

力包括機械理解、動作能力與空間關係等。在較小群因素之下是屬於最低層次的特殊因素。此理論統合了上述Spearman 的 g 因素，Thurston 的群因素和Guilford 的多因論，因此最受美國心理學家的喜愛。

5.其他的理論

Cattell (1963, 1971) 的因素分析理論主張智力可分為流體智力 (fluid intelligence) 和結晶智力 (crystalized intelligence) 兩種。流體智力主要來自遺傳，是各種不同領域所共有的。具有較多流體智力的人能夠做好許多不同的工作。這種智力與一個人的學習和速度有高度的正相關，但它不受環境經驗或文化背景的影響。結晶智力則相反，它是一種經學習和經驗而逐漸累積得來的智力，可被看成是一個人智識和專門性技能之總體。結晶智力是某些領域所獨有的，欲在特定的工作中獲得成就，需要具備此種智力。在固定習慣性的工作中，使用較多此種智力。流體智力通常是採用文化公平測驗測量，結晶智力則採用語文測驗測量。Cattell 研究發現，流體智力發展到 14 歲或 15 歲為最高峰，但結晶智力則持續發展到 25 歲或 30 歲。

另外，Jensen (1969, 1973) 提出二層次理論 (two-level theory) 。第一個層次為聯想智力 (associative intelligence) ，包括機械學習能力和記憶能力；第二個層次為抽象智力 (abstract intelligence) ，包括思考推理與問題解決能力，類似於 Spearman 的 g 因素或 Cattell 的流體智力。

四、題庫擴充時所需之等化與連結

題庫擴充時，新題目的量尺必須與原有題庫的題目參數之量尺相同，因此就需要用到等化與連結方法。

(一)等化

等化 (equating) 是將一式測驗上所得之分數，以統計公式轉換調整至另一式測驗分數量尺上的過程 (Dorans、1990) 。基本上等化乃是決定二個或二個以上不同式測驗 (或量尺) 分數之間關係的過程。換句話說，也就是要將不同式測驗結果的分數放到同一量尺單位上，其過程乃在於調整題目及測驗特徵的差異 (Hung、1990) 。

當測驗結果所做的決定具有很大影響時，基於公平、保密等之考慮，通常必需編製許多平行測驗，如果二式測驗完全合於平行的定義，等化就不需要了，然而編製嚴格平行測驗幾乎是不可能的，故只有在測驗分數經過適切的等化後，才能有意義的、公平的比較測驗結果。而且在發展標準化測驗的過程上，如修訂、增加題目……等，等化均是一項非常重要的工作，藉此新、舊版測驗間才能相互比較，題庫才能加以擴充，有關文獻的連貫也才有可能。等化有下列兩種型式：

1. 平行等化

有某些考試（如托福、GRE 等）需要有多式複本測驗，才能在一年內實施多次測驗。但是，如果不同式測驗分數的分佈並不相同，那就需要決定不同式測驗相互等值的分數。此種為多式複本測驗建立相互等值分數的系統，稱之為平行等化（horizontal equating）。平行等化的目的在於編製含有多式，但在難度上相平行的測驗。

2. 垂直等化

有某些測驗組（test battery）包括不同難度水準的測驗，如「修訂 Otis- Lennon 心理能力測驗」的水準三適用於小學一年級下學期至小學三年級下學期，水準四適用於小學四年級至六年級。在這些心理測驗中，題目難度的設計原本就有顯著的不同，要比較學生在這兩個測驗上的表現，就需決定兩個水準相互等值的分數。此種過程稱為垂直等化（vertical equating）。垂直等化的目的在於編製同一特質而含有不同程度的測驗，並且這些不同水準的量表須使用相同的計分量尺，通常用於對個體長時間學習發展的縱剖評估。無論在理論或實施方面，垂直等化皆比平行等化複雜多了。

(二) 連結

連結（linking）是 IRT 等化中一個必要的步驟，其目的是將所有題目的參數放到同一單位的量尺上，使題目的參數經連結後可互相比較。以下提出一些學者對於 IRT 連結的研究報告：Vale、Maurelli、Gialluca、Weiss和Ree（1981）對共同試題題數及其訊息曲線的形狀與等化結果的關係進行探討，結果發現 15 題至 25 題共同試題是必要的，共同試題的訊息曲線以長方形或常態為佳。尖峰式的訊息效果較差。

Mckinlex 和 Reckase (1981) 認為 5 題共同試題是不夠的，25 題的結果較 15 題好，但如採用同時估計法則 15 題共同試題即已足夠。Raju, Edward 和 Osberg (1983) 於垂直等化的研究中發現 6 至 8 題的共同試題就已足夠。Wingersky 和 Lord (1984) 發現當採用同時估計法時，共同試題甚至於可少至 2 題，只要估計誤差小，其效果與 25 題不相上下。

