

萬 有 文 庫

第一集一千種

王 雲 五 主 編

教 育 測 驗 概 要

華 超 著

商 務 印 書 館 發 行

教育測驗概要

華超著

師範小叢書

編主五雲王

種千一集一第

要概驗測育教

著超華

號一〇五路山寶海上
五雲王

人 行 發

路 山 寶 海 上
館 書 印 務 商

所 刷 印

埠 各 及 海 上
館 書 印 務 商

所 行 發

版 初 月 四 年 十 二 國 民 華 中

究 必 印 翻 權 作 著 有 書 此

The Complete Library

Edited by

Y. W. WONG

OUTLINE OF EDUCATIONAL MEASUREMENT

BY HUA CH'AO

PUBLISHED BY Y. W. WONG

THE COMMERCIAL PRESS, LTD.

Shanghai, China

1931

All Rights Reserved

自序

這本小冊子是根據麥柯爾博士的教育測驗 (How to Measure Education) 的上半部和參考孟羅的教育測驗理論編譯而成的。目的在介紹教育測驗的理論和應用方面。講到測驗的怎樣編造及種種統計的方法都是見教育測驗的下半部，本書並沒有把他介紹。我所以敢於這樣「斷章取義」有兩種原因，第一因為原著的性質是可以這樣分開，第二是因為教育測驗介紹到本國來雖然有好幾年的歷史，但是他的理論和實用方面還沒有得到一般從事於教育的人的深切瞭解。所以想於這方面多介紹些文字，希望多數人明了測驗的真面目以後可以盡量的應用他。

測驗的高深學理——無論是教育測驗或智力測驗或特殊能力（或傾向）的測驗——在現在都還是一種假定，和物理化學方面的最高假定相同，沒有一種實驗可以證明。所以在測驗這門學科最發達的國家如美國，有權威的教育家還時有反對論調發生。在我們教育幼稚的國家裏頭要想推行這種新式工具，自然是更不容易。我這本小冊子倘能對於內地的教育界多少激起一

點研究的興趣，那末我所化於編譯這本小冊子的勞力就可算獲得了相當的酬報。

一九，一二，五，華超識於上海。

教育測驗概要

目次

第一章	測驗在教育上之地位	一
第二章	測驗和分別班級	一五
第三章	測驗在教育診斷上之功用	三九
第四章	測驗在教學上之功用	七五
第五章	測驗和教學效率的估量	一〇二

教育測驗概要

第一章 測驗在教育上之地位

理論一 凡存在的東西都是有量的

學生成績的測驗，內容很是複雜；但是也有他的基本的意義。倘不把這基本意義認清楚，便容易發生兩種教育的不幸事件：一種是過度的熱心於教育之科學的測驗，一種是過度的反對這種測驗運動。現在把這種基本意義歸納爲若干種理論，以便一般沒有充分時間來研究測驗的人可以藉此認識測驗在教育上之真正地位。

第一種理論，就是凡存在的東西都是有量的。這個理論，無須加以多大說明，因爲有健全頭腦的人都是承認的。但是爲明顯起見，不妨再說一說。譬如一個教師教了學生，學生所起的變化一定是有量的。做教師的極少願意承認自己的努力一點沒有產生量的變化。若果他承認是這樣，則徒

然證明他對學生的教學是沒有效率，而我們所持的理論卻仍舊可以成立。若說教育的努力，在性質之改善，這話是有一部分的真理，但是要知道沒有一種性質是屬「非量的 non-quantitative」。

理論二 凡是有量的都能够加以測驗

這個理論，也無須多大的解釋。譬如一個人對於日落時景象的愛好，對於音樂的鑑賞，是多少可以測驗的。又如一個人的創發力，比較價值的評判，領袖才能，以及對於文學的鑑賞等等，——這都是教育努力的目標——也是可以測驗的。我們目前的問題不是這個理論之能成立和不能成立的問題，乃是在這個理論下我們的測驗能不能比較以前主觀的評量更為準確，更為可靠。好在科學一天進步一天，測驗中所使用的工具，所謂 *test* 者，各方面已證實比主觀的評價準確得多了。

