

應用智力測驗的態度與認識

郭為藩主講 林振春記錄

壹、前言

各位老師以及各位參加這次研習會的學員；很高興有機會與大家見面，今天，我同時代表學校歡迎參加這次研習的各位老師到師大來。這兩天，曾一再考慮究竟與各位討論什麼樣的題目，最後決定藉此機會談些智力與智力測驗的基本概念。我希望能唱些反調，讓各位能從不同的角度來瞭解問題。

我想先說明一下我的動機，我們常聽到一個故事，以前人們上山拜師學武藝，在學成即將下山時，老師父總是告誡徒弟，下山以後少顯露自己的身手，免得功夫太好，傷害到別人。各位到此來學習輔導評量的各種技巧，特別是學測驗的技術，我們也是希望各位能善用這些技術，免得傷害到別人。

記得前幾天的報上有個有趣的新聞，說印度有一個醫生，為病人檢查眼睛，病人明明是右眼有毛病，結果竟把左眼挖掉了。像這類的事情，其實也見怪不怪，走江湖的密醫，誤了病人的健康，也是常有的事情。各位不要誤會我把各位比喻成密醫。事實上，很多外科醫生喜歡開刀。因為他們想增加開刀的經驗，特別是很多實習的醫生。因此有些本來不需要開刀的，常被他們抓來開刀，試試身手，以累積自己的經驗，遭殃的却是病人。因之，今天我的意思並不是跟我們講測驗的教授唱反調，而是想提醒各位在做測驗時要慎重其事，非有必要，不貿然從事。

我們常有一種觀念，認為做測驗一定比不做測驗好，英文常有：「Diagnosis is better than no diagnosis」，診斷比不診斷好。還有我們過去常有一個觀念，認為愈早的診斷或測驗，常是愈好，因

為預防勝於治療，可是事實上，今天很多看法不見得如此。有些人做診斷或測驗，祇是為測驗而測驗，為診斷而診斷，那麼這種測驗或診斷便沒有價值。測驗的目的是為了要輔導，診斷的目的也是為了要補救教學，如果沒有接連的補救教學或輔導工作，而祇是新生入學後便例行的做個測驗，我認為這種測驗是害多於益，對教學沒有什麼好處。這不僅是我個人的看法，事實上也是很多人共同的想法。其次我所要強調的一點，就是說我們測驗也好，診斷也好，必須要講求正確，尤其是避免對智力祇取一個大略的量數，而不做內在能力的分析。我今天之所以提出這一問題，也是因為今天很多輔導人員對測驗技巧本身一知半解，常祇求「IQ」這個量數，且認定「IQ」是絕對不變的，並且用來概論學生的學習能力，結果却害慘了學生，所以我們認為一知半解比完全不懂來得糟糕。我們並非反對測驗，而是要建立各位對測驗的正確認識，同時不隨意地應用到學生身上。

各位可能在課堂上提到「自行應驗的預言」(self-fulfilling prophecy)，有時也稱為Pigmalion effect (比馬龍效果)。Pigmalion 是蕭伯納 (Bernard Shaw) 在一九一二年寫的一本書，大抵是說塞浦路斯有一位王子，心目中經常有一美女的影像，這一美女並不存在，可能是他看到許多美女的綜合影像，所以他找了一位雕塑家，依照他心中的影像，雕成一個非常生動的美女塑像。後來這一王子竟然迷上了這一塑像，他的純情感動了天上的神，賦予塑像生命而變成活生生的人，後來王子就跟美女結婚而有了後代，就是今天的塞浦路斯人。

後來有人把這個故事用到一本書上，就是一九六八年羅桑沙（R. Rosenthal）與傑柯布遜（L. Jacobson）合寫的「教室中影像的塑造」（Pigmalion in the Classroom），其意義乃指很多老師對學生常有先入為主的成見。比如說經過測驗或鑑定後，就說這位學生是問題學生或智能不足，或是經由編班編到壞班，老師就對學生有了這種觀念，學生受到老師的影響也就有了消極的學習態度，因此，學生就認為老師既然如此看他他是壞學生，那他自己大概就是壞學生，有了這種自我認同，慢慢的，他的功課及各方面表現也就越來越差。到此，老師却相信學生的表現正印證了他的預言。因之，他又以對付壞學生的態度來教這學生，這就形成了惡性循環。老師先對學生有一種期待，學生接受此影響而形成一種自我觀念。

