

## 第二章 文獻探討

本章之目的在探討與本研究有關之理論基礎與相關研究，全章分為四節，期望透過有系統的文獻整理，呈現本研究的方向。第一節為古典測驗理論與其相關研究；第二節為李克特量表法的理論與其相關研究；第三節為試題項目反應理論與其相關研究；第四節為壓力的理論與相關研究。

### 第一節 古典測驗理論與相關研究

#### 一、歷史源由

測驗理論是一種解釋測驗資料間實證關係 (empirical relationships) 的有系統的理論學說。迄今劃分成二大學派：一為古典測驗理論 (classical test theory, 簡稱 CTT)，又稱傳統測驗理論 (conventional test theory)、古典真分數理論 (classical true score theory)，是以真實分數模式 (true score model) 為骨幹；另一為當代測驗理論 (modern test theory)，以試題項目反應理論為架構 (周文欽、歐滄和、許擇基、盧欽銘、金樹人、范德鑫，民 84)。為瞭解其兩者的內涵及應用，我們有必要從歷史的觀點進行回顧，明瞭測驗理論發展的沿革，以及展望其未來趨勢。

古典測驗理論又稱真分數理論，是以真分數模式  $X=T+E$  為架構，其中  $X$  表示實得分數， $T$  表示真實分數， $E$  表示誤差分數 (郭生玉，民 93)。

比西量表 (Binet-Simon Scale) 是第一個心理測驗，也是心理計量學誕生之始。後經諸多學者的研究與闡述，終於歸納形成古典測驗理論 (周文欽等，民

84)。

## 二、基本假設

真分數模式雖是一種簡單線性模式，在應用時仍必須滿足一些基本假設，這些基本假設便是真分數理論所賴以建立的基礎。依據弱勢假設（weak assumption），個人在測驗上的實得分數（observed score）是由兩部份所組成，即真實分數（true score）和誤差分數（error score）（葛樹人，民 77）。實得分數是研究對象從測驗中實際得到的分數；真實分數無法正確被測量，理論上係指研究對象接受無數次測驗的得分平均數或期望值；誤差分數是指測量誤差，即研究對象實得分數與真實分數的差，它可能是正的、負的或零（楊明宗，民 90）。

其主要假設如下：

1. 實得分數的期望值等於真實分數。

$$\blacksquare E(X) = E(T) + E(E) = E(T + E) = E(T) = T$$

2. X、T、E 互為獨立，無相關存在。

$$\blacksquare r_{XT} = 0; \quad r_{TE} = 0; \quad r_{EX} = 0$$

根據上述假設，分別介紹難度、鑑別度、信度、效度和等古典測驗理論於試題分析時的重要指標（林清山，民 81；周文欽等，民 84；郭生玉，民 93；葛樹人，民 77）。

### （一）難度（difficulty）

指試題是困難或簡單的一種指數，古典測驗理論中的難度（item difficulty），是該題的答對率，也就是所有研究對象中答對該題的百分比，通常以 P 代表試題難易度。

$$\blacksquare P_i = n_i / N$$

其中  $P_i$  是答對題項  $i$  的百分比， $n_i$  是答對該題的人數， $N$  是所有受測人數。因此  $P_i$  值愈大表示愈多人答對該題，即該題愈簡單。答對率深受當次受測者的平均能力之影響：如果該次研究對象的平均能力相對於題項  $i$  普遍較高，則  $P_i$  會較高，顯得容易，反之則顯得較低。由於答對率只考慮受測的總人數及答對該題的人數，因此無論如何劃分各種分項測驗都不會影響它的值。

若  $P_H$  表示高分組答對該題的百分比(全體研究對象當中分數最高的 27%)， $P_L$  表示低分組答對該題的百分比(全體研究對象當中分數最低的 27%)。難度介於 0 1 之間，數值愈小表示試題愈困難。

