

教育測驗與統計

目 次

第一編 測驗概論

第一章 測驗的意義·····	1
一．主觀測量和客觀測量的區別·····	1
二．測驗的優點·····	2
三．測驗的內容·····	3
第二章 測驗的功用和限制·····	4
一．測驗的功用·····	4
二．測驗功用的限制·····	9
第三章 測驗的歷史和種類·····	12
一．測驗的歷史·····	12
二．測驗的種類·····	18
第四章 智力測驗·····	21
一．智力測驗的意義·····	21
二．團體智力測驗·····	22
1．廖氏團體智力測驗·····	22

2. 陳氏圖形智力測驗	32
三. 個別智力測驗	37
3. 陸氏訂正比納西蒙智力測驗	37
第五章 教育測驗	44
一. 教育測驗的意義	44
二. 國語測驗	46
4. 陳氏小學默讀測驗	46
5. 陳氏小學默字測驗	50
6. 廖氏中學文學常識測驗	54
7. 陳氏小學文法測驗	54
8. 俞氏小學綴法測驗	55
9. 俞氏小學書法測驗	60
第六章 教育測驗(續)	66
三. 算術測驗	66
10. 德氏算術四則測驗	67
11. 俞氏小學算術應用題測驗	69
12. 俞朱柯麥四氏算術練習測驗	72
四. 社會自然及常識測驗	74
13. 俞氏小學社會自然測驗	74
14. 陳氏小學常識測驗	76
五. 圖畫和音樂測驗	78
15. 克朗—卡雷圖畫測驗	78

目	次	3
16. 郭華司—盧西音樂作業測驗·····	82	
附一 英文測驗·····	86	
附二 各科混合測驗·····	86	
第七章 其他測驗·····	87	
一. 體力測驗·····	87	
17. 視力測驗·····	88	
18. 聽力測驗·····	89	
二. 品行測驗·····	90	
19. 廖氏道德判斷測驗·····	91	
三. 職業測驗·····	93	
20. 德氏機械的智力測驗·····	94	
四. 測驗式題目·····	96	

第 二 編 測驗實施法

第一章 核定測驗的標準·····	103
第二章 測驗施行的方法·····	105
第三章 示範與實習·····	107
一. 示範測驗·····	107
二. 實習·····	107
三. 分組實習與教學時實習·····	110
第四章 測驗記分規定法·····	111
一. 無巧合機會者不扣分·····	112
二. 有巧合機會者扣分·····	112

第五章 實足年齡核算法	114
一. 計算法	115
二. 查表法	115
三. 概算法	117
第六章 智力年齡和教育年齡核算法	118
一. 智力年齡核算法	118
二. 教育年齡核算法	121
三. 全班年齡核算法	122
第七章 智力商數、教育商數、成績商數	123
一. 智力商數	123
二. 教育商數	125
三. 成績商數	125
第 三 編 測驗上應用之圖表法及統計法	
第一章 圖表法	129
一. 圖示法	129
二. 表列法	137
三. 圖示法與表列法的差別	140
第二章 全體數的統計	141
一. 分配圖	141
二. 分配表	144
三. 順序表	147
第三章 均數的統計	148

目	次	5
一. 平均數	148	
二. 中數	152	
三. 衆數	155	
第四章 離中差的統計	157	
一. 離中差的意義	157	
二. 四分差	158	
三. 平均差	161	
四. 標準差	165	
五. 相對差數	170	
六. 偏態度之測量	171	
第五章 相關係數	173	
一. 相關係數的意義	173	
二. 算法	175	
三. 相關係數的解釋	185	
四. 相關係數的可靠性	185	

第 四 編 TBCF計算法

第一章 怎樣求TBCF	189
一. 陳氏小學默字測驗的TBCF	189
二. 俞氏小學算術應用題測驗的TBCF	196
三. 陸氏訂正比納西蒙智力測驗的TBCF	200
第二章 TBCF的解釋	204
一. T量表的製法與理論	205

