学会, International Association for the Evaluation of Educational Achievement)实施的“国际数学和科学调查”,是对学业成就(Educational Achievement)的实证性调查。这一调查现已成了各参加国教育改革的导火索,特别是对于美国、英国 20 世纪 80 年代开始的、以学力提高为目标的教育改革影响深刻。原本“TIMSS”中的“T”是“Third”用来表示第三次,但由于已经明确今后将继续进行学力调查,所以将“T”改为“Trend”(动向)。

文部省(2001 年起更名为文部科学省)实施学力调查的历史较为久远。1956 年至 1966 年期间,为了掌握学力状况,确立学习指导要领和教育课程相关政策标准,实施了“全国学力调查”。此后,作为抽样调查,进行了两次“教育课程实施情况综合调查研究”(1981—1983、1993—1995)。从 2001 年开始,由日本国立教育政策研究所课程研究中心继续进行“教育课程实施状况调查”,并从 2007 年开始实施了众所周知的“全国学力、学习状况调查”。关于以上学力调查各测试题的分析,请参见田中耕治编著的《解读新学力测试——PISA/TIMSS/全国学力、学习状况调查/教育课程实施状况的分析及其课题》(日本标准,2008)。

## 二、“学力水平”的视角

“学力水平”是指以可以用分数衡量的学力为对象，通过学力调查得到该群体的平均分。一般在谈到“学力低下”问题时，主要就是指这种可测量的“学力水平”的下降。

当然，如果单纯以 IEA 调查中日本的排名顺序一览表(参见表序-1)来看，原先独占金银牌的日本现在仅仅跻身前六名而已，这是很有刺激性的，然而我们也只能接受这个事实。另外，比较“PISA2003”和“PISA2006”，可以看到“阅读理解素养”排名(第 8 名→第 14 名→第 15 名)；“数学素养”排名(第 1 名→第 6 名→第 10 名)；“科学素养”排名(第 2 名→第 2 名→第 6 名)，呈下降趋势。从这一结果来看，日本学生的学力水平在世界排名上确实呈现出若干下滑趋势，这一点毋庸置疑。

表 序-1 IEA 国际数学和科学教育调查中日本排名顺序

| 实施年度      | 调查               | 小学            | 中学            |
|-----------|------------------|---------------|---------------|
| 1964      | 第 1 次数学调查        | —             | 第 2 名(1 年级)   |
| 1970      | 第 1 次科学调查        | 第 1 名(5 年级)   | 第 1 名(3 年级)   |
| 1980、1981 | 第 2 次数学调查        | —             | 第 1 名(1 年级)   |
| 1983      | 第 2 次科学调查        | 第 1 名(5 年级)   | 第 2 名(3 年级)   |
| 1995      | 第 3 次数学调查        | 第 3 名(3、4 年级) | 第 3 名(1、2 年级) |
|           | 第 3 次科学调查        | 第 2 名(3、4 年级) | 第 3 名(1、2 年级) |
| 1999      | 第 3 次调查第 2 阶段 数学 | —             | 第 5 名(2 年级)   |
|           | 科学               | —             | 第 4 名(2 年级)   |
| 2003      | 国际数学和科学教育动态调查    |               |               |
|           | 数学               | 第 3 名(4 年级)   | 第 5 名(2 年级)   |
|           | 科学               | 第 3 名(4 年级)   | 第 6 名(2 年级)   |

注：—为未参加项目。

但是也没有必要对这种下滑趋势做出过激的反应。如果从统计学角度来看“PISA2003”、“PISA2006”的成绩，可以看出日本仍然保持在世界最好

的行列中。此外,由于学生在“国研调查 2003”中,与过去相比对相同问题的解答水平有所提高,因此媒体在结果公布后即分别以“学力改善的征兆”(《读卖新闻》)、“宽松一代成绩提升”(《每日新闻》)为题进行了报道。当然,不能因为这些报道就可以放心大胆地盲目乐观,但至少可以认为日本学生的学力还是处于比较稳定状态的。<sup>①</sup>反过来说,如果只是关注学力下降,而忽视了“学力差距”、“学力内容和结构”以及“学习意愿”的问题,那才是最危险的。

### 三、“学力差距”的视角

所谓“学力差距”的视角,其字面意思是考察学生学力分散程度与差距的视角。即便是学力水平很高,也不见得所有学生的学力水平都高,而整体上的高“学力水平”,往往会掩盖“学力差距”的存在。