我們常看到有些老師討厭學生，學生當然也不可能喜歡老師，慢慢地功課落後，老師更討厭他，學生也就更不喜歡老師了。這是一種惡性循環，我們的輔導工作，就是為了打破這種惡性循環。

其實這種循環的開始，常常是心理學上所謂的標記（labeling or categorizing）。就是說，我們常有意或無意的相信某一類兒童，他們在行為特質上是一樣的；如同從商標推論商品的品質，現在有些雜牌的電視機，同樣是廿吋的，要賣兩萬塊錢，而有些名牌的電視機如新力、國際等，要賣兩萬三，可能您會買兩萬三的，而不買雜牌的，因為您有一個先入為主的觀念，以為新力或國際牌的产品一定比較好，儘管那些雜牌的較便宜。所以只要我們對學生有了成見以後，立刻對他的態度就不一樣。

有時候我們常認為買貴了東西，品質總是較好，事實上，這個想法有待商榷。這種先入為主的成見，在社會心理學上，有個名詞叫「自行應驗的預言」——我們在心理上期待某種現象的存在，可能會真正導致此種現象的發生。如同以前日本發生衛生紙缺貨，有些日本觀光客到台北認為是最後一站，就買了兩包衛生紙提回去。海關人員看到衛生紙缺乏，回去告訴他太太，他太太就多買兩包。每人多買兩包，價格當然稍微上漲，報紙再加以渲染，很多家庭便加以囤積，因之，很多家庭主婦常是一手五

包，兩手十包的提回家。我們的市場上本來供應非常平穩的，經大家一搶購，自然衛生紙就缺貨了，當然價格更為上漲了。另外前幾年的硬幣短缺也是一樣。這種自行應驗的預言，在社會現象中經常存在，有時候我們相信某項現象存在的預言，結果真的在事實上應驗了。假如我們透過測驗，而把一個學生看成是智能不足或能力很差，我們對他的教學法、對他的態度各方面也就慢慢地產生符合那種成見。

接著，我想再把測驗的背景稍微地說明一下。測驗在美國早些時候，成爲一種新興的教育企業，很多人靠測驗吃飯、靠測驗賺錢。很多測驗的服務社在美國非常普遍，除了很多人編測驗可以收到很高的利潤，還培養很多專門做測驗的行家。而形成一種專業。這種專業興起的結果，使得很多人濫用測驗，把學生做了各式各樣的標記。另一方面，測驗的應用還牽涉到一個困擾美國社會的種族混合問題。

各位可能記得數年前，聯邦政府透過法令，用交通車使黑人與白人兒童混合就讀同一學校，那時曾發生很多暴力事件，把學校交通車放火燒掉。雖然動用警力強制執行，維持了混合就讀，可是學校透過測驗，把黑人子弟分到特殊班去，或是加以能力分班，結果某些班仍是「黑漆漆的一片」，因之，這種測驗被誤用而做爲一種種族歧視的工具。另外一種情形與美國特殊教育這些年來發展很快有關，各地教育機構爲爭取特教經費，難免巧立名目，以不同名稱來標記學習障礙的小孩，如學習缺陷兒童，以設立特殊班。

目前，美國所稱的特殊學習缺陷，起初估計其出現率爲百分之一到百分之三；到最近，其出現率竟達到百分之十五，甚至百分之三十，一百人當中，有十五到三十人是特殊學習缺陷，愈講愈多。爲什麼會有這種情形呢？因爲特殊學習缺陷與智能不足一樣，這些兒童常被認爲與神經系統的功能障礙有關，在智力或感官方面是正常的，除了某一方面的學習困難與成就低劣，其他方面可能完全正常。因之，很多人利用這個機會把自己的子弟放到學習缺陷特殊班，以得到好老師的指導。因爲特殊班學生人數較少，老師也多受過專業訓練，具有補救教