理論三 通常說來教育上的測驗是和物理科學上的測驗相同的

這兩種測驗，基本上是相類似的，因為兩者都是測驗物理的表現 (*physical manifestation*)。加減能力的高下和動機的好壞，都不是把寒暑表插入學生的精神媒介體中所可測驗的，但是

我們可以由他的外表的行爲和他的內在的情形，加以測驗。現在心理學家可以由呼吸曲線 (breathing curve) 來斷定證人說的話可靠不可靠，除非這個證人是慣於說慌的。

在最遠的將來，我們對於學生的心理活動的機制或者可以攝成影片，但是就現在而論，只得由心理活動所產生的結果來測驗；遇到必要的時候，再由結果來推想學生的內在情形。

理論四 凡物理的科學中的測驗都不是完全的

物理的測驗，通常是比教育的測驗較為準確。物理科學中所用的測驗工具，儘有是現在教育測驗專家所望塵莫及的。但是這種優秀點，並不是由於兩類測驗性質上根本不同而起。物理測驗實際上往往也免不掉教育測驗中所常犯的錯誤。不過他們的錯誤，數量極小，致通常人不能辨別出來罷了。不但物理科學是這樣，一切科學中都有錯誤，其中竟有爲專家所不能超脫的。桑代客 (E. L. Thorndike) 曾說：①

『這些教育結果的測驗和物理的測驗（如大小、密度、速度、溫度以及電量等等）的互相對比，毫不足奇。因爲溫度的零點的確定不過幾年以前的事，而溫度表的單位的定奪也有賴於多種

內在的，極精微的假定。我敢擅斷說一句，這種假定任何人都不能很清楚的很精細的說個明白。在現在，我們的教育結果的測驗，固不能作為讀法教學法或算術教學法等等的取舍標準，但是其中是有許多已比主觀的評量如法庭判斷某一商標是侵害他一商標權的評量優勝得多了。」

○見美國全國教育研究會年報第十七卷，第二編，第十六面。

但是教育測驗中的缺點，通常說來總比物理學，化學以及其相類似的科學中的缺點是顯著些。有些感情用事的人極端相信標準測驗的完全，是不合於真理的。標準測驗的功用是有許多確定的限度。所以要發起測驗的宣傳運動，與其說是因為測驗的十分完美，不如說是有鑒於現行各種測驗種類不多或尚有缺點須加改良。

教育測驗通常遇到一種反對，就是說僅僅測驗了學生的全部能力的一部分。就這點而論，物理的測驗更不如教育測驗。大多數物理的測驗如長度，廣度，重量和溫度等等的測驗都是極狹隘的，但是不失其為有用。對學生全部能力的測驗固然一定是有用的，但是倘使我們有了這種全部能力的測驗我們也必然進而為能力之分析的測驗。譬如作文測驗是有用的，但是分析作文的要

素的量表也是一定有用的。做教師不但從「全體」上教授學生，他也從小處着想。倘使測驗是能扶助教學的，則全部能力測驗和分析的（或逐步的）測驗是同等有用。

理論五 測驗是科學的教育之發展上所不可少的

各種自然科學所以有極迅速的進步，就是由於有準確的測驗。有人說過，肥皂之使用量是一國文化之指數。測驗之準確度，也可以說是某一種科學之指數。試問沒有密達，沒有格蘭姆，沒有安丕（電流之單位），沒有弗打（電動力之單位），沒有歐穆（電阻力之單位），沒有瓦特（力之單位），七四六瓦特為一馬力，科學還成個什麼樣子。教育之被擯於科學範圍以外，固有種種原因，缺乏相當測驗實是其中原因之最最大的一個。我們要把教育發展成為一種科學，是非謀測驗工具之進步之準確不可。

