

Psychological Testing

History, Principles, and Applications

心理測驗

Robert J. Gregory 著

游恆山 譯



Psychological Testing

History, Principles, and Applications

《心理測驗》的目標是在教導學生關於心理測驗的特性、目標及廣延效應。

除了傳統主題的報導寬度，本書的特色是進一步呈現「神經心理與老人病評鑑」、「測驗的早期使用與濫用」、「學習障礙的評鑑」、「特殊背景的測試」，以及「IQ的種族差異」等主題。

作者描述及批判最廣泛使用一些測驗的最新近版本、檢視施測過程的微妙性，以及探索關於測驗智慧的一些價值方面的議題。

心理
測驗



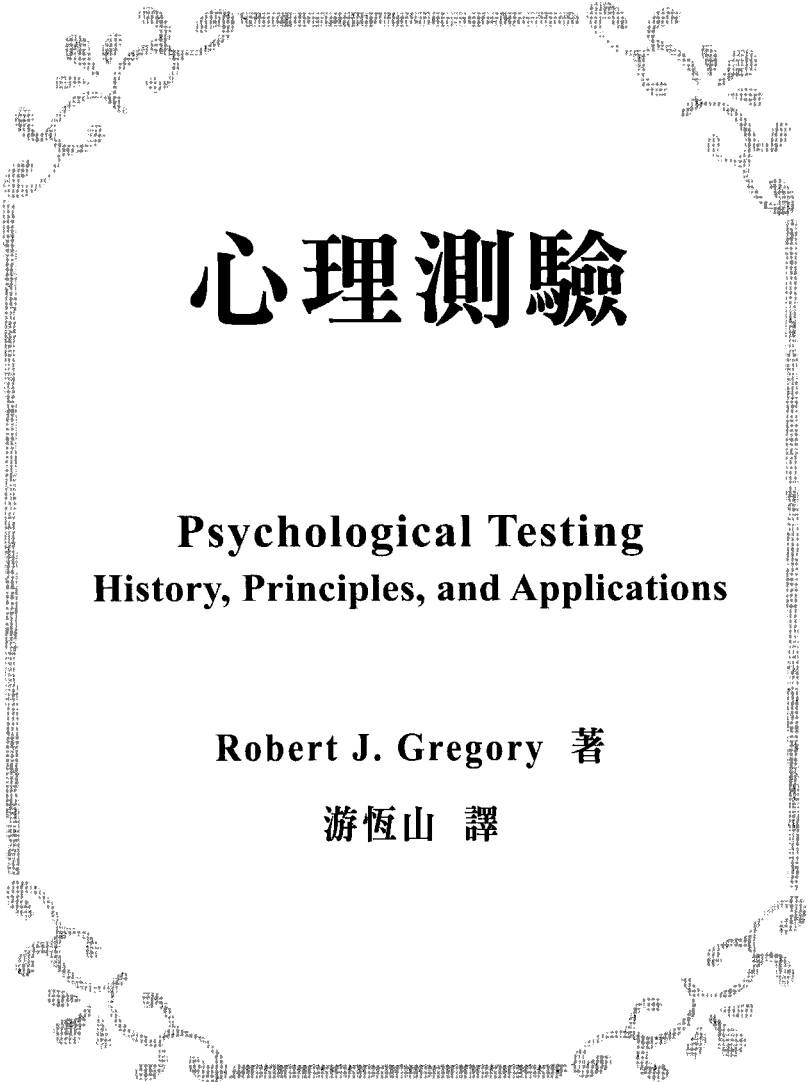
五南文化事業

ISBN 978-957-11-6055-9 (179)



9 789571 116055

五南圖書出版公司



心理測驗

Psychological Testing
History, Principles, and Applications

Robert J. Gregory 著

游恆山 譯

五南圖書出版公司 印行

國家圖書館出版品預行編目資料

心理測驗 / Robert J. Gregory 著 ; 游恆山
譯。

—初版。—臺北市：五南，2010.09.

