



新世纪高等学校教材

心理学基础课系列教材

[美] Lewis R. Aiken 著

张厚粲 黎 坚 译

# 心理测量与评估

PSYCHOLOGICAL  
TESTING AND ASSESSMENT



北京师范大学出版社  
BEIJING NORMAL UNIVERSITY PRESS

# 心理测量与评估

Psychological testing and assessment

(eleventh Edition)

[美] Lewis R. Aiken 著

张厚粲 黎 坚 译

北京师范大学出版社

## 版权声明

本书中文简体版由培生教育出版集团授权北京师范大学出版社出版, 仅限中国大陆境内销售。版权所有, 翻印必究!

本书英文版由培生教育出版集团 Pearson Education Group, Inc. 2003 年出版。

Chinese simplified language edition published by BEIJING NORMAL UNIVERSITY PRESS, Copyright© 2005

Authorized translation from the English language edition, entitled PSYCHOLOGICAL TESTING AND ASSESSMENT, 11<sup>th</sup> Edition by LEWIS R. AIKEN., published by Pearson Education, Inc, publishing as Pearson Education Group, Inc. Copyright© 2003, 2000, 1997, 1994, 1991, 1988, 1982, 1979, 1976, by Pearson Education Group, Inc.

All rights reserved. No part of this book may be or transmitted in any form or by any means electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

## 图书在版编目 (CIP) 数据

心理测量与评估/(美)刘易思·艾肯(Aiken, R. L.)著;张厚粲, 黎坚译. —北京:北京师范大学出版社, 2006. 3

ISBN 7-303-07893-2

I. 心… II. 艾… III. 张… IV. 心理测量学 V. B841.7

中国版本图书馆 CIP 数据核字 (2006) 第 011253 号

北京市版权局著作权合同登记图字: 01-2005-4790

北京师范大学出版社出版发行

(北京新街口外大街 19 号 邮政编码: 100875)

出版人: 赖德胜

北京东方圣雅印刷有限公司印刷 全国新华书店经销

开本: 170 mm × 230 mm 印张: 31.5 字数: 592 千字

2006 年 3 月第 1 版 2006 年 3 月第 1 次印刷

印数: 1 ~ 4 000 定价: 48.00 元

## 译者序

《心理测量与评估》一书的作者刘易思·艾肯 (Lewis R. Aiken) 是美国心理学会 (APA) 理事, 获得过大量的学术和职业荣誉, 并被列入美国名人录。他在佛罗里达州立大学获得了临床心理学学士和硕士学位, 又在北卡罗来纳州立大学获得心理测量和实验心理学的博士学位。随后, 在美国一些著名大学任教 40 多年, 在心理评估、社会 and 人格心理学领域有很多建树, 专著达 13 本之多, 并且有大量的论文发表, 同时他也是一名活跃在临床、教育和工业等应用领域的评估专家。

《心理测量与评估》一书在 1976~2003 年间共发行了 11 版, 是公认的经典之作, 已经成为美国大多数高等学校测量和评估课程的专用教科书, 也是专业人士的重要参考材料。该书在理论和实践上做到了很好的平衡, 围绕“心理测量和评估在教育—学校、临床—咨询以及工业—组织中的应用”来组织材料, 对测验的开发, 实施, 计分, 结果解释, 甚至管理和发行等环节上都有详细的阐述。该书对心理评估领域的描述饶有兴趣、丰富多彩、并且有很强的实用性, 能够引起专业学生的浓厚学习兴趣。而任何打算开发或者使用测验的实践者也会发现, 它不仅是一本优秀的教科书, 更是一本内容翔实的操作实务手册。

