

527.116
508

新考 論試

著煊集史

行發局書智民海上

史美煊 著

考
試
新
論

上海民智書局發行

中華民國二十二年六月初版

考試新論

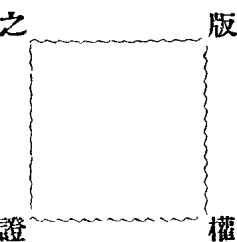
（每冊定價大洋五角五分）

（外埠酌加郵費函費）

著者 史美煊

印刷者 民智印刷所
上海塘山路九百二十六號

發行者 民智書局
上海河南路二〇〇至二〇二號



分發行處 民智書局
南京 廣州 北平 武昌 長沙

分售處 海內外各大書坊

總發行所 民智書局
上海河南路中市

二〇〇至二〇二號

憲	日	之	深	服	總	貫	余
法	所	新	慚	務	理	之	常
則	希	著	愧	於	所	係	謂
功	望	式	耳	政	理	統	學
業	之	足	史	府	想	然	校
更	建	補	美	主	之	後	與
無	設	益	煙	持	考		任
量	方	於	君	考	試		用
矣	針	教	著	政	方		宜
	有	育	考	者	克		吏
民	所	倘	試	四	實		兩
傳	貢	更	新	年	現		種
聞	獻	進	論	尚	惜		考
題	以	一	為	無	願		試
○	弼	步	中	所	奢		必
年	成	於	國	發	而		須
春	玉	聞	空	明	力		有
月	權	平	前	珠	小		一

自序

學校考試，向來是認爲例行的事務，無人注意。直至近二十餘年，教育的研究日趨於科學化，考試方法遂大受批評。若干調查均證明考試分數之不可靠。於是標準測驗代之而興。此種新發展，一部分是兒童研究的興趣之結果，一部分則由於教育統計之最近的發達，乃屬考試方法演進之第一步。這種測驗運動，在近代教育史上佔一重要的地位，是大家公認的事實。歐美各國姑且不論，即在我國，已經出版的標準測驗不下數十種，專論測驗的著作亦頗多，由此可見其盛況。

這種新的標準測驗之優點甚多，凡曾稍加研究或應用它的人都曉得，但亦有其弱點，如不能常用、結果不易解釋、等等。所以最近的趨勢注意非標準測驗，即所謂新法考試。此種測驗之產生，可說是考試方法演進之第二步。它具有標準測驗之特長，而無其缺欠，故頗風行一時。

本書之目的，就在詳述此種新法考試之意義、價值、編造、施行、校閱、與應用，旁及優良的考試之標準、舊式考試之改良法、分數制度、等等。至於教育測驗之高深的理論與複雜統計，則一概從略。所以本書不但便於師範學校的學生與在職的教員閱讀之用，即工商界選擇職員與行政當局考試公務員時，如能加以參考，不無裨益。

著者在二年以前，即已開始收集材料，着手編輯，但因種種關係，屢作屢輟，到現在方才草畢，然不滿意之處甚多，終因友人之催促，乃交書局付印，尙望讀者不吝賜教，幸甚。

本書之完成，得力於附錄中各參考書處甚多，未能一一註釋，謹此聲明，並示感謝。書中有數章曾於中國測驗學會出版之測驗上發表。一切譯名，多沿用舊有者，間有少數新譯者，然均列入附錄譯名對照表中。李君家湛代抄全稿，殊爲感激，特此誌謝。