■  $P = (P_H + P_L) / 2$

答對百分比法所建立的難易度指標，由於單位不相等，只能用來表示試題難易的相對位置，無法指出各難易度間差異大小的數學涵義，為了彌補這項缺失，我們可以假設每個試題所要測量的潛在特質或能力是呈常態分配的，然後將試題的難易度指數，按照常態分配所分佈的機率大小，轉換成具有相等單位的等距量尺指標。將  $P$  值轉換成  $Z$  值，然後再將  $Z$  值轉換成  $T$  值 ( $T = 13 + 4Z$ )，已有學者將  $P$  值、 $Z$  值和  $T$  值之間的轉換製作成表，稱「范式試題分析表」，使用者可以直接查表得到(林育柔，民 90)。

## (二) 鑑別度 (discrimination)

指試題是否能對不同能力的學生反應出其答題的差異，希望得分高的研究對象答對某題的機率比得分低者的機率高，才能顯得該題確能鑑別出研究對象的能力差異。以  $D$  表示鑑別度，其值介於 -1 到 1 之間，愈靠近 1 表示個別試題反應與測驗總分之間的一致性愈高。

■  $D = P_H - P_L$

下列三種鑑別力指數常被用來分析測驗內部的一致性( internal consistency ):

1. 相關 (phi correlation) : 此法適用於試題與效標均是二分變項的情況。

	答對	答錯	
及 格	(A)	(B)	(A + B)
不 及 格	(C)	(D)	(C + D)
	(A + C)	(B + D)	

$$= \frac{BC - AD}{\sqrt{(A + B)(C + D)(A + C)(B + D)}}$$

2. 點二系列相關 (point-biserial correlation) : 即每題答對情形 (1 or 0) 和總分相關的情形 ; 這種方法適用於試題為二分變項 , 而效標是屬於連續變項。

$$r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \sqrt{pq}$$

$r_{pb}$  : 點二系列相關係數。

$\bar{X}_p$  : 為答對的學生在效標上的平均得分。

$\bar{X}_q$  : 為答錯的學生在效標上的平均得分。

p : 為答對人數百分比。

q : 為答錯人數百分比。

$S_t$  : 為全部學生在效標上得分的標準差。

3. 二系列相關 (biserial correlation) : 用於多點計分之題目 , 求每題之得分和總分相關的情形 ; 此法適用於假設學生在試題上的反應表現呈常態分配 , 但為了某些理由 , 強以人為方式將其分為答對與答錯兩種情形。

$$r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \left( \frac{pq}{y} \right)$$

$r_{pb}$ ：二系列相關係數。

$\bar{X}_p$ ：為答對的學生在效標上的平均得分。

$\bar{X}_q$ ：為答錯的學生在效標上的平均得分。

p：為答對人數百分比。

q：為答錯人數百分比。

$S_t$ ：為全部學生在效標上得分的標準差。

y：常態分配下答對人數百分比所在位置曲線高度。

### （三）信度（reliability）

一份測驗重複對研究對象施測，所得到的結果是否一致。信度的主要計算方式算法有三種：重測信度、複本信度和內部一致性信度。

從測量誤差來看，信度就是指測驗的分數反映出真實量數（true measure）的程度，即指測驗沒有誤差的程度。即測驗分數中有多少比例是由於研究對象的能力造成，有多少比例是由測驗的誤差造成。因此，測驗的誤差愈小，則信度愈大。

### （四）效度（validity）

指測驗分數的正確性，也就是測驗是否能夠真正測出研究對象是否具備某種特質的程度。效度主要的的計算方式有內容效度、效標關聯效度以及構念效度。

## 三、優點與限制

由於 CTT 具有淺顯易懂、簡單明瞭、計算方便等優點，加上發展歷史長久，適用於大多數的教育與心理測驗資料，以及社會科學資料的分析，為目前測驗學界使用與流通最廣的理論依據。但也因其模式過於簡化，產生了以下幾點缺失與限制（Guion & Ironson, 1983；Wright, 1977；王寶壟，民 84；劉湘川，民 84）：