面；公分

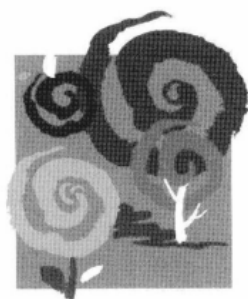
譯自：Psychological testing: history,
principles, and applications

ISBN 978-957-11-6055-9 (平裝)

1. 心理測驗

179.1

99014223



1BWF

心理測驗

作 者 — Robert J. Gregory

譯 者 — 游恆山

發行人 — 楊榮川

總編輯 — 龐君豪

主 編 — 陳念祖

責任編輯 — 李敏華 雅典編輯排版工作室

封面設計 — 童安安

出 版 者 — 五南圖書出版股份有限公司

地 址：106台北市大安區和平東路二段339號4樓

電 話：(02)2705-5066 傳 真：(02)2706-6100

網 址：<http://www.wunan.com.tw>

電子郵件：wunan@wunan.com.tw

劃撥帳號：01068953

戶 名：五南圖書出版股份有限公司

台中市駐區辦公室 / 台中市西區中山路6號

電 話：(04)2223-0891 傳 真：(04)2223-3549

高雄市駐區辦公室 / 高雄市新興區中山一路290號

電 話：(07)2358-702 傳 真：(07)2350-236

法律顧問 元貞聯合法律事務所 張澤平律師

出版日期 2010年9月初版一刷

定 價 新臺幣600元

Psychological Testing

History, Principles, and Applications

Robert J. Gregory

Authorized translation from the English language edition, entitled PSYCHOLOGICAL TESTING: HISTORY, PRINCIPLES, AND APPLICATIONS, 5th Edition, ISBN: 0205468829 by GREGORY, ROBERT J., published by Pearson Education, Inc. publishing as Allyn & Bacon, Copyright © 2007 Pearson Education Inc.

All rights reserved. No part of this book may be reproduced 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.

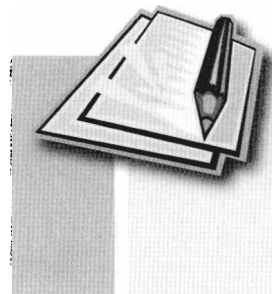
CHINESE TRADITIONAL language edition published by WU-NAN BOOK INC.,
Copyright © 2010

原著由 Pearson Education, Inc 於2007年出版

此繁體中文版為 Pearson Education, Inc 授權五南圖書出版

版權所有 翻印必究

序 言



心理測驗起始是一種畏縮不前的活動，發軔於 19 世紀一些歐州心理學家的學術實驗。從這個不順利的發祥以來，測驗的實踐以不斷加快的步伐擴散到整個工業化世界。如讀者在這本書的字裡行間將會發現的，心理測驗現在已實際上衝擊現代生活的每個角落，從教育到職業以迄於矯治。

如同前幾版，這本書第五版的抱負是在提供讀者關於心理測驗之特性、目的及廣延影響的知識。在追求這個目標上，我編入一些老生常談的傳統觀念，但是也前進到一些新的方向。例如，依據慣常的手法，這本書囊括了常模、標準化、信度、效度及測驗建構等一般主題。再者，我也在這些傳統領域中編纂及批判各式各樣的測驗和測量工具，像是智力、成就、工業—組織、職業及人格測試等領域。

除了前面所列的傳統主題外，我也強調另一些議題、論點及觀念。根據我的意見，它們是深入理解心理測驗的基本所要。例如，本書第二章檢視心理測驗的歷史。這一章的編排強調我的觀點，即心理測驗的歷史跟當今的實施具有實質的關聯。簡言之，對現代測驗的成熟理解只有透過探究它的傳承才能達成。當然，心理學學生往往規避歷史素材，因為這些論題經常是以單調、乏味而迂腐的手法呈現，缺乏跟現今的關聯性。然而，我希望帶有批判性的讀者將會以開放的心胸接近我的歷史章節——我已致力於使得它顯得有趣而切題。