學的技巧與耐心，這點也是測驗被誤用的一例。

當然，物極必反，今天有很多人對測驗提出檢討，前些時，*British Journal of Educational Psychology* 有一位教授寫了一篇文章，檢討五十年來的測驗發展趨勢，就說明最近何以很多測驗走下坡，此類問題，都是值得我們重視的。我們的看法是；測驗並非不好，而是常被許多人濫用。很多用測驗的人常有目的地去印證自己心目中已存在的假設。如同做研究，即有些研究者心理先有一種假設，因之，在求證的過程中，合乎自己假設的資料，明察秋毫，鉅細不遺地把它們蒐羅殆盡，不合乎自己假設的，難免就睜一眼，閉一眼，有意或無意地忽略過去。正如同我們現在想成立一特殊班，老師一定睜開雙眼去看那位學生合乎這個條件，就把他拉進來放在特殊班裏頭，惟恐這個班的人數不足，不能成班。因此，我們並不是不重視測驗，而是要把很多測驗工作與教育診斷串連在一起。

如同美國的伊利諾心理語文測驗（*Illinois Test of Psycholinguistic Abilities, ITPA*），目前應用日廣。這類測驗工具，不求IQ的多少，而分析學生內在能力的長處與弱點。這種分析、診斷的工具，慢慢地取代了祇求智商或總量數的測驗，這是目前表現的一種趨勢。

另外一種趨勢是常模參照測驗（*norm-referenced test*）被效標參照測驗（*criterion-referenced test*）所取代。所謂常模參照測驗，強調測驗結果跟別人常模的比較，以瞭解受試的地位，多數智力測驗，屬於此類。效標參照測驗則是先訂效標，如同分析兒童要學會這門功課。要具有那些能力，把這些能力的因素列出來，我們測驗就在瞭解學生在這些能力方面，那些已足夠，那些還不夠，俾有個較深入的瞭解。還有一種像皮亞傑式的測驗，祇求分析兒童各項認知能力發展到何階段，而不籠統求一概括的智商。此種觀念和作法，目前在國內，逐漸的普遍。而以前那種測驗，會不會再回過頭來，很難瞭解，就像領帶寬了又窄，窄了又寬，物極必反，不過目前的趨勢是比較注重內在能力的分析。瞭解學生智力發展的過程屬於那一階段，而不太注

重數量化，更不太拿個人的成績跟他同班的學生作比較，這種比較，我們認為沒有什麼意義。

此種趨勢，也表現在特殊教育方面，我們較注重教育的導向，代替過去特殊教育受醫學的影響而形成的醫學導向，過分重視分類與病因的瞭解。今天，在教育上，我們重視的是學生能學些什麼？要學些什麼？他在學習上有那些困難，這是每位老師所應該瞭解的。至於他是否智能不足，蒙古症，或有何病名、成因，都跟老師的教學沒有關係，我們所要瞭解的，是他的能力怎樣，我們如何教他，才能教得好，讓他有興趣，這是我們學教育的人應有的態度。

學醫學的人，很重視這是什麼病因，屬於何類，這種作法並不是學教育的人所贊同的。今天祇就觀念的轉變，特別提出來，給各位作一個說明。以上是關於智力測驗的背景，我提出來，主要是先讓各位對智力測驗有一正確的看法，底下我們談到主題。

貳、智力的基本概念

評量智力及估算智商是兩件事情。作智力測驗或透過任何其他評量工具，來瞭解學生的智力是一件事情，求出一個代表學生智力的商數，又是一件事情。我們過去習慣的作法總是不注意去瞭解智力，而祇希望求出最後的結果——智商。就如同很多人看小說，不喜歡從頭看到尾，而急著想看最後的結局一樣，致於書中構局的奧妙及文章的優美，從不想好好去欣賞。今天有很多人作測驗祇為知道一個智商，當知道智商多少，他就滿足了，這種作法是我們首先要避免的，為什麼呢？我想大家應知道測驗是從比奈（*Afred Binet*）開始的，比奈對於智力測驗與心理測量的最大貢獻在於他提出一個整體而概括性的智力觀念（*The concept of a pervasive general intelligence*）。智商並不是比奈提出來的。比奈當時祇應用「心理年齡」，智商是後來的推孟（*L. M. Terman*）所創用的。