从 20 世纪 80 年代开始, 心理测验与评估领域的发展更加快速, 技术上也有了很大的进步、理论上逐渐成熟, 专业责任也在日益提升。同时各种测量工具本身也在发生变动, 新测验层出不穷。最为可贵的是, 《心理测量与评估》的每一版都能紧跟时代的发展, 全面展现最新的领域全貌。本书翻译的原版是该著作第 11 版 (2003 年), 其中对新出现的测验形式, 例如适应性测验、神经心理和发展性测验等进行了详细介绍, 同时也关注到测量理论的新进展 (项目反应理论等) 和新的测验实施平台 (电脑、网络等) 等给评估领域带来的变化; 对当前争论热点, 如“智力测验的利弊”等也有阐述。

第 11 版内容丰富, 几乎是心理评估领域的百科全书, 不仅介绍了几乎所有领域用到的心理测验和评估手段, 而且宣扬了一种测量文化和规范, 如: 测验的公平性, 测量的职业道德, 测验导致的法律诉讼与出台的法案等。它还展现了一个完善的测验经营、管理和服务体系, 包括维护测量专业性与高技术水准的协会, 测验出版商和代理商, 测验信息查询平台, 测验相关的刊物等。中国的测评事业正处于起步阶段, 在制定行业规范, 确定发展方向等方面, 可以从该书中获取很多有益的经

验。

我很高兴翻译这本书，希望它能够为我国的心理测量教学提供教学材料，为广大测评实践者提供重要的参考资料。感谢我的研究生们，包括骆方、王力、张青华、江雯等，正是由于他们的大力支持与帮助，才使得本书的翻译工作能够在较短的时间内完成。但限于我们的水平与精力，书中肯定还存在一些不足和错误，望读者不吝指正。

张厚粲

2005年12月

于北京师范大学

## 序 言

心理测量与评估多年来一直是被批判的对象，特别是标准化测验在教育和雇佣领域中的应用，一再受到攻击。大量的法律诉讼、法庭案例都与测验有关，有些州已经制定了关于测验标准和测验应用的法律。评价人及其行为的相关活动，以及所用的方法和工具，即使其所受到的批判是公正的，但是在这个人口超过 60 亿的世界里，人们对于评估、诊断、预测个体行为的方法的需求仍然是毋庸置疑的。尽管受到很多来自心理学界和教育界内外部的批评，心理评估还是在持续发展并日趋多样化。新的测验、调查表、量表的出现，以及相应的心理测量工具的编制、实施、计分和解释方法的发展进步，都证实了心理测量与评估正处于生机勃勃、充满活力的状态。许多因素促成了这种发展，这些因素包括：随着人口的持续增长，很大一部分人对社会服务的需求增加，以及雇佣、教育和临床领域中的选拔、安置、诊断都需要更为有效的方法。

在过去的几十年中，计算机在设计及编程方面的进步推动了心理测验的发展。自从 20 世纪中期计算机进入市场，就被用于测验计分和分析个体和群体的表现。随后，计算机又被用于测验和其他评价工具的施测以及对测验结果的解释。由此，测验和其他心理测量工具真正打开了心理评估领域的新局面，并且随着更精细的技术和程序的发展，毫无疑问会继续发挥更大的作用。

公众和专业人员对测验有效性和局限性的关注日益增多，这就需要对测验及相关评估材料的设计和传播给予更多关注。需要测验使用者更加训练有素，需要公众和专业人员更加重视心理测验在教育、临床、工业和组织背景中应用时的个人和社会价值，这些需求变得越来越明显和紧迫。心理测量专家和其他在测验及施测方面有丰富知识的人注意到，测验在编制和应用时，应该关注的不仅是它们的技术特征，还应保持对个人和社会总体的需求及权利的敏感性。很多专业组织，譬如美国心理学会，美国教育研究协会，美国人力及管理协会以及美国国家教育测量委员会，已经在出版物中提到此类事宜。

与这些关注和期待相一致，这本教科书的主要目标跟本书 30 多年前发行第一版时的目标一样，旨在提高那些编制测验、实施测验、参加测验、对测验结果进行评分和解释以及根据测验分数作出决策的人们的知识、理解和实践。像前面所有的版本一样，第 11 版主要是为大学生设计的一本教科书。它适合作为心理学、教育

