

教育評價概論

橋本重治 著

橋本重治 著
橋本重治 著
橋本重治 著

臺灣書店發行

教育評價概論

橋 本 重 治 著

侯 靜 遠 譯
李 聰 明

臺 灣 書 店 印 行

中華民國六十三年元月初版

教育評價概論

基 價：壹元貳角陸分正

著 者	橋 本 重 治
譯 者	侯 靜 遠 李 聰 明
印行者	臺 灣 書 店

地址：臺北市重慶南路一段十四號

電話：三 七 八 一 二 〇

郵撥：七 八 二 一

版權所有・翻印必究

目 次

原序

譯言

第1章 教育評價的意義	1
I 教育評價的歷史	1
II 評價與測驗	4
III 教育評價的功能	6
IV 教育評價的範圍	8
第2章 評價的程序	11
I 評價目標的分析	11
II 評價資料的搜集	13
III 評價資料的處理與解釋	17
第3章 評價的用具 (1)	21
——教師製作的測驗——	
I 論文式測驗	21
II 客觀測驗	23
III 問題情境測驗	30
IV 問卷法	32
V 人物推定法	34
第4章 評價的用具 (2)	37
——標準測驗——	
I 標準測驗的意義與種類	37
II 標準測驗的製作方法	39

■ 標準測驗的基準及其結果的解釋方法	41
IV 標準測驗的選擇利用注意事項	46
第5章 評價的用具 (3)	49
——觀察法及其他技術——	
I 觀察法	49
II 評定法	52
III 軼事記錄法	56
IV 接談法	57
V 投射法	59
第6章 智能與性向測驗	61
I 智能的意義與分析	61
II 智能測驗的用具	63
III 實施智能測驗應注意事項	66
IV 智能測驗的利用	67
V 性向測驗	71
第7章 學習的評價	76
I 評價目標 (學力) 的分析	76
II 各種學力的評價方法	80
III 單元學習的評價方法	89
IV 標準學力測驗	92
V 評價結果的處理與解釋	93
第8章 人格的評價	99
I 人格評價的目標	99
II 性格測驗法	102
III 行為與道德的評價方法	108

IV 個案研究法	111
第9章 教育診斷	114
I 教育診斷與學業不振	114
II 缺陷(症狀)的診斷	115
III 妨礙學習進步的原因	117
IV 原因診斷的方法	119
V 矯正與預防	123
第10章 學校評價	128
I 學校評價的目的及其由來	128
II 評價目標的分析	129
III 評價的方法	131
IV 結果的整理與解釋	135
第11章 教育統計	138
I 次數分佈表與分佈圖	138
II 代表值	140
III 標準差	142
IV 常態分佈曲線	143
V 百分等級	148
VI 相關關係	148
VII 推定與檢定	153
第12章 指導要錄與通訊簿	161
I 指導要錄的意義	161
II 指導要錄的理論基礎	162
III 指導要錄的評定方法	164
IV 通訊簿的功能	166

V 通訊簿的內容.....	168
---------------	-----

附 錄

小學校（國小）兒童指導要錄

中學校（國中）學生指導要錄

第 1 章 教育評價的意義

I 教育評價的歷史

所謂教育評價，係自本世紀黎明以來的數十年中，對教育測驗急速發展的又一個新的里程。其發展的過程，正和教育思想的變遷，以及心理學的發展史等，相為呼應。

(1) 教育測驗運動

用口試或個別談話方式來判斷教育效果，乃自有教育歷史以來，共同嘗試的現象，後來才逐漸發展到用筆記測驗。筆試在歐美教育史上，則是屬於近三百年以來的事。英國的劍橋大學於 1702 年實行 Bentaly 制，是為筆試的濫觴。美國波斯頓市教育委員會曾於 1845 年採用此制，以測驗該市的教育效果。但筆試法極為主觀，不可信賴，於是十九世紀後半葉起，開始作各種研究，以證明使用筆試法的不盡妥當。