考試新論目錄

戴序

自序

第一章 緒論

考試之批評及其答辯——考試之目的——考試之種類

一

第二章

優良的考試之要素

特效——可靠性——客觀性——廣博性——便易性——用途——融洽性

一七

第三章

考試分數之研究

三一

學校的分數之研究——教員的分數之研究——校閱同一試卷的結果之研究——同一教員前後所定的分數之比較——問題的價值之研究

第四章 舊式考試之改進……………五一

舊式考試之優點——舊式考試之限制——舊式考試之改良法

第五章 新法考試……………六五

新法考試之優點——新法考試之限制——新法考試之種類及其舉例

第六章 新法考試之編造……………八一

新法考試之編造——新法考試之施行——新法考試之校閱——結果之處理

第七章 追憶測驗……………一〇五

一字答測驗——填字測驗

第八章 真偽測驗……………一一七

真偽測驗之意義與性質——真偽測驗之優點與限制——真偽測驗之

編造——真偽測驗之施行與校閱——真偽測驗之變像——精度與反面暗示的影響

第九章 選答測驗……………一三二

選答測驗之意義與性質——選答測驗之優點與限制——選答測驗之編造——選答測驗之施行與校閱——選答測驗之變像

第十章 雜式測驗……………一四三

對偶測驗——類比測驗——重組測驗——指認測驗——訂誤測驗——

第十一章 綜合測驗……………一五七

綜合測驗之優點與限制——綜合測驗之編造——綜合測驗之施行與

校閱

第十二章 分數制度……………一六五

測驗分數與等級分數之區別——分數制度——化測驗分數爲等級分數法

附錄 譯名對照表

.....

一七九

參考書目

.....

一八九

考試新論

史美煊著

第一章 緒論

考試之批評及其答辯

考試並不是一種近代風行或新發明的方法，它的淵源甚古。在多數的學校中，成績之測量，已有其相當的地位。大概世界最早的考試制度，當推我國古代的官吏考試制。三代以上（約在西歷紀元前二二〇〇年）我國就用此法取士。他如希臘與羅馬，亦曾實行此法。但是現代歐美各國的學校考試，則發源於中世紀的大學考試。

考試雖有這樣悠久的歷史，而其價值却未確定。對於考試之批評，由來甚久。一七六六年牛津大學有一學生以爲其學位考試是一絕對的趣劇，包括一很容易的西伯來文問題與一歷史題。美國哈拉斯梅恩（Horace Mann）在一八四五年亦表示對於當時的考試方法不滿。然而討論之烈要算近世。近數十年

來，研究考試的文字，佔歐美教育著作之一重要部分。標準測驗之產生，即由於此。

我國自五四運動以後，有一部分教育家與在校的學生頗反對考試制度，甚至少數學校已經實行廢除，但是教育當局則力主保存之，國民政府更確立考試制度，以拔取一般官吏。故考試之存廢問題，實有研究之必要。

造成此問題之大原因，自然是由於若干調查考試與分數的價值的結果之公布。這些調查皆證明教員所給的分數之不可靠與差異，將於本書第三章詳述之。

我們歸納反對考試之理由，不外下列數點：

(一) 考試有礙於被試的健康，因為考試促成過度努力、氣餒、焦急、及其他不應有的身心影響。考試疲勞的結果，往往發生神經衰弱、神經興奮、恐怖不安等症。據茂慈(Melzi)的調查，在二百名女子小學生中，其卒業前一月，與卒業後體重之比較，平均減少七百五十克。

(二)考試的問題之內容，不適合公認的教育目標，却鼓勵強記。完全記憶事實，而無推理或其他高深的思想歷程。此種強記不但不能持久，反傷身心。

(三)考試常使本身成爲目標，學生以爲求學之最大目的是考試及格，而非精通各種學科。

(四)考試鼓勵作弊。考試時，學生每有夾帶、抄襲、竊視、或他種欺騙的舉動。

(五)考試造成應用不良的國語與惡劣的書法之壞習慣，因爲要想數量衆多，書寫迅速，遂致忽略良好的形式。

(六)考試之時間可作他種用途，如上課、溫習、研究、等等。

(七)教學的結果是無形的，無法測量。

(八)考試非必須的，優良的教員可以不用考試評判其學生的工作。

我們細閱上述八點，就知道它們都是反對考試之誤用，而非反對考試本

身。換言之，我們可改善考試之形式與方法，而不必廢除之。所以我們對於諸反對理由之總答辯，是編造、施行、校閱、及應用得法的考試，對於這些弱點均可避免。茲再分別答辯如後：