学或其他相关领域的大学生以及低年级研究生的测量和评估课程用书。心理学家和其他编制、施测和评估工具并解释和运用测验结果的专业人士，也可以把这本书作为有用的参考和评论材料。

写这本教科书的时候，我尽量做到全面而不是面面俱到。使用这本书的教师会发现，学生拥有这本书并不代表不再需要教师的指导，因为此书留给教师充分的空间来选择、解释、阐述和拓展。每一章后面的总结部分都有对本章材料的回顾和复习，并且还有问题和练习部分作为材料的扩展和补充。

《心理测量与评估》第 11 版的基本结构和过去的版本相似。熟悉早期版本的教师可以发现，尽管到处都有过改动，但是这些改动并非彻底的改变，因此他们仍然会感到很熟悉。使用过早期版本的读者很快也会注意到，这个版本比以往的版本有更多的章节（共 18 章）。前 5 章的结构和之前的版本非常相似，但是书中其他部分的材料则重新安排过了。第 10 版中的第六章到第九章以及第十三、十四章中的部分内容在第 11 版中变为 6 章（第六章到第十一章）的内容。第 10 版中第十一章的材料在第 11 版中被分成 3 章（第十四章到第十六章）的内容，第 10 版中第十二章的材料在第 11 版中被分成第十七章和第十八章。章节数的增多并不只是增加了新材料，更大程度上是由于将早期的章节分成更短小的部分、并在第六至十八章中吸收了第 10 版的第十三、十四章中相关内容的缘故。作者认为这样的重新安排是有意义的，能够促进读者学习和理解第六至十一章关于认知能力的材料，以及第十二至二十九章关于人格、兴趣、态度及其他相关的材料。

近几年来，教育和心理测量领域发生了一些值得注意的变化，此书对这些变化给予了足够的关注。这些变化是：大学入学考试在内容和形式上的改变，各种类型测验新版本的出现，以及对“智力测验的利弊”的探索。对于适应性测验、项目反应理论、微机在心理测量中的应用，神经心理和发展性测验，以及在各种背景下测验的应用方面，人们也给予了更多的关注。我们的目标是将心理测量与评估作为一个有趣的、重要的领域，介绍给打算进入任何一个编制或使用心理评估的专业领域的学生，为了有助于实现该目标，我们将新的重点放在测验在教育—学校背景、临床—咨询背景以及工业—组织背景中的应用上。

《心理测量与评估教师手册》第 11 版包括几百个多选题和是非题，可以从出版者那里获得（Allyn & Bacon, Pearson Education, 75 Arlington Street, Boston, MA 02116, 电话：848-6000）。与课程内容配套的、包含几十个关于心理与教育测验的应用程序的磁盘可以从作者那里获得。如果你需要该磁盘，请将你的 DOS 操作系统中格式化过的磁盘以及一个贴有邮票的回邮信封寄往下列地址：Lewis R. Aiken, Ph. D., 3300 Blue Bridge Court, Thousand Oaks, CA 91362。最后，本书配

套的学习指导也可以通过上述地址联系作者付费获得。也可通过电子邮件: laiken@prodigy.net 或传真: 805-370-8015。

关于此书我已经说了很多,就此搁笔吧,在这里我非常感谢在本书 10 次出版过程中付出辛勤劳动并给我提出中肯批评和建议的学生和同事,以及为第 11 版审稿的: Rochester 学院的 Angela Hazel; Concordia (肯考迪娅)学院的 William Mahler; 以及 Shorter 学院的 William Warley。我还希望表达我对发行编辑, Faye Whitney-Lussier, 及校订编辑 William Thomas 的感谢,感谢他们不懈的努力和精湛的专业技能。我真诚希望这些同仁和我的努力在这一成果中有所体现。并且非常欢迎读者提出有助于完善此书的评论和建议,我将非常感激。

刘易思·艾肯 (Lewis R. Aiken)