企圖促使試驗方法客觀化的先驅，為英國的費雪 (Fisher, G. F) 及美國的萊斯 (Rice, J. M)。費雪在 1864 年設計一種量表簿 (Scale Book)，評定自己學校裏學生的答案與作品的成績，儘量改良他的記分方法，使之成為客觀的分數，以資改良教學方式。萊斯在美國派爾芝麻市當教育局長時，為適應時代進步的需要，改良教育課程，編訂豐富的教材，於是對於當時花費時間最多的作文、計算、英語等課程，認為極須要減少學習時間，以便增加科學與自然科等教材。他對於此事加以研究之後，為測量這種種科目究應減少多少時間才為適合，乃設計了一種綴字測驗 (Spelling Test)，是為測驗的創始。因此之故，後來即有人稱他為測驗的創始人。萊斯實驗的結果，每天須學習卅分鐘的綴字練習，即使減少練習時間至 15 分鐘，亦不致影響學生的成績，於是提出減少綴字教學時數的科學證據，呼籲改善科學教材的合理分配。此一事件，正說明教育測驗是為解決教育

問題的科學方法，表示教育測驗乃是屬於教育研究與教育改進的手段，已成為歷史的事實，堪為吾人注意。

關於教育方面客觀性測驗的蓬勃展開，乃是進入本世紀以來的事，亦就是說進入廿世紀之後，教育上有待研究改進的問題漸多，例如升學指導，升降級，教材教法之選擇，教育課程，能力分組，教育診斷，治療教育，等種種問題，都感覺到必須要有正確的科學理論，於是漸漸刺激教育界對於教育測量與客觀考試方法的研究發展。

美國的桑戴克 (Tnorndike, E. L.) 於1904年發表「精神與社會測驗學導論」一書，即是為測驗學上劃時代的鉅著，之後桑氏又與其門弟共同研究教育效果的數量測量，於1918年發表。在其所著“The nature, purpose and general methods of educational products”論文中曾說：「凡是存在的東西，必然含有數量」又說：「含有數量的東西，必定可以測量。」這就是桑氏的教育測驗信條。此後他相繼制訂並發表了種種標準測驗與評定量表，造成教育測驗運動 (Measurement movement) 的潮流，後人因而稱他為教育測驗之父。

以後桑戴克一派的教育測驗運動成果，越來越多，如斯騰 (Stone, C. W.) 於1908年，柯地氏 (Courtis, S. A.) 於1909年，相繼發表算術標準測驗。桑氏本人又於1909年發表習字量表 (Riting Scale)。此後在美國即逐漸流行標準測驗 (Standard Test)，迨至1928年，即斯騰算術測驗發表後之20年，在美國已經製作了近千三百多種的教育測驗。

又當時一般教師所採用以考試學生成績的論文式考試，評分極不客觀，史達克 (Starch, D.) 和耶里歐得 (Elliott, E. C.) 曾將一份幾何考試的答案拿給116位教師評分，其結果最高評82分，最低評28分，證明教師所採的考試評分之不客觀。

於是教育測驗運動更進一步影響到教師對於考試方法以及評分法的反省與改進，遂有教師自製的客觀測驗 (Teacher-made obj-

ective Test)之普遍發展。這就是1920年，麥可爾(McCall, W. A.)所倡導的新法考試。

以上所述，乃是以學力測驗為主的教育測驗，此外1905年法國的比納(Binet, A.)與西蒙(Simon, T.)所竭力倡導的智能測驗，同時傳入美國，遂確立了教育測驗的基礎。

(2)由測驗到評價

引起教育測驗運動的中心思想，不外想將傳統上不客觀的考試方法加以改良，使之成為客觀化。凡是不能持有客觀資料與科學基礎的教育計劃或教育改革，都是非常危險的，從這一點來看教育測驗運動的動機，不可謂不正確。

但是，新的心理學說的發現，與新教育運動的發展，又給予測驗運動一大轉機。新的心理學構造觀，對於人格測驗的片斷性，加以懷疑。新的心理學或新教育學，認為人類教育非僅知識教育的一面，應當包括知識的、社會的、情緒的、身體的，以及統一的全人格之教育為目標，測驗不過是對於含有數量性的知識分數的測量，如不從根本上打破偏頗的分數觀念，勢必不能適應新教育理想。尤以測驗的最大缺陷，即在「教育測驗根本未曾意識到教育目的，教育理想，甚至教育價值，這許多方面」。教育本身有其目的與理想，是則無論測驗也好，考查也好，都必須要密切地配合教育目的，教育理想，從根本上加以考慮才有意義。

