

心理與教育測驗

第一章 緒論

第一節 測驗的意義

(一)測驗的性質 教育上有三個重要問題：第一，教育的目標是什麼？這是教育哲學和教育社會學的問題；第二，用什麼方法去達到教育目標？這是教育心理學，教育行政學和教學方法的問題；第三，教育目標已否達到？這就是心理與教育測驗的問題了。所以測驗是教育上的一種工具，用以度量教育的結果，以為改進的根據的。

度量物質的工具，已經有很多種了。例如權衡輕重有秤，度量長短有尺，計算時間有鐘錶。甚至於溫度的高低，電流的多寡，都有測量的工具。自然科學之所以能够日益進步，未始不由於此。我們要想改進教育方法，提高教育效率，那麼，對於度量教育結果的工具和方法，自不能不加以研究。

(二)測驗的可能 不過屬於物質方面的東西，顯而易見，度量時並不困難；而教育的結果，隱而難明，度量時卻很不容易。因此，有人對於心理測驗的可能性，發生懷疑。他們以為智力、品格、性情、興趣、意志等抽象的對象，是否可以用科學的方法去測量？書法的優劣，文章的豪邁，學力的高低等，是否可以用數字去計算？這些都是可懷疑的問題。其實，心理測量，也和物質測量一樣是可能的。我們祇要明瞭測驗的兩個

理論的基礎，我們的懷疑就可以渙然冰釋了。現在把心理測驗在理論上的兩個基本原則畧述如下，以闡明心理測驗的可能性：

(1) 任何現象，祇要是存在的，總有數量。這個原則是美國心理學家桑戴克(E. L. Thorndike)於1918年提出的。他說：「凡物苟能存在，必有其數量」。(Whatever exists at all, exists in some amount.) 例如智慧有高低，品質有優劣，文章有工拙，學力有高下，在高低優劣之間，必有程度的不同。所謂程度的不同，就是數量上有差異。

(2) 凡有數量的現象，都可以測量。這個原則是美國測驗專家麥柯爾(W. A. McCall)於1922年提出的。他說：「凡有數量的東西，都可以測量」。(Anything that exists in amount, can be measured.) 這一條原則，在理論上是毫無問題的，但在實際上，目前尚難完全做到。教育的結果——知識、技能、習慣、品格、理想、態度、欣賞、興趣等——，我們現在還不能一一加以測量，這是不必否認的。但這是由於測驗的歷史很短，測驗的工具尚未完善的緣故，我們決不能因為測量某種教育結果的工具還未發明，而說某種教育結果是無法測量的。

(三) 測驗的方法 就上面兩條原則看來，便可以知道心裏測量是和物質測量一樣可能的。不過心裏測量的方法，和物質測量的方法不同。物質方面的東西，是固定的，具體的，可以直接去測量，例如用尺量布，用斗量米，布和米都是固定的具體的東西，所以能够直接的度量；而教育的結果——知識、技能、品格……等——是抽象的，時時變動的，所以不容易直接去測量，祇可以間接去測量。例如「智力」是一種「無聲無臭」的東西，不能直接去測量。祇可以從一個人的作業上去估定。所以智力測驗所測量的，雖名為先天的能力，實係後天的作業。由後天作業的成績去推測一人的先天能力，便是間接測量的方法。

總之，測驗是教育上的一種新工具，可以用來度量教育的結果，以

謀教育方法的改進。但測驗的歷史很短，測驗的方法，尚未達到完善的程度，然而這祇是工具完善不完善的問題，並非測驗的可能性有什麼可懷疑之處。測驗在教育上可能的貢獻很大，我們還要繼續努力研究，以求達到盡善盡美的地步。

第二節 測驗應具的條件

據測驗專家的意見，一個完善的測驗，須具備下列幾個條件。

(一)測驗的正確性須高 一個測驗，須能確實的測量到它所要測量的能力。例如一個智力測驗，應該純粹測量智力；一個算術測驗應該純粹測量算術能力。倘若一個測驗，真能確實的測量到它所欲測量的能力，可謂之「正確性」高，或「效度」高(Validity)。正確性或效度是一個測驗最重要的條件。要編造一個正確的測驗，對於測驗的材料，須經過審慎的選擇。編造算術測驗時，所用的材料，必須包括算術上各種重要的計算能力；編造歷史測驗時，所出的問題，必須是歷史中重要的部分。而且在測驗中，應當把與所測量能力不相關的分子完全除去。

(二)測驗的可靠性須高 所謂測驗的「可靠性」(Reliability)，又稱測驗的「信度」。這就是說，一種測驗，測量同一個兒童兩次或數次，若所得的結果相同，則這種測驗的可靠性很大，或信度很高。測驗的可靠性是很重要的，不可靠的測驗，就沒有多大的用處。假如用尺去量一塊布，第一次量得一丈，第二次量之為八尺，這把尺還有什麼用處？心裏測驗也是這樣的。假如用一種智力測驗去測量某兒童，今日測量的結果為智力年齡八歲，明天測量的結果為智力年齡九歲，這個智力測驗就沒有什麼功用可言了。

(三)測驗材料的難度的分配須寬廣，俾能鑑別學生的程度。例如有兩個兒童，他們的算術程度相差很微，普通的觀察不能鑑別二人算術能