列却特 (Richards) 在美國科學改進社 (American Association for the Advancement of Science) 的演說辭中，曾有下列一段話：『好多年以前，柏拉圖已經承認沒有權衡和尺度無論何種藝術便空無所有。近代科學所以着重於量的張本就是這種注重權衡的返響；一種科

學於量的上面格外準確，就變爲更科學的。』

果真，測驗之和教育如同一對學生的孩子，這對學生子的頭髮已被母親編結在一起，他們無論那一個不能單獨行動只能有共動的作爲。

有些人以爲教育苟發展而爲一種科學，將侵害生活中的情緒及優美的知覺。這種見解正如恐怕提倡了婦女參政權或工業中雇用了女子，將毀損男子的勇敢一樣可笑。要知道這些優良的性質在人性中已根深蒂固，決不是科學所能破壞。科學之基本的功能在幫助我們以最簡捷最經濟的方法到達我們的目標，不論我們的目標是物質的或是精神的。科學之所以每被人看做是唯物的，主要原因就因爲只有那探求物質目標的人已善意的去獲得了科學的幫助。幾百年以來，教育上的變化真是千端萬緒。教育上有許多大思想家從柏拉圖，裴斯泰勞齊，福祿培爾以至現在各國的教育名家。舊秩序去了，新秩序起而來替代。但是似乎沒有人知道究竟舊的好還是新的好。就事實而論，竟有人懷疑以爲我們是在一個圓圈中行動。這些教育領袖並沒有解答了問題。他們所探討的問題，在別人心目中並沒有發生。他們只提起了預備試驗的問題。各個教育問題之最後的

解決，一定要交託給教育測驗的專家，一定要等到教育發展而成爲一種科學。

○見一九一九年正月三日科學 (Science) 中 The Problem of Radioactive Lead 篇。

理論六 除測驗外教育中尚有幾種不可忽視的動因

鼓吹教育之科學的測驗的人竟承認這個理論，這在旁人看來，大可驚怪。其實不然。教育中至少還有三種重要的動因，就是（1）學生，（2）方法和材料，（3）目標。做教師的必須明瞭學生的心理，訓練學生時所取之適當目標，以及使學生從開始能力 (initial ability) 到達希望的目標時所必須使用的適當的方法和材料。測驗對於這些教學過程，有些什麼關係，試讀下文便知。

理論七 倘使學生的開始能力是不能加以測驗的則我們對於學生是無從瞭解

上文已經說過，做教師的對於學生須有一種最直接的可能的知識，以便應用適當方法和材料去幫助他極迅速的達到所希望的目標。倘使只知道學生所有的能力是什麼，我們只是部分的了解學生罷了。部分的了解，只知道學生有某種能力，並沒有知道某種能力之量有多少。他這種能力是够得上得博士學位麼，還是僅够進幼稚園學習，我們都無從說起。所以我們不但要知道某

種性能的存在，並且要知道這種性能在何等量數以內存在，知道得愈準確愈好。這樣看來，測驗對於心理學之實際知識是必不可少的。

理論八 不可把捉的目標是沒有價值的目標

關於教育目標，我們至少要解答三種事情，第一就是，目標的價值是怎樣？第二就是這個目標的地位是怎樣？第三就是學生的行動是趨向於目標還是和目標背道而馳？這三個絕對重要的問題的解答，都離不了測驗。現在假定有個教學目標，在使學生產生書法能力。這個目標是完全不可捉摸的，除非確定了他的地位。就是要確定了書法能力究竟要多少，書法的速度是怎樣，書法的品質是怎樣，有了一個地位這個目標纔是能把握的目標。更有進者，沒有確定的地位，目標的價值也無從斷定，因為價值隨量之多少而不同。「要多少」這三個字實含有測驗的意義和須要測驗的意義在裏頭了。至於學生的行動是否合於目標這個問題，尤其非測驗不能解答。