心理測驗代表兩方之間的一種契約。一方是施測者，經常占有施加權力於另一方的位置；另一方則是受測者。基於這個原因，施測者需要以最大靈敏度處理測驗，注意受測者的需求和權益。為了強調這個關鍵點，我花費不少篇幅於論述施測過程的微妙性，包括像是「建立融洽關係」和「觀察對測驗結果之不當的環境影響力」等議題。此外，我也透過檢閱施測上的專業議題和倫理準則以強調衡鑑的契約本質。

這本書所強調的另一個主題是神經心理衡鑑，它是臨床心理學正在快速發展的另一個次場域，現在已穩固地建立起它的專業地位。神

經心理衡鑑絕對是一個成長的領域，它現在已構成當今心理測驗的重大用途之一。我已專致整整一章於這個重要主題。因此，讀者可以更良好理解神經心理衡鑑的範圍和目的。

這本書不僅是在論述測驗和它們的信度及效度，我也探索許多與測驗智慧有關之價值方面的議題。心理測驗之所以有爭議，正是因為施測結果可能具有傷害性，不僅是對個體而言，且或許對整個社會結構（**social fabric**）也是如此。我沒有迴避跟心理測驗使用有關的一些爭議，這些各自議題包括智力的遺傳與環境促成、IQ 之種族差異的起源、測驗偏差與效度之外的關涉、團體成就測驗上的作弊，以及心理施測的倫理議題。

總之，在這一版中，我的目標是 (1) 增添關於既定測驗之最新近的發現；(2) 編入在先前版本中被忽略，但值得探討的主題；以及 (3) 包羅施測上的革新及進展的報導。我相信我已達成所有這三項目標。



目錄

序言

第 1 章 心理測驗的用途與影響 1

主題 1A	心理測驗的本質與使用	2
一	測驗的影響	2
二	測驗的定義	3
三	測驗的進一步區分	6
四	測驗的類型	7
五	測驗的用途	11
六	影響測驗健全性的一些因素	13
七	測驗施行的標準化程序	14
八	測驗施行的適宜程序	15
九	施測者的影響力	17
十	受試者的背景及動機	18
主題 1B	心理測驗的倫理及社會意涵	21
一	專業測試準則的理論基礎	22
二	測驗發行者的責任	24
三	測驗使用者的責任	25
四	文化與語言方面少數族群的測試	32
五	再訪：負責的測驗使用	36

第 2 章 心理測驗的歷史 39

主題 2A	心理測驗的起源	40
一	測驗的重要性	40
二	測驗於西元前 2200 年在中國的初步形式	41
三	心理測驗的精神病學前例	41
四	面相學、顱相學及心理描述圖	42



五	測驗的銅管儀器年代	43
六	智能不足在 1800 年代的變動觀念	47
七	比奈的早期研究對他的測驗的影響	48
八	比奈和高級心智歷程的測試	49
九	修訂量表與 IQ 的來臨	50

主題 2B 美國的早期測驗 52

一	測驗在美國的早期使用和濫用	52
二	Howard Knox 與 1900 年代早期的非語文測驗	54
三	斯一比測驗：IQ 的早期台柱	56
四	團體測驗與第一次世界大戰陸軍新兵的分類	56
五	早期的教育測驗	57
六	性向測驗的發展	58
七	第一次世界大戰後的人格與職業測驗	59
八	投射測驗的起源	60
九	興趣量表的發展	62
十	測驗歷史上重要事件的摘要	63

第 3 章 常模與信度 65

主題 3A 常模與測驗標準化 66

一	原始分數	67
二	基本的統計概念	68
三	原始分數轉換	75
四	挑選常模組	83
五	標準參照測驗	87

主題 3B 信度的概念 89

一	古典測驗理論與測量誤差的來源	90
二	測量誤差的來源	91
三	測量誤差與信度	93
四	信度係數	94
五	相關係數	95
六	相關係數作為信度係數	97
七	信度作為時間穩定性	97
八	信度作為內部一致性	99