## 目 录

## 第一章 历史和专业问题

历史回顾 2

心理测验行业 7

测验标准和道德规范 12

小结 17

问题和练习 18

## 第二章 测验的设计和编制

制定测验计划 19

编制测验题目 25

合成和复制测验 35

口头测验 38

操作测验 40

小结 41

问题和练习 42

## 第三章 测验的施测与计分

测验的施测 44

测验计分 53

小结 61

问题和练习 62

## 第四章 项目分析和测验标准化

项目分析 64

测验标准化和测验常模 75

测验等值 82

小结 83

问题和练习 84

## 第五章 信度和效度

信度 87

效度 96

人事决策中的测验使用 101

小结 106

问题和练习 107

## 第六章 标准化成就测验

成就测验的基础知识 109

标准化成就测验的类型和选择 117

成套成就测验 120

特殊领域的成就测验 123

小结 132

问题和练习 133

## 第七章 智力测验

历史、定义和理论 135

个体智力测验 141

团体智力测验 155

小结 163

问题和练习 165

## 第八章 心理能力的个体与群体差异

智力落后、超常和创造性 166

心理能力人口统计学相关的研究 174

生物因素和心理能力 183

小结 188

问题和练习 190

## 第九章 发展心理和神经心理评估

婴儿和儿童早期的发展性评估 192

学习障碍 198

神经心理失调及其评估 200

小结 208

问题和练习 209

## 第十章 特殊能力的评估

特殊能力的概念 211

感知和心理运动测试 215

机械能力 219

文书和电脑相关能力 223

艺术能力与音乐能力 225

多重能力倾向成套测验 227

小结 235

问题和练习 236

## 第十一章 能力测验的应用和问题

教育情境中的评估 238

能力测验中的批评和问题 243

教育测验的其他问题 251

职业测验和偏见 257

小结 261

问题和练习 262

## 第十二章 职业兴趣

兴趣测量的基础 264

兴趣调查表的效度 267

斯特朗兴趣调查表 270

库德兴趣调查表 275

兴趣和人格 277

其他一般和特殊用途的兴趣问卷 283

兴趣调查表在咨询中的使用 285

小结 288

问题和练习 288

### 第十三章 态度、价值观和个人定向

态度的测量 290

价值观的测量 300

个人定向 303

小结 305

问题和练习 306

### 第十四章 人格评估的起源、应用和问题

伪科学以及人格评估的其他鼻祖 310

人格理论 312

人格评估的使用和误用 319

临床评估 323

人格评估应用的其他领域 325

人格评估中的问题和争议 330

小结 336

问题和练习 337

### 第十五章 观察与访谈

观察 339

传记资料 345

访谈 347

行为分析与评估 357

小结 359

问题和练习 360

## 第十六章 核查表与评定量表

核查表的特征 361

核查表的种类和样例 366

编制评定量表的策略 372

评定量表的种类 373

评定的问题 379

标准化评定量表 380

Q 分类与角色构念库测验 381

小结 382

问题和练习 383

## 第十七章 人格调查表

真实性、信度及效度 387

单一结构的症状量表 389

以内容效度为基础的多维调查表 392

应用因素分析的调查表 393

明尼苏达多相人格调查表 397

其他效标参照的人格调查表 406

小结 411

问题和练习 411

## 第十八章 投射技术

|                        |     |
|------------------------|-----|
| 词语联想与造句                | 415 |
| 墨迹测验                   | 419 |
| 主题统觉测验及其变式             | 422 |
| 其他统觉测验                 | 424 |
| 投射技术的问题                | 426 |
| 人格评估的前景                | 426 |
| 小结                     | 428 |
| 问题和练习                  | 428 |
| 附录 A 问题和练习答案           | 431 |
| 附录 B 正态曲线下的面积          | 435 |
| 附录 C 提供心理与教育评估工具的商业化机构 | 438 |
| 附录 D 与心理测量和评估相关的组织的网站  | 444 |
| 附录 E 测验索引              | 446 |
| 附录 F 词汇表               | 457 |