理論九 各種教學方法和材料的價值是不知道的除非測驗了他們的效力

方法和材料之主旨固然在幫助學生向着目標進行。但是這個方法和材料是否實踐了這種

目的，我們要施行測驗之後纔可定奪。有時候一種方法我們以爲是很有益的，就決定採用，但容或可以是弊端百出。譬如採用了一種方法，這種方法不但不引致學生向目標進行，並且足使繼任的教師採用別種方法時發生困難。好幾種方法和材料集在一起，而要下一個比較價值的評判時，也非使用測驗，決定每一方法每一材料對於學生的效力不可。這就是說教育中對於各種最經濟的教育情景之實驗的選擇都離不掉測驗。

教育中和日常生活中，到處都是測驗。測驗不是新奇的或突如其而來的事情。從早上起來到晚上睡眠的時候，測驗都伴隨我們。指示路程遠近的標誌，錶的指針，廚房中的不值錢的杯子，草地上樹影子的延長，一年四季的遞傳，凡此種種都是表示測驗之無所不在。教育的全過程中有測驗，生活中也無處不是測驗。教育中除測驗以外固然還有別的東西，但是這別的東西和測驗脫離了關係，便是沒有價值的。

理論十 成績測驗應在教育視察之前

初等教育之主要目的在使學生產生某種變化。爲達這個目的起見，所以我們有種種工具，如

書本、圖畫、地圖、玩具以及各種教授方法，聯同使用這些工具的教師。這些工具真的能產生我們所希望的變化麼？此地我們有兩種測驗的方法。一種是間接法，就是測驗某種書本之應用，或某種圖書儀器之設備，或某種教授能力（如激起學生動機或使學生有創發性之能力），或某種專業訓練（指教師說）之存在是否於學生的變化爲不可少的條件。不可少的條件不知道，則我們就無從斷定不可少的條件多滿足了，自然能產生所希望的變化。一種是直接法，就是直接測驗學生的變化，換句話說就是測驗學生的成績。間接法可以確定教學法，教學工具和學生變化的因果關係，但是迂緩而費時直接法是最經濟又最合實用。所以舉行教育視察之前應當先施行成績測驗。

理論十一 測驗不是教育上新奇的事情

從反對測驗和極力宣傳測驗兩方面看來，測驗好像是絕對新奇的事情。其實不然。教育上的測驗並不是突然而來的，是有他的逐漸的演化。曾有人說過：從事實的立場說，教育測驗是古代的；從過程的立場說，教育測驗是中世紀的；從科學的立場說，教育測驗纔是近世的。這幾句話已把測驗的演化概括無遺了。幾百年以來，中國的考試制度和近來的學校考試都是一種測驗。不過他們

都是些因襲的形式的考試，並沒有準確的標準足以和現在教育測驗比擬罷了。就是論標準測驗的產生也不是一朝一夕的事。以美國而論，從拉哀斯（Rice）博士於一八九四年把許多學校中的教學結果做一個比較的測驗之後，方纔產生測驗中的所謂常模（norms）。桑代客教授利用客太爾，福爾董的等距理論（Cattell-Fullerton equal-distance theorem）方纔產生教育成績測驗中的一種量表單位（scale unit）。科學的教育測驗從此有了一個發端。一九〇八年斯東（Stone）氏的算術測驗公表於世；一九〇九年桑代客又發表書法量表。隨後桑代客的學生對於統計技術和教育量表的構造又多所研究。於是希來格斯（Hillegas）勃金漢（Buckingham），脫拉別（Trabue）和華迪（Woody）等等又各自造成作文量表，拼法量表，語言量表，和算術四則測驗。一九一一年至一二二年紐約市的學務調查全用克的斯（Courts）的算術測驗，頗為時人所信從。從此以後，測驗的應用範圍愈加擴大，幾乎凡學務調查都信賴測驗是估計學校效率的主要工具。多數大學設有教育測驗一學程。有幾個大學且特設研究部和各學校共同研究教育測驗。著名的洛克弗樂基金社（Rockefeller Foundation）對於教育測驗的運動又極力協助。到了美