比奈當時認為智力是一種普通能力，祇表明一個兒童智力發展的層次到達何種階段，這種想法與皮亞傑的觀念是相當一致的。他不重視最後

的量數——智商。智商之類的概念雖然很有用，但是危害性也很大。我們看到很多學童爲了名次，不知受到多少委屈。現在很多學校老師的考試很少具有鑑別能力，學生常因一分之差而從第二名落到二十幾名；可是學生家長不瞭解，一看到這情形，就以爲是學生成績退步。事實上，很可能只是學生不小心做錯一題，可是家長却焦急萬分，責怪子女的不是。教師對學生的看法也不一樣，連帶影響學生對自己的看法，形成消極的自我觀念，這可以說名次害苦了很多小孩，尤其在能力分班中，分在高能力班的學生，競爭的目的只在計較名次的前後。因此，他們把書本背得滾瓜爛熟，惟恐稍一疏忽，名次就退了很多。那時便會受到各方的壓力，這在教育上是很不正常的現象。今天，我們做測驗的目的是在瞭解學生的智力，但是過分注重智商，祇是給學生貼上一個標記，對學生並沒有多大的意義。

其次，我們知道，智商是相對的，我們常把智力測驗稱爲 intelligence scale, “scale” 是一種量尺，譬如我們想量這張桌子有多寬、多長、多大，我們必須靠標準化的量尺。但是一個小孩的智商有多少，那是相對的，智商是拿一個小孩的智力發展與他同年齡、同文化背景的很多小孩來作比較，所得到的一個商數。這種比較是相對的標準，而不是絕對的標準。舉一個較爲具體的例子；我國南方小孩身高一七〇，我們說他很高了，可是把這小孩拿來跟美國或我國北方的小孩比，這一七〇就不見得高了。我們也看到今天的小孩在營養方面較好，身高普遍增高，可是早十年的小孩就沒有這種成長環境。十年前的小孩，身高一四〇可能就很高了，可是今天則算相當矮，智力的情形也是一樣，如果您拿十年前編的測驗所訂的常模，來測驗今天的小孩並加以比較，您會發現處在今天所謂「電視時代」的小孩，其智商普遍有偏高現象。

所以我們知道智商是相對的，一個小孩的智商九十五，是跟他同年齡、同時代的小孩相比較，他的智力是在中等或接近中等的地位。譬如有一百名小孩，他可能排在倒數四十五名，有五十五個小孩超過他。所以智商是相對的，是建立在

一個常模之上，如果常模不標準，不隨時代而修正，那藉此所得到的智商將毫無意義可言。

再來，我們談到智力評量的三個層次，赫伯（D.O.Hebb）於一九四九年把智力分成智力甲（Intelligence A）與智力乙（Intelligence B），智力甲代表智力發展的潛能，是一個人天生俱來的潛能，由基因所決定；智力乙是代表現在已發展出來的能力。甲、乙之間，一定有差距存在。心理學家卡特爾（Raymond Cattell）則分爲流體的智力（fluid intelligence）和結晶化、具體化的智力（Crystallized intelligence），略等於智力甲和智力乙的意義，這兩者之間必然有差距。如果我們再加另一個智力丙，即 Intelligence C，也就是從智力測驗所得到的，用智商或其他量數代表的智力，我們相信A與C之間固然會有差距，智力乙與智力丙也必然有差距，因爲就是最好的測驗也不可能把完整的智力百分之百的量度出來。

但是今天有些人強調智力是一種潛能，有些人強調智力是現在具有的能力，又有些人強調操作性的定義（operational definition），認爲智力只是智力測驗評量出來的量數。這種情形，好像我們所講的民主與社會主義共產國家他們所稱民主共和國的民主，並不一樣，他們雖然開口閉口民主，可是却不民主，其民主的條件跟我們真正民主國家所講的民主完全不一樣，每個人用同樣的名詞，可是觀念却不同，在智力方面也是一樣，何況今天很多智力測驗的編製者，他們對智力的觀念和看法也不完全相同。例如瑞文氏所欲評量的智力與魏克斯勒（D. Wechsler）所欲評量的智力並不盡同；他們的智力觀念跟比奈、西蒙當時所構想的智力也略有差異。如同這是一張桌子，有人度量它的左半部，有人度量它的右半部，雖然同樣測量這張桌子，所得到的結果必然有別。所以如果我們做三個智力測驗，所得到的智商完全一樣，那反而是很奇怪的巧合了。我們同時做瑞文氏測驗，魏氏兒童智力測驗，比西量表，雖然這些好的測驗應該有較高的準確性，但如果所得出來的智商完全一樣，實在是一件非常令人驚異的巧合，我們常說瞎子摸象，今天我