(一) CTT 在試題分析時所採用的指標，如難度、鑑別度和信度等，都是一種樣本依賴 (sample dependent) 的指標，這些指標的取得，會因為研究對象樣本的不同而有所差異，因此，針對不同潛在特質的研究對象，同一份測驗很難獲得一致的難度、鑑別度、或信度等指標。

(二) CTT 以一個共同的測量標準誤 (standard error of measurement) 作為每位研究對象的測量誤差，忽略研究對象反應的個別差異，對於具有高、低兩極端潛在特質的研究對象而言，這種指標較為不合理且不精確。

(三) CTT 對於非複本 (nonparallel) 但功能相同的測驗所測得的分數間，無法提供有意義的比較，只對複本 (parallel) 且功能相同的測驗所獲得之研究對象分數，才能提供有意義的比較。

(四) CTT 對信度的假設，是建立在複本 (parallel forms) 測量的概念假設，但是這種假設往往不存在於實際的測驗情境。因為，在實際的測驗情境下，研究者不可能要求研究對象接受同一份測驗無數次，仍然假設每次測量間都彼此獨立不相關，況且，每一種測驗並不一定同時都有製作複本。

(五) CTT 忽略研究對象的試題反應組型 (item response pattern)，認為原始得分相同的研究對象，其能力必定一樣。其實不然，因為總分相同的研究對象或總和相同的試題，其試題反應組型不見得會完全一致，因此，試題反應組型所顯示的意義也不會相同，所估算出的潛在能力和試題參數估計值，也應該不同。

(六) 無法確知不同研究對象答對某一試題之機率，因此在建立題庫 (item bank) 試題等化 (item equating) 偏誤 (bias) 上有所困難。

因此古典測驗理論與傳統項目分析其有限制如下：要有一個具代表性的樣本；難度與鑑別度是全部人通過的平均率，無法看出個別受測者；僅有一個單一的信度資料；難度是對於相似測驗的比較，會有樣本組成及試題難易的問題。

## 第二節 李克特量表法的理論與相關研究

### 一、歷史源由

人格或態度的測量，因為研究對象可以隱藏內心的真正態度，因此一直是許多心理學家想解決的問題。塞斯通（Thurstone）和柴夫（Chave）於 1929 年提出以等距量表法（equal-appearing interval scaling）測量態度，然而使用這套方法必須符合一些統計的先決條件。例如：敘述句的得分必須獨立於回答者的態度且為常態分配等。但是，若要符合這個統計條件，就會使得整個編製計分的方式變得極為複雜（郭生玉，民 93）。

為了讓整個問卷編製及計分方式更為簡潔，李克特提出了簡便的多點量表，就是大家所熟知的李克特總加量表。此量表自從 1932 年李克特發表以來，因為有著尺度建構方便明確、編製簡單、所耗時間較少的優點而被廣泛使用至今，許多測量人格或態度的問卷，大多採用此方式（葛樹人，民 77）。

### 二、基本假設

李克特認為一個人對於事情的反應，應在某種範圍內變動，並非固定於一個值，而量表是用來測量研究對象對事情反應的強度，而非像塞斯通的量表中詢問研究對象的絕對反應（顏乾明，民 92）。

李克特量表法主要基於以下三個概念（郭生玉，民 93；張文豪，民 91）：

（一）態度是對某個外顯行為的傾向。

（二）這個外顯行為可用語言替代表達。

（三）每一個題目所測量的態度具有同等的數值，而研究對象可以對每一個題目表示不同程度的態度。

(四) 假設所選擇的項目代表研究者所預測量的變項。

(五) 假設每一個反應類目 ( response categories ) 之間等距。

而李克特量表法的結構包括兩部分：第一部份是對測量主題所編擬的積極和消極的態度敘述句，用以表達研究對象對每一個句子陳述內容的贊同程度。第二部分則是用來表示看法的評分表，贊同程度的量尺由兩點至七點不等，由研究者自行決定，通常是以預計受測母群的特性來考慮。譬如最常用的是設計成五個量尺，分成「非常同意」、「同意」、「稍微同意」、「不同意」、「非常不同意」，並依序給予不同的分數 ( 葛樹人，民 77 ; Horne , 1985 )。由於李克特量表法有了等距性，因此可以把原先屬於順序量尺的數據，轉換成等距量尺的數據，進一步可以將所得數據做運算。