二. B量表的製法與理論·····	211
三. C量表的製法·····	216
四. F的理論·····	217
五. 各種公式的檢討·····	218
第三章 TBCF的應用·····	221
一. 化非標準測驗或考試的分數爲TBCF·····	221
二. 化教師估定升級序數爲T分數·····	224
三. 計算全班 TBCF的方法·····	226
四. 學校行政上之各種計算法·····	227
第四章 餘論·····	232
一. 年齡量表與T量表·····	232
二. TBCF命名的意義·····	234

附 錄

本書外國人名中西對照表·····	236
------------------	-----

教育測驗與統計

第一編 測驗概論

第一章 測驗的意義

一、主觀測量和客觀測量的區別

1. 測驗是什麼 什麼叫做測驗？簡單說來，就是根據心理學和教育學的原理，去測量人類精神特質的一種新方法：比如布用尺量，米用斗量，物用秤量，難道人的智慧和學力，就不能量麼？學校或各種機關的成績，就不能量麼？若要量我們該用什麼標準去量？該用什麼方法去量？量完以後，又該用什麼法子去計算牠的結果？這些事都是測驗範圍以內要研究的問題。

2. 舊社會裏的測量 從前的社會裏邊，對於人類的精神特質，亦何嘗不測量呢？比如某商店用一個夥友的時候，如認為他可以用，就說某人「好」「很好」「最好」，學校裏考試的時候，教員就把考卷上面記上「七十分」「八十分」的分數，這不是

一種測量麼？但我們反過來要問他那「好」「很好」「最好」的意義，是以什麼為標準？所謂「好」與「很好」相差多少？「很好」與「最好」又差幾分之幾？某人對於辦銀行的事情好？還是貨行的事情好？是辦會計「好」或「很好」呢？還是辦交際「很好」或「最好」呢？再問考卷上的分數為什麼不記上七〇・〇二五六……，八〇・一三四九分數，普通只記整數的分數呢？這些問題，恐怕沒有一個人，能給我們真確的解釋。

3. 主觀的測量和客觀的測量 進一步的說來，從前社會裏邊的測量都是主觀的：我認為他好，他就是好，我認為他壞，他就是壞。測驗則不然，牠是想把這主觀的測量變為客觀的測量，換句話說，就是無論千百萬人去測量某人或某團體，要教他測量的結果是一樣的，這就是測驗能夠長足進步的原因，亦就是牠能在教育科目中，佔一地位的緣故。

二、測驗的優點

測驗既是用科學方法，求客觀的標準，以矯正主觀方法的弊病，已如上述。現在把這種標準的好處分述如下：

1. 客觀標準是普遍的 我們如談起某天才生的成績，便說他很好，談到紐約八十五層的帝國大樓，便說他很高；可是「很好」或「很高」並表示不出他們的程度，應當說「幾多好」或「幾多高」纔能有個標準，去與其他一切比較。不過房屋高低，是可以用尺量的，故能確定「幾多」。學校裏沒有那支「尺」來度量成績，所以通常說不出「幾多」。用了測驗，各校便可以找到

單位相同的量表 Scale，來度量或比較各種課業的成績，與兒童的智慧。

2. 客觀標準是精密的 主觀的測量，終有許多疏漏的地方。科學的方法，比較來得精密。比方編製一個測驗時，須先搜集了各種教本，與擔任該學科的教師談話，再徵集專家的意見，然後纔決定教材中的要點，和測驗題目的取舍。及取定題目後，又要依所編測驗的性質，來決定排列的次序。測驗既成，其施行的手續，都有詳細的說明；測驗終了，其記分與核算的方法，也都有詳細的規定。所以有許多觀察不到的地方，用了測驗，便可以推原竟委了。

3. 客觀標準是公開的 照上面說來，測驗不但是量表劃一，材料劃一，就是施行的手續，與計算的方法，也都是劃一的。所以測驗的結果，不能因時因人因地而異：比方今天我測驗得這樣結果，下次依舊這樣，不會有什麼上落。不單是我個人如此，別人用了同樣的方法測驗時，也可得到同樣結果的。