們對智力的觀念，也有一點類似。

各位知道基爾福（J.P.Guilford）把智力分為三個層面（three dimensions），可分析成一二〇個因素，任何一個測驗在有限的時間內，只能包括很有限的題目，欲求把所有一二〇種因素完全衡量出來，事實上是很難的。我們的基本假設在這些能力間有高度的相關，但不可能在不同的測驗裏，能測量出一致的東西，此外，塞斯頓（L.L. Thurstone）倡智力多因說，他認為智力有語文瞭解（verbal comprehension）、語言流利度（word fluency）、空間關係（space）、數字應用（number）、記憶（memory）、推理（induction）等重要因素，有些智力測驗偏重於衡量某部份的因素，另有些智力測驗則偏重於另一部份的因素。如比西量表在不同階段年齡的題目所測量的能力各有所重，由此我們得到一個觀念，智力測驗在衡量普通能力，但是不同智力測驗所評量出來的普通能力並非完全一致。智力為一整體的概念，各種能力間雖有高度相關，但不宜視為一單純的能力，而實有其不同的層面，這點認識對我們來說相當重要。因此，我們做智力測驗，不在求最後的結果，而應進一步做內部因素的分析。如我們做魏氏兒童智慧測驗（WISC），應就每個分測驗評量的結果加以比較、分析，因為每個分測驗所評量的能力並不一樣。甚至在比西量表中雖沒有分測驗，但是我們也可以從不同的項目來分析內在能力的結構。我舉這些例子，主要目的仍在強調不要祇重視測驗的最後分數，而應去分析受試者內部能力的長短處與差異。

智力也牽涉到一個問題，那就是遺傳和後天發展孰較重要的爭論。有些重視遺傳的學者，特別強調智力測驗在評量個人的潛能（智力甲）；重視後天環境的學者，強調智力測驗在評量智力乙（Intelligence B），以我個人的看法，智力測驗與其說在評量潛能，不如說在評量現在的能力，但是所評量出來的與事實仍有相當的差異，如同在智能不足觀念剛流行的早期，很多人強調遺傳的重要性，在一九二九、三〇年代，很多學者認為智能不足，有百分之八十至百分之九十

屬於遺傳。在一九四〇年代，杜爾（E. Doll）推斷智能不足的成因，約百分之三十屬於遺傳，這種比率後來逐漸降低，可見隨著時間的改變，我們對智力或智能不足的看法也有差別。

接著我想提出一個有關智力的重要事實，我們的智力包括很多種能力，所以有些小孩一、兩種能力特別優異或特別低劣，都不足以一概而論，我們常在報上看到某一神童計算能力很強，可以跟計算機比較，這不足以立刻斷定他就是天才；同樣的，某個學生在某一方面的能力特別差，我們也不能立刻說他是智能不足。譬如我個人在短暫記憶能力（short-term memory）方面特別差，很多同事祇要上課點一、兩次名，全班同學的名字差不多都可以叫得出來，可是我點了十次名仍然記不住，雖然我記性很差，可是各位不要以此事認為我就是智能不足。有個笑話，有一位老教授去拜訪朋友，按了門鈴之後，他却突然忘了主人的姓名，門房問他找誰，他支吾半天仍想不起來，祇好說：「我找您們主人！」那門房問他：「您找那個主人？」他真是尷尬萬分，剛好主人探頭出來看，才解除了這個窘境。

有些學生記憶能力特別差，千萬不要以此論定他智能不足，很可能他在其他方面學習能力很不錯。就拿記憶來講，有些人視覺記憶較差，有些人聽覺記憶較差。譬如有些學生背書要誦讀出聲才能記得起來，要是他不唸出聲音來，他就很難記住，這是偏於聽覺記憶；有些人則傾向視覺記憶，背誦十遍不如眼讀一遍。這種差異也好像吃飯，有些人吃得快不見得不消化，有些人吃得慢也不見得消化好。吃飯速度的個別差異很大，學童的學習能力也是一樣。有些人學得很慢，可是穩紮穩打，學得很切實；有些人學得快，可是也不見得就很潦草。所以在測驗裏要慢慢去分析，從分數看不出這類長短處。