### 三、優點與限制

根據郭生玉 ( 民 93 ) 指出，李克特量表法的優點如下：

(一) 編製容易。

(二) 量表的同質性較高。

(三) 信度較塞斯通式量表高。

(四) 研究對象可以表達其情感的強度。

(五) 有較大的變異性。

李克特量表法一方面希望將模糊認知直接以數值加以表示，以明確的度量代替模糊的語意性措辭 ( linguistic terms )，但是此種轉換是否能合理地描述人類語意表達？

另一方面，它要求研究對象採用二元邏輯的思考方式，對所討論的議題給予

單一的選擇，以態度量尺數為五點時為例，若假設一填答者在某試題的潛在特質為 2.74，則此人可能選填的選項為 2 或 3，計分值則分別為 2 或 3，當填答者填寫完該問卷時，將各題分數加總得一總分，即可進行分析，這種「非此即彼」的二分法，進而造成估計或分析上的誤差（顏乾明，民 92）。

最後，它假設區間尺度等距，如表 1 所示，研究對象語意的差異性與模糊性，有意或無意的被忽略掉，因此存在著語意膨脹或貶值現象，可能會導致結果的偏差（王舜傑，民 90）。

表 1：李克特量表法在不同態度量尺點數的設計

量尺點數 \ 語意	非常同意	非常不同意
四點		
五點		
六點		
七點		

#### 四、相關研究

以下就李克特量表法的相關研究及各學者的看法作一概述。

徐秀珍（民 77）比較態度量尺為四點與六點的不同量表，發現兩者的信度並無差異。而態度量表题目的設計，最重要也最根本的原則是題目選擇與題意陳述的明確與具體化。此外，在編製態度問題的題目上，尤其是針對政治性對社會道德且隱私議題的內容，增加「無意見」選項的安排是必要的。

彭仁信（民 83）以高中生與大學生在疏離量表上的反應，探究探討李克特式量表中選項形式之問題。研究結果顯示，選答的分布隨量表點數的不同而改變，兩點的量表選答的分布，易集中在兩端；而當量尺點數為五點時，較能明確反應出填答者的態度強弱。而一般問卷不同反應類目上的選項呈現非等距的關

係。

吳毓瑩（民 85）指出，量尺的偶數點或奇數點，並沒有一致的好與壞。在壓力較大的態度意見測量上，四點量表較少被方法變異量所干擾，表現優於五點量表；相反地，在情緒的測量上，由於五點量表的中間項提供一個中等量的程度指標，表現優於四點量表。

翁儷禎（民 89、民 90、民 91）的研究結果顯示評定量尺的標示語點數可能影響研究對象的反應，研究者建構評定量表時，應考慮研究對象在該量表可能的反應異質程度。而針對量尺點數為 2 到 14 點做信效度的模擬研究，其結果顯示，量尺點數主要影響真分數與觀察分數相關平方的估計值，雖然量尺點數對其兩種信度估計值的影響並不大，但可發現當量尺點數超過五點後信度的增加量變小，七點之後增加量更是微小；在檢視信度的平均值與標準誤後發現，偶數點量尺採用六點，或奇數點量尺採用七點的結果應會較穩定。

王舜傑（民 90）以電腦模擬的方法探討模糊語意變數計分和傳統計分之李克特氏量表在問卷的信度以及對研究對象內心潛在特質的估算精確度。研究結果發現模糊語意變數計分問卷資料在信度及估算潛在特質精確度兩點上皆優於傳統計分，且模糊語意變數計分較不受樣本數及問卷題數等變因影響。