三、測驗的內容

照第一節所述，測驗是測量人類的精神特質。但是精神特質的範圍很廣，所以測驗的種類亦很多。有測量智力才能的，叫做智力測驗；有測量學業成績的，叫做教育測驗；此外如測量身體的狀況，則有體力測驗；測量個人的品格，則有品性測驗；應用於職業方面的，則有職業測驗。其中智力測驗與教育測驗二者，在中國已粗具規模；至於其他各種，尙正在研究試驗時期。現在

測驗的內容，大概說來，不過如此。但就美國言，如品性測驗職業測驗等，功效亦已顯著，所以其前途是未可限量的。

研究問題

1. 測量和考試根本上有什麼異點？
2. 客觀標準的好處，是些什麼？
3. 測驗中那兩種測驗已粗具規模？

參考書報

1. 廖世承陳鶴琴合編：測驗概要第一章及第四章。（商務）
2. 華超：教育測驗綱要第二章。（商務）
3. 周調陽：教育測量法精義第二章。（中華）
4. 龔啓昌：編造中小學教育測驗之理論與方法—中華教育界二十卷八期一（中華）
5. Mc Call, W. A. : How to measure in Education, Chapter I. (The Macmillan Co.)

第二章 測驗的功用和限制

一、測驗的功用

測驗的用處至廣，茲撮要分述如下：

I、智力測驗與教育測驗聯合的用處

1. 分別班次 同班的學生，必須智愚齊一程度相仿，纔能謀教學上及訓育上之順利進行。智力測驗能查出個人先天所具有

的才能；教育測驗能查出個人後天獲得的教育；故利用兩種測驗之結果，可做分班之標準。

2. 設立特別班 學校中之高才生及低能兒，均不應使他和普通學生，同在一班讀書，所以優良學校裏邊，常設有特別班。我們亦應用智力測驗與教育測驗查出低能兒及天才生，組織低能兒班及天才生班以利教育。

3. 設立試驗班 欲知教育上某種組織的好壞，必須根據牠產生的結果，纔敢斷定：比如要決定設計教學法的價值，應先用智力測驗和教育測驗挑選一班優等生，組織設計法試驗班，施行設計教學，經過相當時日，再用測驗查其結果如何，然後方能判定價值的高低；又如要比較設計教學法和道爾頓制教學法的價值，亦必須經過測驗，組成兩班的學生，去作試驗，過了若干時期，再各用測驗測量數次。看兩班成績的高下，以定這兩種教學法價值的大小。

4. 分別才能 兒童的智力，各不相同，所以他們的興趣和需要，亦各有差異，我們應當因材施教，使各得發展。不過用什麼方法去分別各種才能呢？自然是用智力測驗與教育測驗，比用主觀的考試和教師的估量正確的多。所以就從這一點看起來，測驗是很有價值的。

5. 入學試驗 普通入學試驗，只憑一二篇文章，四五道算題，分別去取，試驗的工具既不可靠，錄取的標準焉能準確？其結果亦許使聰穎者落第，僥倖者濫竽，實在是危險萬分。所以現

在美國由小學以至大學，入學試驗均採用測驗方法，即一面用智力測驗，以測量學生的才能，一面用教育測驗，去考查他學業的程度。二者參用，以定取舍。

6.規定升級降級 一班中之學生，不但須各個人之學力相若，纔可受同樣之教育；且須智力相若，纔能望有同一之進步：故學校中用智力測驗和教育測驗以定升級及降級之標準，較為精確公允。

7.教員優劣的甄別 教員教學的優劣，固可由學生成績之良否而定；但是學生天資之良否，亦應顧及：因為天資魯鈍的學生，其學業絕難和聰明的學生相較。故欲甄別教員教學之效率，對於其所教之學生，必須用智力測驗與教育測驗的結果，求出其成績商數或努力數，然後再去相比，纔可以免去智愚不同的差誤（詳第二編第七章及第四編第三章）。此外尚有應注意者，即斷定教員之優劣時，除測驗學生之智力及教育成績外，尤須注意教員之品格，人生觀，及生活之興味等等。因學生所受之教育，不僅在教員之正式教學，此外教員之人格感化，亦為良好教育的重要元素。