今天在智力的鑑定上，注重多層面的能力鑑定。瞭解兒童某一特別方面的能力很好，就要加以發展。今天人們對資賦優異的觀念也是多元的看法，有些創造力特別高，這些人的普通能力智商不見得特別優越，雖然兩者之間往往有高度相關，然而在智商超過一二〇以上，關連性就減低

。故創造力高的人，多偏重在擴散性思考（*divergent thinking*），腦筋動得很快，可是功課不一定特別好；另一種普通能力高的人，偏重在聚斂性思考（*convergent thinking*），比較專心致力在某一方面，功課可能較好。所以有時候我們說胡適之先生是博學多識型的學者，各方面都懂，就不見得在某一方面特別精通；董作賓先生是專家型的，專心致力在某一方面，成就很高，也不見得學問淵博。如要胡適之先生做董作賓先生的考古工作，他可能做不來；反之，要董作賓先生做胡適之先生的工作，他也不一定做得好，所以學者可以分為兩類，一種是思想家型的，海濶天空，涉獵很廣，各方面都有興趣，却不容易深入；另一種是專家型的，把精力投注在某一專門領域上。以上的例子，用來顯示智力雖是同一樣東西，可是每個人在智力結構上却有相當大的個別差異，我們求出來的，祇是一種概括的事實，如同我們在坐的老師分為三組來做智力測驗，可能得出來的平均智商假定都是一〇〇，可是標準差就可能有很大的差別，這就是從團體來看，三個團體的平均數可能一樣，但是其組合就不會一樣，因此，三個學生，同樣的智商，其代表的學習能力，也不會一致，這是我們應該瞭解的概念。

接著我們來探討純智力是否存在。我們都瞭解智、情、意三位一體，原是很難加以分開，祇有在學術討論上，我們才將之劃分。在實際的情境，特別是很多測驗的情境，或是工作的情境，智情意是不能劃分的，因此進行一個智力測驗的結果，常常與情緒的穩定性和成就動機有密切的關連。測驗的結果，不完全是智力，還有其他的配合條件，如同情緒及動機方面的因素。我們常看到有些小孩求好心切，或過去有較多挫折的經驗，常常在測驗的情境顯現高度焦慮，因而壓抑了他的智力功能，沒有正常的表現，所以我們常看到有些孩子平常很好，可是到了測驗的時候，表現不很理想，因為焦慮壓抑了智力的表現。因此在智能不足觀念裏，常有假性的智能不足（*pseudo-retardation*），亦即不是真正的智能不足，而是可能在智力測驗的時候，因焦慮而壓抑

了他智力的正常表現，也可能是因為聽力障礙或其他心理衝突而使得他在一段時間內，智力功能無法達到應有水準。成年人如因家中有事煩心，在辦公室工作的表現便不會很好，有些人因為失眠，在一段期間內工作都做不好，假使剛好上司來考查他的能力，他不就倒楣了嗎？有些學生也有這種情形，記得在一次台北市小學啟智班的研討會上，有位小學校長說：「我們有一位學生入學時，智商是六十三，經過特殊班三個月的輔導，老師愛心、耐心的教導，智商提高至九十七。」他以此來說明特殊班的效果。其實，這兒童開始的六十三，可能是評價偏低，後來的智商九十七，可能是偏高，或者可能是他們的測驗有問題。特殊教育雖然有效果，可是三個月內，智商提高那麼多，那是不可思議的事。

這裏還有很重要的原因，那就是可能這個學生開始在普通班，老師見學生考試不好便要責罵，無形中使學生對考試產生一種抗拒與焦慮，所以考試情境中不可能把真正能力表現出來，測驗也是一種考試情境，所以結果不合理想，智商當然不可能高。等到換了一個教學情境，焦慮漸漸減除，學習動機也提高，老師又用成就感教學來鼓勵，培養他的自信心，每次考試後，老師總要稱贊他幾句，因此他對考試很感興趣，全心全意地做，結果當然不同。由此，我們可以知道，測驗的情景，受試者與主試者之間的關係具有重要影響。很多沒有經驗的老師，學生一進來，便開門見山的問他是什麼名字，然後把測驗卷發下來。三言兩語解釋過後就做一、兩個小時的測驗，然後求出一個結果，這種測驗一定不可靠。尤其學生沒有做過測驗，心裏沒有準備，這種測驗方式是很危險的。