王嘉寧（民 90）的研究在探討 2、3、5、7 點的量尺的反應點數與题目的分配特性對因素分析結果的影響。在因素負荷量的估計上，分別比較主成份法，主因子法，疊代主因子法，與最大概似法所估計負荷量的精確性。研究結果顯示量尺反應點數與變項分配型態對於因素分析結果均有重要的影響。態度量尺點數與變項分配型態對於因素分析結果均有重要的影響。當態度量尺點數越少，偏態程度越大。

張文豪（民 91）的研究指出，李克特式量表的點數對於分群效果指標的運算有顯著的影響，較為適合的是四點或五點的李克特量表法。而且正向反應類目

應該在前，中間反應類目居中。

顏乾明（民 92）採三角模糊數的模糊語意變數，針對不同量表的態度量尺點數、試題數、填答者人數與填答者潛在特質分布等變因，分別比較模糊語意量表計分與傳統李克特式量表計分，在各因子組合下的信度差異。研究發現：以模糊語意計分所得之信度值，皆較以傳統李克特式量表計分所得之信度值要高。而在模糊語意計分模式下，影響信度的主因是試題數的增加，而非填答者人數。

Komorita & Graham (1965) 指出若題目的同質性高，則信度與態度量尺反應類目的多寡無關；但若題目的同質性相對較低的話，則增加態度量尺反應類目可以提高量表的信度。

Berdie (1986) 對於態度量尺反應類目多寡的問題，提出以下看法：在大多數的情況下，態度量尺點數為 5 點的量表是最可靠的，超過 5 點的話，一般人很難有足夠的辨別力去進行填答，5 點量表則正好表示了溫和意見與強烈意見之間的區別性。

Anderson (1990) 的研究指出，增加態度量尺反應類目就是讓填答者有更多的選填機會，以增加量表的內部一致性，而在李克特量表法中增加態度量尺反應類目，與在二分法 (dichotomous measures) 的認知測驗上增加題目意義是相似的。

綜合以上的研究論述可知，較為合適的是四點或五點的李克特量表法，量尺是要選擇偶數點或奇數點，並沒有一致的好與壞。但超過五點的量表，一般人很難有足夠的辨別力去進行填答；而適當數目的反應類目可以讓研究對象有更多的選擇機會，以增加量表之內部一致性。故以李克特量表法設計題目時，最重要的原則是題目選擇與題意陳述的明確與具體化。

### 第三節 試題項目反應理論與相關研究

#### 一、歷史源由

為了克服古典測驗理論的缺失，當代測驗理論因此誕生。自從 Lord (1980) 發表第一本以「試題反應理論」為名的專書後，當代測驗理論正式以試題項目反應理論為其中心架構，故測驗學者逐漸以試題項目反應理論為當代測驗理論的代表。在此之前，試題項目反應理論有一別稱：「潛在特質理論」由於潛在特質理論一詞還包括「因素分析」( factor analysis )、「多元度量法」( multidimensional scaling ) 與「潛在結構分析」( latent structure analysis ) 等，涵蓋面甚廣，無法精確反應出研究對象在試題上的反應狀況。當代測驗理論的內涵，主要是以試題項目反應理論為理論架構，依據強勢假設 ( strong assumptions ) 而來，其理論的發展為時稍晚，理論模式也不斷的在發展當中，所採用的計算公式複雜深奧、艱澀難懂，為一立論與假設均合理與嚴謹的學說，所適用的測驗資料種類雖屬有限，但深受測驗學者的青睞，已有逐漸凌駕古典測驗理論之上，甚至進而取而代之勢 ( 周文欽，民 86；余民寧，民 86 )。

試題項目反應理論雖然自 1980 年才正名成立，然而在 30 和 40 年代，試題項目反應理論便已有初步的理論架構。茲將對試題項目反應理論發展有實際貢獻的代表性作者及著作，條列簡述於下表，由表 2 的內容便可獲知試題項目反應理論的發展概況。