這就是由測驗到評價(Evaluation)的根本構想。這一觀念1930年在美國當時與新教育運動同時展開。

從測驗運動到評價的新發展，其中為評價運動確立基礎的美國進步教育協會所進行的八年研究(The Eight-Year Study)及其所獲致的功績，不能不大書特書。八年研究就是從1933—1940年在教育的內容和評價方法上的一大實驗。為使新的觀念——教育評價獲致實效，由當時有名的泰勒(Tyler, R. W.)教授領導研究實施。結果證明教育評價，乃是「教育的科學經營與教育的科學研究所必須的惟一手段」。

I 評價與測驗

教育評價與教育測驗，兩者之間究竟有何關係。爲明瞭此點，對於教育評價與測驗等種種概念，首先須有加以明確介紹的必要。

(1) 教育評價的意義

前述由教育測驗到教育評價的演變過程，已足以引人覺察教育評價的意義。本來，所謂教育事業，即是有目的的經營，有價值有理想。教育的目的或價值，不但指導教育計劃的方面，而且也同時貫有評定教育效果的骨幹。教育計劃或教育指導的實施，莫不是爲了實現教育目的，測量實施效果，故它一方面實施教育，一方面也在評價教育。爲了貫徹教育目的，對於重要環節之一的評價作用，不得不予以特別重視。

評價 (Evaluation) 一語，無論中文也好，英文也好，都是表示評判某件事或物的價值 (Value)。在所謂評價的當中，有如史密斯 (Sims V. N.) 所說的，「它含有能被視察的事件」與「觀察者的價值標準」二種條件。所以教育評價，也就是「將兒童、學生們的行動或學習，甚至教育計劃上所規劃所希望達到的實效，加以價值的衡量」。

(2) 教育測驗的意義：

一般所謂測量，不外對於所能觀察的事物，用一定的量表 (Scale) 測其數量的結果的一種工作。而量表須具備一定的測量單位和測定的基點 (通常用○點)。

教育測驗所使用的尺度，如偏差值，百分比等指數，則遠不如物理的測量所用的尺，或秤，來得精密，因爲教育測驗的尺度單位不比長度的公分 (Cm) 重量的公克 (g) 等，應用於任何地方都能量出相等的結果，所以教育測量總是不如物理測量那樣的準確。

還有教育測驗所用的尺度基點，也不比重量或長度的基點那樣有絕對性，教育測驗的尺度基點，有時可以任意加以調整做相對性的規

定，例如以偏差值來說，它的測定基點放在常態曲線下 $-5 S.D.$ 的地方，再者，教育測驗雖然也是科學的測驗但它的測量都是屬於間接測量（indirect measurement）的性質。以長度為單位的公尺或以重量為單位的公斤，能用於直接測量（direct measurement），但是智能、學力、人格等方面的測量，還是要從行動表現來進行間接測驗。所謂教育測驗，如斯而已。

（8）評價與測驗：

概要地說，測驗是用心理學或測驗學的技術，去測得客觀的數量資料。評價則不然，評價是為診斷教育工作究竟達成幾分目標，得到多少成就，以便作為改進的依據。測驗是評價法的一種技術，它可以測得一種數量的結果。將測驗的結果，所得到的價值，去和教育目標相對照，在對照之後，加以解釋分析。所以，教育測驗是專為測定客觀的數量作為參考的資料，教育評價則是對於測得的資料加以解釋，它的重點，應放在診斷意義之上。

是故教育評價並非排斥教育測驗。凡是能够利用測驗方法，都希望能儘可能地用客觀的測驗，去求得客觀的資料以為評價依據。不過在能够評價的事情之中，儘管如桑代克所說「凡存在的東西，都含有數量，凡含有數量的東西，都可測量」的信條如何冠冕堂皇，像態度、價值觀、鑑賞、習慣、行動、性格等項，還是難以客觀測量。而且像態度這類教育效果往往含有不可測量的東西在，站在教育評價的立場，這些則正是評價的重要對象。所謂現代教育評價法，就是對於能够產生客觀數量的測驗法，與對教育本質上的變化結果的主觀觀察手段，都須兼容並採，那一方面都不可排除，在適當的場所活用適當的方法，以進行評價。客觀的測驗，自然還是受到相當的尊重，而很公正的主觀觀察法，也同樣獲得尊重，換句話說，就是對量的資料和質的記載，都同樣予以重視。