記得以前我在國外學做心理測驗，一種測驗差不多要學一個月，連續做很多次，老師且一再告誡。測驗開始前一段時間，先要跟學生聊聊談談，使學生對您有信心，喜歡跟您接近，此種安全感很重要。這跟第一次見面便開門見山，單刀直入，差異當然很大。由此，我們可以知道，智力測驗所求得的不完全是純智力，智商也不是純智商。我們常發現，智商高的人，學習能力不一

定高，如果他情緒不穩定，注意力有問題，缺乏學習動機，就常有這種情形；反之，智商低的人，如果成就動機強烈，也能勤以補拙，學習一些。因此，我們對於智商不要遽下結論，應先存疑。這是我一再向各位提醒的。

叁、有關智力測驗的認識

我們對智力測驗應該有什麼樣的看法？第一點：智力測驗的結果必然的存有少許誤差。庫龍巴（Lee, J Cronbach）寫過一本很有名的心理測驗（Psychological Test）的書提到，每次智力測驗都可能有五分的誤差，因此測驗結果，智商如為九十，應該稱之為八十五至九十五之間，也就是說，受試的智商可能落在八十五至九十五這段範圍，這對量數絕對論者是一個很重要的提示。另外，即使是主試者受過訓練，由於受試者的身心狀況（如極度疲勞），當時的精神狀況（如分心、抗拒），測驗當時的情景（如不時有人打擾、電話聲），也經常產生誤差。例如一個學生剛上了四節數學課，已經累得很，您要他中午來做測驗，或他中午沒吃飯、昨晚沒睡好，或測驗過程的外來干擾，都可能影響測驗的結果。第三點是測驗主持人對受測者的態度有很大的差別。美國的研究發現，白人做黑人小孩的測驗與黑人做黑人小孩的測驗，得到的結果常常不一樣；白人做黑人的小孩，其結果往往偏低，黑人做黑人小孩，結果則易偏高。當然這牽涉到種族的成見，問題可能很複雜。不過，我們可以確信，假使有一種期待的心理，會受到一種期待效果（expectancy effects or halo effects）的影響。譬如說：我們作學術研究，合乎你研究假設的資料，你自然而然的會把它用進來，跟你假設相出入的事實，則難免睜一眼閉一眼的，盡量不去理睬它。今天如果你做測驗的目的在找特殊班的兒童，自然而然的會對受測兒童產生期待效果。一個人的態度對某人產生懷疑時，祇要他稍微表現某一行動，即會附會於這種涵義。譬如這位你所不喜歡的朋友，很可能是真誠來向您道賀，可是您會懷疑那是一種諷刺。一個人對測驗的態度有別，一定也會影響他對測驗結果的解釋。在社

會心理學上，態度影響於行為是一種明顯的事實。如果在做測驗之前，就假定他大概是智能不足，結果也差不多會得到相同的結論。

另外一種對智力測驗的結果發生影響的是教練（coaching），一個小孩是否做過測驗，或經常做測驗，或有人補習測驗，都會產生影響。在資賦優異班裏頭，很多家長幫學生補習測驗，甚至現在有很多補習班，把測驗拿來給學生做。因之，做過測驗的學生，難免佔了便宜，從來沒做過測驗的學生便要吃虧。有時候常做新式考試也會有所幫助，因為考試的情境與做測驗的情境很相似。有時候，我們說某人是考試專家，每次考場，他都是無往不利。考試也有技巧，在組織能力強，能抓住要點，寫得清楚，總佔一些便宜，所以我相信考試有專家，考試經過訓練可以有較佳的結果，所謂「教練」，多少有效。

外國大學人員常奇怪中國學生考托福英語能力測驗，常超過六百分的成績，可是很奇怪，一到美國又用不出來。我不好意思說這是練習應付考試的技巧方面，我們比美國略勝一籌。由此我們可以知道考試情境與測驗情境很相近，有系統的訓練會影響結果。