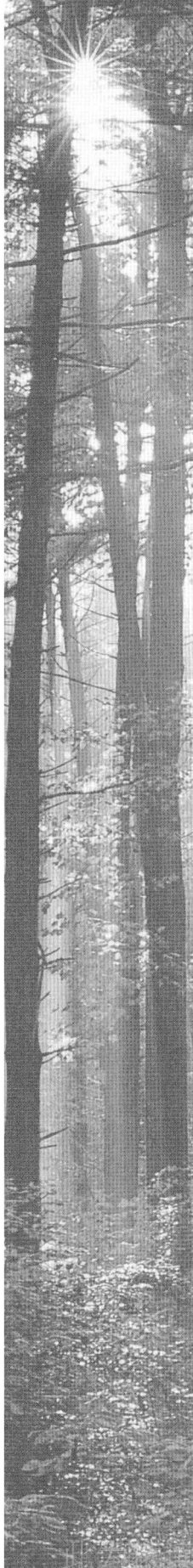
九	信度估計上的特殊情況	105
十	信度係數的判讀	107
十一	信度與測量標準誤	108

第4章 效度與測驗編製 113

主題 4A	效度的基本概念	114
一	效度：定義	115
二	內容效度	116
三	效標關聯效度	119
四	構念效度	127
五	構念效度的驗證法	128
六	效度之外的關涉和測驗效度廣延的範圍	137
主題 4B	測驗建構	140
一	界定測驗	141
二	選定量表編製的方法	142
三	具代表性的量表編製方法	144
四	建構題目	149
五	檢驗題目	153
六	修訂測驗	158
七	發表測驗	160

第5章 智力與成就的理論及個別測驗 163

主題 5A	智力理論與因素分析	164
一	智力的定義	165
二	因素分析的初階	168
三	高爾頓與感官敏銳性	173
四	斯皮爾曼和 g 因素	173
五	賽斯通與基本心理能力	174
六	卡泰爾—洪恩—卡洛 (CHC) 理論	176
七	皮亞傑與適應	179
八	吉爾福德與智力結構的模型	180
九	同時性與連續性處理的理論	181



十 智力作為一種生物構念	182
十一 葛德納與智力多元論	183
十二 史騰伯格與智力三元論	185

主題 5B	智力與成就的個別測驗	188
一	個別智力測驗的定向	188
二	魏氏智力量表	190
三	魏氏分測驗：描述與分析	192
四	魏氏成人智力量表—第三版	202
五	魏氏兒童智力量表—第四版	206
六	斯一比智力量表：第五版	209
七	考夫曼兒童評鑑組合—第二版	212
八	考夫曼青少年與成人智力測驗 (KAIT)	215
九	考夫曼簡明智力測驗 -2 (KBIT-2)	215
十	個別成就測驗	216
十一	學習障礙的本質與評鑑	218

第 6 章 團體測驗與能力檢測上的爭議 223

主題 6A	能力與相關概念的團體測驗	224
一、	團體測驗的本質、前途及陷阱	224
二、	能力的團體測驗	226
三、	多元性向測驗組合	237
四、	研究生甄選測驗	242
五、	教育成就測驗	244
六、	指導對性向測驗分數的影響	246
主題 6B	測驗偏差與另一些爭議	249
一	關於 IQ 的錯誤觀念	250
二	測驗偏差的問題	251
三	智力之遺傳與環境的決定因素	254
四	美國黑人與白人 IQ 差異的起源	260
五	智力的年齡變化	264
六	智力測驗分數的世代變動	269

第 7 章 施測特殊人口 273

主題 7A	嬰兒與學前的評鑑	274
一	嬰兒能力的評鑑	275
二	學前智力的評鑑	279
三	嬰兒與學前評鑑的實際用途	284
四	入學準備度的篩選	286
主題 7B	施測失能人士	291
一	特殊人口之測驗的起源	291
二	非文字測驗	292
三	非閱讀式與動作簡化的測驗	297
四	施測視覺障礙的人士	299
五	智能不足之適應行為的評鑑	300