# 第一章

## 历史和专业问题

任何在过去半个世纪中曾就读过大学或学院、在军队服役、或应聘某项工作的人，都毫无疑问地参加过一个或多个测验。测验已在全世界范围，对人们的生活和职业方面产生了十分重要的影响。然而，心理评估工具，不仅仅局限于已公开发表的测验，还包括许多未发表的测验、已发表或未发表的问卷(questionnaires)、调查表(inventories)、评定量表(rating scales)和核查表(checklists)等都是可以使用的。

当需要信息来帮助人们制定决策或协助他们选择未来的教育或职业的课程时，就可能会使用某些形式的评估工具。测验和其他评估工具在学校、临床心理、工业以及国家和军队部门广泛运用，服务于诊断、评估、选拔、安置及升迁目的。这些测验除应用于实际的决策制定外，还大量应用于研究工作。

由于测验的诸多功能，其本身已成为一个庞大的商业领域也就不足为奇了。美国出版商联合会(The Association of American Publishers)公布的数据显示，仅2000年一年，用于从幼儿园到12年级的标准化测验，总销售额就大约有23 410万美元，并且该数字每年增长7%左右。附录C上所列的机构，是专门从事发行和销售各种心理测量工具的商业组织，这些工具可用于评估各年龄段和不同环境下人们的各种能力、人格、兴趣和其他个体特征。附录D上所列的专业组织则关注于如何依据不同的应用和研究目的正确使用各类测验。



## 历史回顾

自人类诞生以来，人们便认识到人类在认知能力、人格特性和行为上存在差别，并且发现能够通过某些方式对这些差别进行评估。柏拉图（Plato）和亚里士多德（Aristotle）在2 500年之前就阐述过个体的差异，但他们仍晚于古代的中国。早在公元前2 000年，就由中国皇帝建立了一种行政机构的考试系统，用于考核政府官员是否恰当行使其职责。每三年一次对官员们进行评测，评估他们对音乐、射箭、马术、书法、算术及公共和私人礼仪及庆典的熟练程度。这一系统，由中国后来的统治者保留下来，其内容涵盖了民法、军事、农业、税收、地理、文章和诗歌（Green，1991）。这些测试都是口试而非笔试，评估内容不仅包括被测者对问题的回答，还包括他们回答问题的方式。19 世纪，英国、法国和德国政府也模仿早期中国的测试系统制定了他们自己的行政测试系统。

从科学角度来讲，针对个体的考评在中世纪的欧洲几乎不存在。在中世纪欧洲的社会结构中，人们的活动由其出生的社会阶层所决定，个性的表达和发展受到极大限制。然而到了16 世纪，欧洲社会发生了较大程度的改革，传统教条逐渐弱化；人们越来越认识到个体是独特的，具有各自不同的天赋，并且有能力改善自己在生活中的地位。在文艺复兴和随后的启蒙（The Enlightenment）时期，不仅学习和创造的热情再度觉醒，并受到鼓励，个人主义同样也获得新生。个人自由和个体价值的精神，在资本主义和民主主义提供的政治和经济刺激下开始繁荣，表现在艺术、科学、哲学和政府工作的各个方面。但直到19 世纪晚期，对个体能力和人格方面差异的科学研究才真正开始。

### 19 世纪的心理测量

19 世纪早期，科学家们把在感知觉和心理能力上的个体差异视为一种误差来源。在精确、自动测量和记录物理事件的工具发明之前，科学测量时间、距离和其他物理变量的精确性在很大程度上都依靠观察者的知觉—运动能力。绝大多数观察者都经过专业训练，并在进行观察时十分认真，但是由不同观察者或同一观察者在不同情况下进行的测量仍然存在一些差异。当自然现象的测量不精确或不可靠时，就为发现自然普遍规律带来了困难，于是物理科学家们将关注点转移到制作更为准确和稳定的工具上。比如，由约翰·哈里森（John Harrison）与同事发明的计时钟，基本不会受轮船运动以及温度和湿度变化所带来的误差影响，可以精确地测量经度，减少了航海的危险（见 Sobel & Andrewes，1998）。