下面談到智商的穩定性。過去心理學上經常有智商能否改變的爭論，個人認為智商在相當幅度下改變是可能的。其原因如下：（一）有些人發展有特別早熟或晚成的情形，有些人開始發展很快，有些人是大器晚成，智商是與常模相比較而得的，我們如以身高作比喻，如果某班學生係以身高來排座位，某生一年級的時候排在第四十位，可能二年級時進到第三十五位，三年級進入第三十位，四年級進到第二十五位，五年級進到二十位，也就是身高高於一般的水準，那他到六年級可能進到十名內了。各位老師都知道，有些小孩子的身高往往發展不一致。座位也就要跟著調整向前或向後，智力何嘗不是這樣。智力拿來與同年齡、同背景的團體常模相比較，也會有調整，智商也就有差異。

（二）特殊環境對智力發展會產生影響。如我們選擇兩組智力相當的小孩，一組放在一般家庭，一組放在寄養家庭，經過三年後，我們再拿來比

較，智商便會有顯著的差異。又如在美國實施補償教育，對語文、知覺動作加以訓練，或受到好老師耐心、愛心的指導與鼓勵，對智商也會產生改變。

ㄅ對測驗的訓練、熟悉、或適應了測驗情境，也能影響測驗結果。雖然這不是智商真正的改變。因之，我們說智商並不是不能改變的，因此在心理學上談到智力發展的預測時，都清楚在學前階段做智力測驗，它的預測可靠性很低，愈是往後，預測的可靠性愈高。還有預測一個團體的智力發展，比預測一個人的智力發展較為可靠；因為預測團體時，團體成員互相產生補償作用，正負相抵消，而有趨中的現象。

我現在將有關智力測驗應有的認識歸納為下列幾點：

(一)我們千萬不要一試定終身，不要祇測一次便認定他為智力不足或資賦優異，我們知道智商在相當程度下可以改變，最好每一年測一次。同時不要每次都做同一種測驗，最好多做幾種不同的測驗，再來衡量他的智力水準。

(二)不要過份強調最後的量數，而要分析內部能力的長短。

ㄅ應該多用教育診斷式的測驗或效標參照測驗。少用常模參照測驗或純為求智商的測驗。為何要用診斷測驗呢？我們知道一個學生成績不好，可能是整個學習過程中某一階段的障礙，如同「生產線」的原理，在一連串的生產品中，假使可分為十個階段，若其中一個階段發生故障，當然會影響到整個生產線的功能。我們若祇注意到整個生產效率低，而不知去改善有故障的階段，當然難以發揮效果。例如我們拿撐竿跳高來看，看起來是一個動作，若拍成影片以慢動作來分析，可以分成踏步、起跑、接近竿時拿竿的動作、撐竿的動作、翻身的動作、下來的動作，至少可以分為十數個動作，我們祇要能分析他在那個動作上有毛病而加以修正改進，就能夠影響他的成績進步。很多學生的學習成績不好，不完全是智力低，而是智力的某項因素或注意力、或視知覺與視動協調的問題。但是一個階段的不足，便可以使整個結果受到影響。因之，我們知道教育診斷之所以需要，在它看出毛病，一般的智力測驗，祇標示受試的等級而不能找出真正的毛病。所以我今天不是暗示各位不要做智力測驗，而是建議大家，兼採教育診斷測驗來幫助學生學習。

(主講者現任師大校長)

一九八一年國際特殊教育研討會

研討時間：民國七十年五月廿九日至五月卅一日

研討地點：國立台灣師範大學綜合大樓國際會議廳。

研討主題：“七十年代特殊教育的展望”。

主辦單位：中華民國特殊教育學會及國立台灣師範大學。

贊助單位：教育部、台灣省教育廳、台北市教育局、高雄市教育局。

與會人員：國外部分—來自美、加、日、韓、香港、新加坡國外學者專家共十四位，包括 Dr. Goldstein, Dr. Reynolds, Dr. Nix, Dr. Aschcroft, 三木安正、佐藤親雄、金誠一、陳啓祐等人。

國內部分—特殊教育專家、特殊學校校長、特殊班教師代表及教育行政機關代表約 100 名。

研討內容：五月廿九、卅兩日，研討重點以國外學者專家為主，各就智能不足、聽障、視障等各類特教問題進行專題演講或報告。

五月卅一日為本國特殊教育研討會，進行分組研討以下問題—特殊教育法之研擬；教師進修與研究發展；人員、經費與教材教法；視障教育、聽障、語障及智能不足教育問題。