第 8 章 人格測試的起源 307

主題 8A	人格的理論與測量	308
一	人格：概論	308
二	人格的精神分析理論	309
三	人格的類型論	313
四	人格的現象學理論	315
五	行為與社會學習理論	317
六	人格的特質構思	319
主題 8B	投射技術	323
一	投射假說	323
二	投射技術的初階	324
三	聯想技術	325
四	完成技術	329
五	編造技術	331
六	表達技術	334



第 9 章 人格的結構式評鑑與其他特性 337

主題 9A	自陳式量表	338
一	理論引導的量表	339
二	因素分析推衍的量表	346
三	效標導向的量表	351
四	態度與態度的評鑑	359
五	道德判斷的評鑑	363
主題 9B	行為與觀察的評鑑	366
一	行為治療的基礎	367
二	行為治療與行為評鑑	367
三	結構式晤談量表	374
四	系統化直接觀察的評鑑	376
五	非語文行為的評鑑	379
六	模擬的行為評鑑	382
七	生態瞬間的評鑑	383

第 10 章 神經心理評鑑與篩選工具 385

主題 10A	神經生理學概念的初階	386
一	大腦的構造與功能	386
二	神經病變的起因	388
三	腦部顯像技術	392
主題 10B	神經心理測驗、測驗組合與篩選工具	396
一	大腦—行為關係的概念模式	397
二	感覺輸入的評鑑	398
三	注意與專注的測量	399
四	學習與記憶的測驗	402
五	語言功能的評鑑	407
六	空間與操作能力的測驗	409
七	執行功能的評鑑	413
八	運動輸出的評鑑	415
九	神經心理評鑑的測驗組合	417

十 酒精使用疾患的篩選	419
十一 老年人心理狀況評鑑	421

第 11 章 特殊性向測驗與興趣的測量 425

一 特殊性向測驗	426
二 興趣的測量	430

第 12 章 電腦化的評鑑與測驗的未來 439

一 電腦在施測上的應用：縱覽和歷史	440
二 電腦本位的測驗判讀：現今的情況	442
三 高解析度的影像和虛擬實境：CAPA 的新視野	450
四 電腦本位測驗判讀的評估	452
五 電腦化適應的施測	455
六 測驗的未來	456

附錄 459

A. 心理測驗歷史上的一些劃時代事件	459
B. 美國主要測驗出版機構的地址	462
C. 主要的一些測驗及它們的出版社	464
D. 在常態分配中，幾種常見的標準和標準化 分數與百分等級的對應關係	469

英文書幾乎都附有厚厚的參考資料，這些參考資料少則 10 頁以上，多則數十頁；中文翻譯本過去忠實的將這些參考資料附在中文譯本上。以每本中文書 20 頁的基礎計算，印製 1 千本書，就會產生 2 萬頁的參考資料。在地球日益暖化的現今與未來，為了少砍些樹，我們應該可以有些改變——亦即將英文原文書的參考資料只放在網頁上提供需要者自行下載。我們不是認為這些參考資料不重要，所以不需要放在書上，而是認為在網路時代我們可以有更環保的作法，滿足需要查索參考資料的讀者。

我們將本書【參考文獻】放在五南文化事業機構（www.wunan.com.tw）網頁，該書的「資料下載」部分。

對於此種嘗試有任何不便利或是指教，請洽本書主編。

第 1 章

心理測驗的 用途與影響

Psychological Testing

主題 1A

心理測驗的本質與使用

- | | |
|------------|----------------|
| 一、測驗的影響 | 六、影響測驗健全性的一些因素 |
| 二、測驗的定義 | 七、測驗施行的標準化程序 |
| 三、測驗的進一步區分 | 八、測驗施行的適宜程序 |
| 四、測驗的類型 | 九、施測者的影響力 |
| 五、測驗的用途 | 十、受試者的背景及動機 |