Cook 和 Eignor (1981) 對等化設計、共同試題、其估計標準誤和考生能力分配形狀進行研究，發現同時估計法與特徵曲線轉化法並無差異，但同時估計法較不需要太多的共同試題。共同試題估計標準誤儘可能接近原測驗的估計標準誤，以長方形分配為佳，尖峰式的分配較差。Wingersky, Cook 和 Eignor (1986) 以三參數模式和模擬資料進行研究，發現以較長的共同試題進行連結時，能提供較穩定的結果。Hung (1990) 以模擬資料進行研究，結果發現就難度而言，連結對於真值與估計值及預定值與估計值間誤差的減小，是有幫助的；但對鑑別度則不然，連結會造成更大誤差、偏差、甚至系統性的偏差。在四種連方法中，以特徵曲線轉化法和平均數、標準差法二者對難度的連結結果最好關於考生特質方面，當樣本中所缺少的考生的能力位於該測驗能提供較豐富訊息的能力點上，則連結對難度量尺可有較大的助益；但如缺少位於測驗訊息較少的考生，則連結是不必要的。

另外，採用 IRT 作為等化之依據時，還需假定所要等化或連結的測驗是單一向度的。

本研究同一年級甲、乙兩式測驗之等化屬於平行等化，五、六年級之等化屬於垂直等化。由於本研究的學術性向測驗包括語文、數量及圖形三個分測驗，此三個分測驗之內容性質不同，因此採用的共同試題題數較多，預試時，兩式測驗有 24 題共同試題（每一分測驗 8 題），正式題本有 20 題共同試題，其中語文有 6 題，數量及圖形各有 7 題。在各種題目參數連結方法中，本研究採用「平均數、標準法」及「同時估計法」（或稱聯合估計法）。同一年級「高雄市與台南市」及「台灣省與台北市」四區同一測驗題目參數之轉換採用「平均數、標準差法」，同一年級男、女生題目參數之相互轉換（題目偏失研究時）也採用「平均數、標準差法」；

五、六年級題目參數之相互轉換採同時估計法，所得題目參數量尺屬於高年級，而非五年級或六年級。

(三)題目參數連結方法

根據原始分數的等化無法滿足公平、對稱和不受樣本變動影響等條件。如果受試反應的資料適合某種題目反應模式，那根據題目反應理論的等化可以解決原始分數等化的那些問題。

依據題目反應理論，受試能力 θ 不受施測的試題子集之影響，而題目參數則不受受試群體之影響。當題目參數是已知的，由於 θ 的估計值是一致的，因此， θ 的估計值不會受到不同的試題子集影響。所以，不管受試接受較難或較易的測驗測試，所得的結果並沒有什麼不同。就題目反應理論來說，不管是水平或垂直等化的情境，只要將試題參數轉換至同一量尺即可，並不需要進行測驗分數的等化。但是，對不同受試群體施測，所得的題目參數表面上看來是不同的。所以會有這種差異是因為我們任意固定 θ (或 b) 的零點。不過從兩組受試所得的題目參數和能力參數存在著一種直線關係。此直線方程式的常數可用下列方式決定：

- a. 迴歸法 (regression method) ；
- b. 平均數、標準差法 (mean and sigma procedure) ；
- c. 強韌平均數、標準差法 (robust mean and sigma procedure) ；
- d. 重複強韌平均數、標準差法 (iterative robust mean and sigma procedure) ；
- e. 特徵曲線法 (characteristic curve method) ；

本研究僅使用「平均數、標準差法」及同時估計法，故僅就此二法略作介紹，其餘請參閱吳裕益（測驗的等化，民 80）。

1. 平均數、標準差法

此種方法是採用估計的題目難度值分配中的二級動差，即找到斜率 B 及截距 A 將測驗二的難度參數估計值的平均數和標準差轉化為與測驗一的難度參數估計值的平均數和標準差。

$$\text{斜率 } B = \frac{S_{b_1}}{S_{b_2}}$$

$$\text{截距 } A = \bar{b}_1 - B \bar{b}_2$$

$$b_2^* = B b_2 + A$$

其中 S_{b_1} ， S_{b_2} 分別為測驗一及二之難度(b)的標準差。

\bar{b}_1 ， \bar{b}_2 分別為測驗一、二難度(b)的平均數。

b_2^* 為轉化到測驗一量尺後的新難度值。

b_2 為測驗二量尺上的難度值。

此種線性轉化方法優於迴歸法，因為它是對稱的，而其邏輯在於對共同試題組而言，二份測驗的估計值應是完全相關的，其差異來源只在於原點和單位不同 (Hung, 1990b)。

2. 同時估計法

本研究所採用同時估計法是以 MicroCAT 測驗系統中的 ASCAL 程式同時估計兩式測驗所有題目的三參數。其使用的步驟與一般用法無異，只是在資料檔上的編排有些不一樣。由於甲、乙式有 20 題共同試題，因此進行同時估計法時，測驗的題數就以 100 題來估計，其資料檔格式如圖 2.2 所示。

圖2.2 同時估計法資料格式

控 制 列			
正 確 答 案			
選 項 數			
測驗所包含的題目之代碼			
受試者之代碼	甲(乙)非共同試題受試者之資料	甲(乙)共同試題受試者之資料	未答完之代碼
	未答完之代碼	乙(甲)共同試題受試者之資料	乙(甲)非共同試題受試者之資料