8.供給精確的記錄和報告 學校對於學生的成績，應有適當的記錄並報告於學生的家長。從前的辦法，或用考試的分數，或用主觀的評語，標準既不一致，記錄與報告，便失其效用。若用教育測驗，則其成績自不能因時因人因地而異，既能確示其學業詣造的程度，又可表示其學業的進度；若再和智力測驗的成績

相較，更可知道努力的度數，以此而作記錄和報告的根據，自然十分可靠。

9. 爲行政人員調查學校效率的利器 一校的校長，對於校內的實在情形，欲求真確的明瞭，往往苦於無法考察；督學或其他教育行政人員調查學校，更是走馬觀花，下幾句主觀的考語，這又有何用處？如要真正調查學校的效率，莫若用智力測驗與教育測驗：因爲這樣調查的結果，都是根據事實而來，當然正確可靠。近來歐美各國視學多用此法，最近上海市教育局也曾測驗過全市的小學，都是這個道理。

II、教育測驗的用途

1. 改進教學方法 (1) 經過各種測驗之後，教員能知各學生能力差異之所在，而注意在個人指導，以適應其個性。(2) 施行教育測驗之後，能查出一部分學生對於某種學科之成績是否優良。其成績優良者，則隨學科之性質，而增加其學習的材料或核減其學習之時間；其成績不良者，則又必須增加其學習的時間或用特別訓練，以資補救。

2. 鼓勵學生作業 (1) 教育心理學上有一原則，就是學生須認識他本身作業時的進步，纔可增加他學習的動機。學校於規定時間內，施行測驗，使各人保有他的成績，並作進步曲線圖，使他們自己知道自己進步的情形，藉作比較，定能促進他們上進的興趣，並可鼓勵他們的競爭心。(2) 從學習的效率言，練習愈多，應用愈廣，則進步愈速。故欲令學生之進步加速，非多予以

測驗，以增加其練習之機會不可；所以教育測驗是刺激進步之原動力。

3. 考查成績 用記分法考查成績，我們早已知其不正確，倘用教育測驗去評判兒童作業的成績，雖不能說完全正確可靠，但比較不統一的記分法，終勝數籌，故教育測驗，可代替舊法考試而一矯其弊。

4. 釐訂各科各學年的學力標準 倘若能充分的利用教育測驗，便可以得到各科各學年的學力標準：就是說在某學年裏的兒童，對於某學科之成績，應該有怎樣的標準，如果不及這個標準，教員對於他們的成績，就要設法補救的。

5. 測知學校的效率 評判學校效率的優劣，我們都知道不該用個人主觀的評論法，該用客觀的標準測驗；因為用測驗來評判學校效率，都是有事實足以證明。（其法以教育測驗成績與學校經費去比例，計算法詳第四篇第三章）不是偏袒任何一方面的。

III、智力測驗的用處

智力測驗，除用之於教育事業外，尚有其他用處如下：

1. 智力測驗與職業 向來工商各業收用夥友生徒，毫無科學的標準，全憑荐保之言，以定取舍，如不能勝任，則交替更換。這種方法，不獨對於店舖不經濟，就是對於夥友，也是虛度光陰，枉費心血。中國工商各業，不發達的原因固多，人不擇業，業不擇人，也是一個重要的原因。若要救濟此種缺點，應當（1）

先用智力測驗，觀其智力的高下，以定其擇業的大概範圍。(2) 再用職業測驗，決定其該習某種職業。

2. 智力測驗與軍事 智力測驗對於軍事，亦有卓著的成績：美國於1917年與德奧戰爭時，就利用智力測驗，來編制軍隊，此次試驗過共計一百七十餘萬人，他的結果，(1) 甄別智愚：使智的特別升調，愚的盡行淘汰。(2) 支配才能，使各為其所，無濫竽的弊病。當年美軍所以那樣精銳，不能說與智力測驗無關。吾人所謂軍隊測驗，亦即指此而言。