考試的結果妨礙被試的身心福利，誠屬不可諱言的事實。但是修學以及他種學校工作與校外的一切工作也是一樣，却沒有人主張因此事而完全廢除各種學校工作或校外工作。若在衛生的情況下——如溫度得宜、空氣流通、光線充足、座位適當之教室中——舉行考試，而考試的時間又不太長，實可避免此種弊端。

考試之所以不常符合公認的教育目標，乃由於編造者之無知與粗率。草擬測量事實的知識之問題，自比編撰測量推理能力、適應能力、應用能力，及心理發展者為易。然而測量後數種能力之考試，亦屬可能的事。故救濟的方法不是廢除考試，却是明定目標與給予教員以編造符合目標的考試之訓練。若說考試鼓勵強記，至少可有二種解釋。第一，強記視考試的情形而定。

。如果單純事實知識之測量達其正常地位，即僅爲考試之一很小的機能，也可減少強記，使之不佔重要地位。第二，所謂強記之某部分並不是完全不需要。學習之一原則，是常須溫習已習得者。倘使考試鼓勵溫習，則已達到一合理的目的。其實，強記總比不記爲佳。再自另一見地觀之，強記是在較短的時間內閱讀巨量的材料而保存之，乃屬一有價值的能力。律師預備案情，醫生診治病症，以及他種職業中，均很需要此種能力。

的確，有時考試本身成爲一目標。然而考試是否成爲惟一的目標，全恃教員重視考試的程度而定。如使學生認考試之成敗爲莫大的榮辱，這種態度與結果自不佳妙。在他方面，如果學生視考試爲課業之一部分，僅佔學期或學年分數之一小部分，考試非但不礙目標，且爲一合理的。尤有進者，若考試能精秘的包括已教的材料，則考試的目標，很類似精通課程的內容之目的。這些目的自然不能認爲不必需。

學生在考試時，每有作弊或欺騙情事，但是他們在其他日常工作，遊戲

、交際、及各種活動中，也如此作，却無人以為應該阻止學生工作、遊戲、交際、等等。如果考試編造得法，使被試難於作弊，並且對於作弊者予以嚴厲的處罰，大可避免此種弱點。苟欲達到此步，必須教員方面之批評的與審慎的判斷。學生是否欺騙，全恃三種原素而定：即其所受的德育、欺騙之需要、及欺騙而不被查出之機會。倘若校風不利於作弊，使被試感覺分數之增加、不敵自尊之損失；教員不太着重考試，使無欺騙之必要，而考試之編造，又使不易欺騙，自然無庸顧慮此點。

考試常注重於書寫之分量與速率，而忽略良好的形式。故考試不僅一任惡劣的書法、標點、文法、等等不加糾正，甚且發展之。其實，施行得法的考試可免此弊。後面將提出應用文字的良好習慣與發展表述的能力乃屬考試之一機能。簡要的救濟辦法，是予以充分的時間，訂出高的標準，及側重於合理的目標。

關於利用考試的時間作他種活動一點，也與對於他種活動同一見解。換

言之，有些教員用太多的時間測驗學生的工作，最好將其一部分用在他種活動上。在他方面，有些教員對於此事所支配的時間太少。考試的時間之比例，依科目或所教的科目之某方面、學生以前的預備及其成熟程度、教學的方法、與其他原素而異，故無定例。在較小的班級中，教員易用他法與各生接觸，故考試的時間可少。依照各種理論與實驗的證據，都證明時常測驗的學生比不常測驗的學生更達較高的成績標準。

『教育的產品是無形的，且是無法測量的』一語，我們承認其有相當的真理。若把它們與工業及科學的出品比較，實是無形的。然而它們的區別是屬於程度的而非性質的。在某些工作方面——如教育之類——的確不能測量其結果達一極精密且完全的程度。以前對於正確的科學也是同樣的情形，就是現在，它們的測量之準度也非絕對的，却僅比較高。此事自不能使我們主張廢除教育測量，却應引起研究與改良測量工具之決心。雖然心理的發展與成績不能達到高度，重量、面積、溫度、等等測量之客觀、可靠、與特效，然