这种“学力差距”的问题,在 20 世纪 70 年代以后成为“学力问题”的争论焦点之一。对于这一问题提出最有力实证数据的是原国立教育研究所。该所根据“小学语文、数学学业成就调查”(1982 年实施)的结果指出:“学力差距”从小学 4 年级开始显现,小学 6 年级学生中,分别有 24.8% 和 16.9% 的学生跟不上语文和数学的教学(其成绩低于 5 年级学生的平均成绩)。并且跟踪调查表明,这些学习落后儿童的状况一直延续到中学仍然没有改变。<sup>②</sup>

分析近年来的三项调查“TIMSS2003”、“PISA2003”、“国研调查 2003”的结果表明,虽然学力差距还没有出现明显的两极分化(双峰骆驼的驼峰形),但是仍然有 1/3 强的学生出现了学习落后现象,而且还可以看到,随着学年增高学习差距固定化和逐渐扩大的倾向。

---

① 作为考查日本学生高学力水平原因的书,参见斯蒂文森、苏提格拉著,北村春朗等编译的《小学生学力的国际比较》[金子书房,1993(原著 1992 年出版)],另有反映日本教师“教学研究”优点的《向日本的数学教育学习》[苏提格拉、詹姆斯·施瓦特著,渚三郎译,教育出版社,2002(原著 1999 年出版)]。另外对出人意料排名很低的“阅读理解能力”,参见《与生存能力有关的 PISA 型阅读理解能力》(有元秀文,载《BERD》,第 6 号,2006),该文批评日本的教育对于此种欧美型阅读理解能力(整体把握全文、批判性地分析文本的能力)没有进行充分的指导。

② 参见天野清、黑须俊夫:《小学语文·数学能力》,秋山书店,1992。

比如,通过各国素养测评得分的分布情况,能够比较清楚地说明“PISA2003”(数学素养)中的学力分散状况。日本学生的高分和低分之间差距较大,显示分散度的标准偏差值为101,而世界排名第一的芬兰的偏差值为84,韩国为92,荷兰为93,超过100的有比利时(110)、德国(103)等。PISA结果表明,日本高一学生的学力差距,与各自所在学校之间的差距、各学科课程之间的差距是一致的。在强调提高“学力水平”、解决“学力低下”问题时,既不能忽视“学力差距”给“低学力”带来的影响<sup>①</sup>,也不能忽视“学力差距”的存在。重要的是,要构建能保证所有学生学习能力的教学和教育评价体系。本书主要通过“相对评价”与“基于目标的评价”的对比来考察学力差距问题。

#### 四、“学力质量与结构”的视角

如果说“学力差距”视角,主要是分析“低学力”问题,那么“学力质量与结构”的研究视角,则主要是让那些一直被认为水平很高的学生的学力(考试能力)问题浮出水面。这一视角很早就被大田尧质疑为“问题与答案之间的距离太短(缺少思维环节)”,中内敏夫则称之为“病态学力”,近年来佐伯胖等人将其作为“学校的规定知识内容”问题进行研究,对于学力等于死记硬背能力的应试学力的问题性、脆弱性进行了严厉的批评。<sup>②</sup>

在学力调查中,人们将学力的质量与结构设想为应有的学力模式(参见第四章第二节),并据此选择评价方法,进行调查问卷的选择、编制与分析。因而,学力质量与结构是分析学力调查的基本视角,同时这样才能如实反映出教育评价研究的水平,对于近年来的学力调查中反映出来的教育评价研究的详细考察,本书将在相关部分详述,这里先介绍其简单概况。

从“学力调查所涵盖的学力模式究竟是什么”这一视点出发,考察近年

① 当代社会中,学力差距与学生所处家庭的阶层地位差距相一致,有人指出,学生的“意愿差距”甚至影响着“希望差距”。割谷刚彦:《阶层化的日本语教育危机》,有信堂,2001;山田宏昌:《希望差别的社会》,筑摩书房,2004。

② 参见大田尧:《何谓学力》,国土新书,1969;中内敏夫:《什么是学力》,岩波新书,1983;佐伯胖等:《为了学校的重生》,1、2、3,东京大学出版会,1992。

来的三次学力调查,我们发现最简单的学力模式是“国研调查 2003”。小学数学中包括:“数学学习的兴趣、意愿和态度”、“数学的思维方法”、“数量图形的表现与处理”和“数量图形的知识与理解”四个方面。这种分类与日本“学习指导要录”(评价手册,详见本书第六章)中的“分类学习状况”栏目相对应,但与下面将要介绍的“TIMSS2003”和“PISA2003”的学力模式相比,未免显得有些粗糙。