3. 智力測驗與社會問題 犯罪行為乃社會中的重大問題，從前大家都推想犯罪的人，多半是心理缺陷的人。但是心理怎樣缺陷？都不了然。自智力測驗產生後，纔證明智力缺乏，實為犯罪的一重要原因。

4. 智力測驗與各種學術 (1) 世界上各種人類的智力，是否相等？如白人與黃人究竟智力孰優，這個問題，研究種族心理學和人種學的學者，至今尚未解決。(2) 又如謂「健全的精神，常寓於健全的身體之中。」身體健全乃生理學與醫學上的問題，精神健全，則為心理學與智力的問題。究竟身體健全與精神健全有若大的關係，現在亦無定論。我們觀上述二項，可以想到智力測驗與各種學術的關係，而有待後起者之發揚光大，實無疑義。

二、測驗功用的限制

測驗的用處，雖然廣大無窮，但亦有牠的限制，並不能視為萬能：

1. 智力測驗的效力爲後天的經驗所限制 智力爲精神方面的遺傳，是先天的，所以智力測驗，正是要測量某人先天的才能，以斷定他將來發展的限度。換言之，即某人幼時智力低，預料其成長後，亦必愚笨；某人幼時智力高，預料其成長後，亦必聰明。但反對這種說法的，以爲人類智力的生長，從縱的方面來說，固然是到成熟時期爲止，但從橫的方面去講，先天的智力與後天獲得的經驗，是同時並進，不可劃分的。所以智力與經驗既不能截然劃分，更不能藉智力測驗來斷定人的一生！關於這一點的解釋，我們認爲智力與經驗，雖不能劃分，但各個兒童在同一教育環境之下，有的進步較速，有的進步緩慢，必與智力之高低有關。智力測驗所用的材料，雖大半根據後天的經驗，但後天的進步，既因先天的關係而不同，所以根據後天的經驗來做材料，亦可以測量先天的智力。因此智力測驗，雖不敢說能斷定人的一生，但亦不能謂其無用。

2. 智力測驗不能精確的選擇特殊才能 智力測驗不能斷定一個兒童應爲雕刻家，或爲建築家，或爲文學家，他所能表示的，祇是把智力的高低，分成等級，然後斷定某個兒童屬於某級。倘若問某個兒童，具有某級的普通智力，足以適應某一類職業上的需要，這是智力測驗所能解答的。若要問對於某種職業上的興趣、特殊才能等等，則非智力測驗所能知道，而有待於完成後的職業測驗。

3. 測驗不能完全評定兒童的成業 一個兒童作業的成功，

不能完全憑恃他的智力，此外如忠誠、勇敢、堅毅、勤勞、健康等良好的習慣，均極重要。這些良好習慣，無論智力測驗或教育測驗，均不能測驗出來。即就現有的教育測驗而論，亦不過測驗兒童對於某種技術能力的大小，如教師對於兒童的人格感化等等，就無從測驗到。其實此種影響對於兒童將來事業的成功，或比學業成績，更屬重要。凡此等難題，待品行測驗與體力測驗完成後，或有解決的希望。

研究問題

1. 測驗對於分班的功用若何？
2. 用測驗甄別教師優劣，其方法為何？為什麼？
3. 督學視察學校時，用測驗有何好處？
4. 測驗與新教學法關係如何？
5. 我們可否憑了智力測驗，斷定某人宜於學習某一種職業？
6. 為什麼智力低的人，容易犯罪？
7. 測驗的功用，有些什麼限制？
8. 近人對於智力測驗的效力，有何種辯論？

參考書報

1. 廖世承陳鶴琴合編：測驗概要第二章。（商務）
2. 周調陽：教育測量法精義第三章。（中華）
3. 李鼎輔：教育測驗與統計法第三章。（北平師大第二附小）