表 2：試題項目反應理論代表性作者及著作

作者 ( 年代 )	代表作及其貢獻
Tucker ( 1946 )	第一位提出試題特徵曲線概念的人。
Lord ( 1952 )	第一位導出兩個參數常態肩形模式的參數估計公式，並考慮試題項目反應理論應用性的人。

Rasch ( 1960 )	試題項目反應理論中 Rasch 模式創始者，影響深遠。
Lord & Novick ( 1968 )	第一本介紹古典與當代測驗理論模式的經典作品，引發學者對「潛在特質」概念的重視與研究。
Wright & Panchapakesan ( 1969 )	美國地區第一篇介紹 Rasch 模式的參數估計法，並發展有名的 BICAL 電腦程式的代表作品。她的一系列作品描述新的試題反應模式及其應用，其中包含處理多分法與連續性資料的模式，甚至擴展到多向度的試題反應模式，為一艱澀難懂的重要著作。
Samejima ( 1969 )	
Bock ( 1972 )	提供許多估計模式參數的新概念。
Andersen ( 1973 )	歐洲地區談論測驗模式的重要著作。
Lord 等 ( 1976 )	創作第一版有名的電腦程式：LOGIST。
Journal of Educational Measurement ( 1977 )	第四季出版一冊專門探討試題項目反應理論的專輯
Baker ( 1977 )	第一篇評論試題反應模式參數估計法的文獻探討。
Wright & Stone ( 1979 )	第一本描述各種 Rasch 模式理論及其應用的專書。
Lord ( 1980 )	第一本以試題項目反應理論命名的專書，是當代測驗理論發展的里程碑。
Weiss ( 1980 )	第一本編輯成的論文輯，專談試題項目反應理論的實際應用課題-電腦化適性測驗。
Andersen ( 1980 )	對測量模式參數估計法有貢獻的方法學專論。
Bock & Aitkin ( 1981 )	提出邊緣的最大近似值估計法-EM 估計程序，對參數估計法的改進貢獻不少。
Masters ( 1982 )	第一位發表部份知識計分模式，對改進 Likert 式評定量表的計分與次序反應資料的計分貢獻不小。
Wright & Masters ( 1982 )	闡述 Rasch 模式的各種模式成員，證明皆與部份計分模式相通，對 Likert 式評定量表與次序反應資料的計分方式改進不少。
Mislevy & Bock ( 1982 )	發表另一有名的電腦程式：BILOG。
Applied Psychological Measurement ( 1982 )	第四季出版一冊專門探討試題項目反應理論及其應用的進階專輯。
Wainer & Messick ( 1983 )	編輯論文集以表揚 Lord 一生對試題項目反應理論的貢獻，並兼論該理論的應用與未來。
Weiss ( 1983 )	專談試題項目反應理論的應用與未來，並介紹它在電腦化適性測驗上的應用。
Hambleton ( 1983 )	編輯而成的論文集，專談試題項目反應理論的模式與