“TIMSS2003”把学力分为“内容领域”(在校学习的内容)和“认知领域”(学生在学习该内容时所表现出的、以及所被期待的行动)。比如在数学中,“内容领域”包括:数、代数(在小学数学中为规则与关系——作者注,以下相同)、测量、几何、资料的呈现与分析、概率;“认知领域”则包括:(1)认识事实与顺序(简称为“知”);(2)运用概念(简称为“用”);(3)解答规则问题(简称为“解”);(4)进行推论(简称为“推论”)。显而易见,按照知、用、解、推论的顺序的学力程度逐渐提高,而频频遭到批评的“死记硬背学力”所对应的只是“知”的学力。这种将学力分为“内容领域”与“认知领域”的方法,所依据的是布卢姆(Bloom, B. S., 1913—1999)的教育目标分类学(taxonomy)。<sup>①</sup>

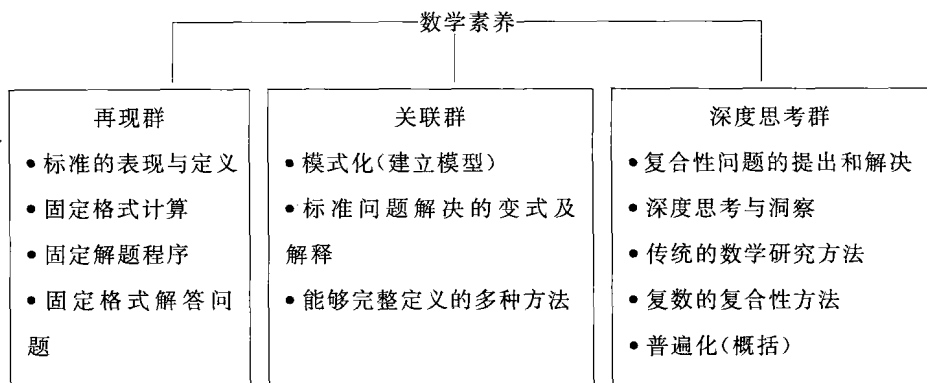
再来看看“PISA2003(‘PISA2006’基本相同)”。对相当于学力概念的“数学素养”的界定是“在复杂状况或者情境中,能够使用数学提出问题、列式、解决、解释问题的能力”,学力调查由内容、能力、情境三部分构成<sup>②</sup>,学力模式的能力则分为“思考与推理”、“论证”、“交流”、“建立模型”、“提

① IEA 调查的学力模式,从第二次调查(1980—1981)开始,有意识地采用了二元模式(内容与目标——计算、理解、应用、分析),第三次调查采用了三维模式(内容;行动期待“了解、简单程序、复杂程序、解决问题”;未来展望“由于难以编制调查问题,因而只进行了课程调查”),而“TIMSS2003”再次采用了二元模式。对于其中的原因,请参见长崎荣三、濑沼花子:《从 IEA 调查看我国中小学生的数学学力》,载《国立教育政策研究所纪要》,第 129 号,2000,以及木村元、北林雅洋、斋藤里美、平冈五月:《作为教育调查的 TIMSS》,载《教育目标·评价学会纪要》,第 12 号,2002。

② 关于 PISA,参见国立教育政策研究所编译:《PISA2003 年调查评价框架》,行政出版,2004(原著 2003 年出版);国立教育政策研究所编译:《PISA2006 年调查评价框架》,行政出版,2007(原著 2006 年出版)。



出和解决问题”、“表现”、“符号语言、公式语言、技术语言、演算”、“使用辅助手段与工具”八项指标，并且将主动应对具体问题的能力归纳为组群（集合体、结构化），分为“再现群（将经过练习所掌握的知识再现出来）”、“关联群（解决身边的实际问题）”、“深度思考群（策划、解决异地的实际问题）”等几种类型（参见图序-1），接下来我们来看实例。



图序-1 能力群概览

摘自国立教育政策研究所编译：《PISA2003 年调查 评价框架》，39 页，行政，2004。

在此，我们以日本学生感到困难的“盗窃事件”试题为例（参见图序-2），在这道题上日本学生的得分低于被调查国家的平均分。这道题的题型是“开放题”，情境设定为“公共关系”，考查的能力是“关联性分析”。日本学生的完全正确率为 11.4%（OECD 平均为 15.4%），正确率超过 20% 的国家有芬兰（26.5%）、意大利（23.3%）、新西兰（23.2%）等国。问题是根据现实生活中经常发生的情况设定的（具有真实性的问题），是测试数学全面素养的好问题。正确答案除要求答出“不适合”之外，还要从绝对意义和相对意义两方面对数值进行解释说明。这是一个难度较高的题目，在这次调查处于第一梯队的国家中，日本的成绩较差。<sup>①</sup> 包括这道“盗窃问题”在内，日本学

<sup>①</sup> 要正确回答这个问题，除了要选择“不正确”外，最好还要对“柱状图只提供了部分信息”、“难以换算成比例、百分比”等加以解释。国立教育政策研究所编：《生存的知识与技能》，2，119～120 页，行政出版，2004。