假使你問一般民眾「你對心理測驗知道些什麼？」他們可能提到一些事情，像是關於智力測驗、墨漬測驗，以及廣為熟悉的 MMPI 這類「真－偽量表」等。大部分情況，他們對測驗的理解將是集中在「量化智力」和「偵測人格問題」上，因為這是關於測驗在我們社會中如何被使用的一般觀點。當然，測驗所涵括的不只這少許事實，人格和智力的測量依然是心理測驗不可或缺的台柱。然而，現代測驗編製者已針對多元化而富有想像力的目的開發許多種類的測驗，有些甚至是測驗的早期開拓者所預期不到的。這一章的目的是討論心理測驗的多樣化用途，也審視這個行業的倫理議題和社會影響。

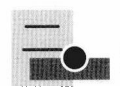
一 測驗的影響

從出生到老年，我們在生活的幾乎每一個轉捩點都會遇到測驗。嬰兒在出生後不久接受的第一份測驗是「Apgar 測驗」，它是對心跳、呼吸、肌肉張力、反射感應性及膚色等施行快捷而多變項的衡鑑（Clarke-Stewart & Friedman, 1987）。整體 Apgar 分數（從 0 到 10）有助於決定是否需要任何立即的醫療關注。稍後，先前在 Apgar 測驗上拿到低分的幼兒可能需要接受發育障礙的衡鑑。學前兒童可能接受「就學預備性」測驗。一旦學業

生涯展開，每位學生在畢業之前需要承受數以百計（或許千計）的學術測驗，更不用提關於學習障礙、資賦優異、職業興趣及大學入學等可能測驗。畢業之後，成年人可能面對工作應徵、駕駛執照、安全許可、人格功能、夫妻相處、發展失調及大腦機能障礙等測驗——幾乎無止境的一份名單。有些人甚至在晚年面對自己脆弱的一項無禮對待，即透過測驗決定他們管理財務的行為能力。

測驗在全世界的幾乎每一個國家被派上用場，像是諮詢、甄選及安置（placement）等用途。測驗發生在各式各樣的環境中，像是學校、行政機關、公司企業、醫療診所及諮商中心等。心理測驗結果以一些深切方式改變個人的命運。個人是否能夠申請進入一流大學？是否被提供特定工作或職位？是否被診斷為憂鬱症病人？所有這樣決定至少部分地取決於測驗結果的意義——隨著位居能行使權力之地位的人們所進行的解讀。簡而言之，心理測驗結果改變了生活。基於這個原因，慎重的做法是（事實上，幾乎是強制性的），施測人員應該認識心理測驗的當代用途和偶爾濫用。在個案提示 1-1 中，它透過幾個真實的個案史說明心理測驗改變生活的餘波。

測驗的概念因此是我們文化的一個普遍元素，我們視為理所當然的一項特色。然而，一般人關於測驗的觀念不必然符合心理計量學家所持較嚴格的觀點。心理計量學家（psychometrician）是心理學或教育學上的一種專家，他們編製及評鑑各種心理測驗。



測驗的定義

測驗（test）是摘取行為樣本的一套標準化程序，然後以類別或分數描述該行為。此外，大部分測驗擁有常模或標準，以便測驗結果可被用來預測另一些更重要的行為。我們在接下來的篇幅將詳盡闡述這些特性。

測驗在格式和用途上有很大變動範圍。儘管如此，大部分測驗擁有這些界定的特性：

- 標準化的程序。
- 行為樣本。
- 分數或類別。
- 常模或標準。
- 對非測驗行為的預測。

我們打算檢視的這些特性將特別是針對常模參照測驗（norm-referenced tests），這樣測驗使用良好界定的個人母群作為解讀架構。然而，對於標準參照測驗（criterion-referenced tests）的特殊案例而言，它們的界定特性稍微不同。這樣測驗測量的是個人