再則，所謂教育評價，並不一定完全是照教育的價值基準為判斷。例如智能測驗的結果，絕不可能自教育價值來衡量，IQ的高低祇

是說明教育的可能性，並非教育的效果。或如對於學生行動的評價，亦不見得一定要評定學生行為的等級為滿足，行動評價的主要目的，還是在於瞭解兒童行為之動機，從心理學的理論基礎上，將其發生的因果關係加以推測，以便作為指導的依據。學業成績的評價也同樣地要時時注重分析學習過程，學習困難之所在，與其說為評價不如說為指導。評價的概念，並不完全根據數值來解釋。它是為測定一種教育結果，和其他各方面究竟有多少關係，然後透過教育全體的關節，來加以解釋，這便是教育評價。

要之，為了對兒童日常教育指導，以及對教育的改良起見，必須從更廣的範圍內，施行各種測驗診斷以求達到教育所期的一切效果，這便是現代教育評價的真面目。是以教育測驗被包括在教育評價的結構之內。

II 教育評價的功能：

無論教育測驗的時代也好，或者進步到教育評價的時代也好，我們都可從前述教育評價的演變歷史中，獲得一項重要事實，即教育評價是為使教育經營合理化和效率化，其終極目的，是在於促使教育效果之向上，是企望教育獲得實效的一種手段而已。

要提高教育效果，不但要建立正確的目標，更要本着目標作正確的指導。但是單有目標和指導，還是不夠，必須要對教育計劃，以及教育指導法從本身加以評判，對教師的教學以及學生的行動加以評判，俾有充分反省的機會。對於評判的結果，在不滿意的地方加以最大的改進，同時對於發生缺陷的原因及影響的條件加以摒除，從而修正改進教育。惟獨如此，教育效果才能向科學的正確目標邁進。如其不然，教育目標徒然空懸，教育計劃也徒具形式。因為教育效果沒有正確的測量，教育作用沒有循着正確的依據加以改進，則教育便無法走上軌道。教育必須事先擬具有良好的目標與計劃，進而作良好的指導。有無良好的計劃與良好的指導，對於效果方面必須加以評價才能瞭

解，評價之後當能發現不良的指導法或不良的計劃，從而加以修正。教育評價，教育目標，與教育實施恰好是一連環節，對教育負有一種重大的調整作用。

教育評價能發生下列主要功能：

- (1)使學習指導表裏一致，以提高學科教育的效果。
- (2)瞭解學童的性格，行動和生活環境，使生活指導能獲得更高的效果。
- (3)使學童能對自己的進步與缺陷有所確認，藉以激勵，並培養其自我改善之動機。
- (4)體察教育計劃，指導法、教材、教具等之有效性的程度，作為改良與求進步的參考。
- (5)對於學童的知能、性格、身體、學習等方面之個別差異，有更明確的瞭解，使能適合個別指導。
- (6)使學校經營計劃、組織設施和管理等得到改善與進步。
- (7)依據評價所得的結果，回答學校所在地的社區或父兄等的批判，以獲得該社區的協助。
- (8)使用標準測驗，以便和別的學校比較成績，使本校的學生程度或本班的學習成果，有比較反省的機會，並且藉以維持一定水準，使教師得到安心。
- (9)製作指導錄，通訊簿，或進行能力分組的依據。和作為升學或入學選拔的資料。

評價在本質上要能實現以上所述各項的效能，無論如何，在教師的評價過程，或學生的學習活動中，似尚須幾種「評價效力素」藏在當中才行。而評價效力素，要由教師與學生共同認定，或者一方對另一方的認定，詳述如下：

- (1)促使理解自己的效力。
- (2)引起並持續動機的效力。
- (3)努力取得重點主義的效力。

(4)強化既存學力的效力。

(5)付與安定感的效力。

(6)指導態度（學生方面為學習態度）及指導方向（學生方面為學習方向）控制的效力。

有效果的評價，不能不具有這些效力，所以教育評價應確實保有以上所舉的效力才好。

IV 教育評價的範圍：

教育評價的範圍，雖可分為許多項，但這不過是為便於說明而加以分項而已。在這些項目中，各個項目往往互相密切關聯，解釋評價結果時，如不把這些相互關聯的地方加以申說，便往往摸不到真象。

(1)知能測驗：

(2)性向測驗：

這兩方面都是注重素質的檢查，可將測得的學童個別差異程度的資料，提供作學習指導或升學指導的根本參考。

(3)人格的評價：

行為評價，或性格的評價，是評價兒童的行為特質，瞭解性格、態度、興趣、適應性等情形，以便給學生更好的指導。

(4)學習的評價：

學習評價，或學科成績評價，或更狹義說是學力測驗，是為評價各種學科的學習效果，指導計劃，及指導法等，直接或間接地都有促使指導或學習方面提高成績的作用。

(5)身體的評價：

這是為身體發育或促進健康而實施，也是為更合理的指導而着想。

以上五項都是產生於「各個學童之成長、發展的評價」的一大概念。

(6)環境的評價：

能够影響兒童成長，發展的家庭狀況，社區狀況等方面之評價，亦為教育評價的一種重要範圍。

(7)學校評價：

學校是對學生教育上最重要的環境，必須單獨予以評價。其範圍是：學校行事的種種計劃，教職員，設備，管理，圖書館等的一一評定，以作為改善學校教育的資料。

(8)教師的評價：

教師在學校環境中，以及對學生的指導上，是最有關係，又是最為重要的對象。對教師的評價是要考量以教師對學生指導，如何地從良心上善盡自己的責任，有無經常作最大努力的現象，這是教師評價的前提。

此外當然還有其他範圍可以考慮，但一般地說，上述各項應當是屬於最重要的部份。對於各項範圍的評價法，俟第六章以後，再為詳述。

※要做更進一步研究的指引與參考書，介紹如下：

(1)要更進一步理解教育評價與教育測驗的意義與性格，可參考下列論文或參考書。

小見山榮一 心理學と測定（東京教育大學心理學教室編，教育心理學第二集）。

現代教育心理學大系 測定、評價（理論篇）中山書店 昭和33年V測定の基礎）。

田中寬一著 教育的測定學，松島三松堂，大正15年第二編3章。

大伴茂著 教育測定の原理と方法，培風館，昭和6年第1編。

(2)要詳細理解教育測驗的發展歷史，可參考下列各書：

橋本重治著 教育評價法總說，金子書房 昭和34年第1章。

四方實一著 教育評價法 日本文化科學社 昭和26年，第3章。

Ross, cc, Measurement in Today's School, 1947, Chap.2.

(3)由教育測驗到教育評價的轉移之經緯方面有：

長島貞夫 新教育と評價 東京文理科學大學兒童研究會編 學習指導，

金子書房，昭和22年。

(4)由教育測驗到教育評價，在教育史上乃是受到用科學方法研究教育問題之後，發現有種種問題必須要解決，而發展演變起來的。這一點有特別加以注意的必要，要調查這一方面的歷史事實，最好參考下列各書：

橋本重治著 教育評價法總說 第1章。

橋本重治著 學力檢查法 金子書房 昭和27年 第3章第I節
Ross, cc., Measurement in Today's School, Chap. 1

(5)美國的八年研究，是教育評價研究的起源，這是一致認為最值得注目的事，關心這一方面的人，可參看：

Smith, E. R. and Tyler, R. W. Appraising and Recording Student Progress, 1942.

(6)關於教育評價的意義與作用，最好更進一步去參看下列各書：

橋本重治著 教育評價法總說，金子書房，昭和34年第1章

教育心理學講座 教育評價と測定，金子書房 昭和28年「教育評價の機能と教育の問題」。

教育大學講座 教育評價，金子書房 昭和25年「評價の教育的意義」。

Ross, cc., and Stley J. C Measurement in Today's School's 1754 Chap 11. 12. 13 and 14.

(7)要更進一步地明瞭有關評價效果的基礎理論，可參看：

橋本重治著 教育評價法總說，金子書房 昭和34年，第3章

(8)關於教育評價的領域或對象的確定方面，可參看：

Remmers, H. H. and Gage, N. L. Educational Measurement and Evaluation, 1943. Part 1.