	應用。
Hulin, Drasgow, & Parsons ( 1983 )	為一本試題項目反應理論的教科書，增加對「適合度測量」概念的說明與應用。
Embretson ( 1985 )	編輯論文集，專談試題項目反應理論的未來發展。
Baker ( 1985 )	一本導論性的試題項目反應理論教科書，專為沒有數學訓練基礎的讀者而作，並附有 CAI 的電腦教學磁片。
Hambleton & Swaminathan ( 1985 )	一本進階的試題項目反應理論教科書。
Crocker & Algina ( 1986 )	談論與比較古典與當代測驗理論的導論性教科書。
Wainer & Braun ( 1988 )	專談有關效度方面的論文集，也談試題項目反應理論在效度上的應用。
Linn ( 1989 )	負責主編第三版的「教育測量」( Educational Measurement ), 其中增加一章專門介紹並評論試題項目反應理論。
Freedle ( 1990 )	專談人工智慧及其在當代測驗理論上應用之論文集
Suen ( 1990 )	介紹各種測驗理論方面的教科書。
Wainer 等 ( 1990 )	專談電腦化適性測驗方面的入門書，也談試題項目反應理論在電腦化適性測驗上的應用。
Hambleton, Swaminathan, & Rogers ( 1991 )	試題項目反應理論方面的入門書，適用於非數學主修的初學者閱讀。
許擇基、劉長萱 ( 民 81 )	出版「試題作答理論介紹」一書。
Laveault, D., Zumbo, B. D., Gessaroli, M. E., & Boss, M, W. ( 1994 )	Modern theories of measurement : problems and issues. Ottawa, Canada : Edumetrics Research Group, University of Ottawa.
王寶壟 ( 民 84 )	出版「現代測驗理論」一書。
Master, G. N., & Keeves, J. P. ( 1999 )	Advances in measurement in educational research and assessment. New York : Pergamon.
McDonald, R. P. ( 1999 )	Test theory. Mahwah, New Jersey : LEA.
Embretson, S. E., & Reise, S. P. ( 2000 )	Item response theory for psychologists. Mahwah, NJ : LEA.
Susan E. Embretson & Steven Paul Reise ( 2000 )	Item Response Theory for Psychologists. New Jersey : Lawrence Erlbaum Associate.
Wim J. van der Linden &	Handbook of Modern Item Response Theory. New

---

Ronald K. Hambleton ( 2000 )	York : Springer.
Thissen, D., & Wainer, H. ( 2001 )	Test scoring. Mahwah, NJ : LEA.
Frank B. Baker, Seock-Ho Kim ( 2004 )	Item Response Theory : Parameter Estimation Techniques ( Statistics, a Series of Textbooks and Monographs ) . Marcel Dekker.

---

( 修改自余民寧，民 80 )

## 二、基本概念與假設

試題項目反應理論的主要假定是：受測者答對某一題的機會，決定於其（被測量的）能力及該題項的特質（如測量的內容，型式等）。IRT 建立在兩個基本概念上（許擇基、劉長萱，民 81；張育璋，民 91）：

（一）研究對象在某一測驗試題上的表現情形，由一組因素來加以預測或解釋，這組因素叫作潛在特質（latent Traits）或能力（abilities），常以  $\theta$  來表示。

（二）研究對象的表現情形與這組潛在特質間的關係，可透過一條連續性遞增的函數來加以詮釋，這個函數便叫作試題特徵曲線（item characteristic curve，簡稱 ICC）。具某能力  $\theta$  之試題特徵曲線所表示的涵義，即是某種潛在特質（能力）的程度與其在某一試題上正確反應的機率二者之間的關係；這種潛在答對測驗試題的機率為  $P(\theta)$ ，能力愈高，即  $\theta$  愈大，答對的機率愈高，反之，則愈低。試題特徵曲線能清楚扼要地表示試題參數（item parameters）與能力間的關係變化，此即 IRT 的精華。試題項目反應理論最大的特點，在於研究對象答對某一試題的機率，是由其能力和試題的特性所共同決定的。

IRT 具有下列基本假設（Hambleton & Swaminathan, 1985；王寶壙，民 84；余民寧，民 81a）：

### (一) 單向性 (unidimensionality)

單向度是指測驗中的各試題只測驗一種共同的特質或能力，這種影響測驗結果僅有單一能力或潛在特質的假設，必須包含在測驗中，便是單向度的假設。

Hambleton & Swaminathan (1985) 認為只要測驗資料有一個「主控」因素或成分，就可算符合單向性假設，而這個主控因素便是特質或能力。

Crocker & Algina (1986) 認為單向性就是試題間統計依賴，也就是對整體受試而言，試題間互相關聯時才有單一特質的存在。如果測驗是單向性時對整體研究對象而言，試題間為相互關聯，故應只有一條迴歸線，以表示只有一個特質存在。

單向性假設檢定的方法通常為因素分析法，透過因素分析中的主成份分析，可尋找到一份測驗中影響研究對象得分的因素有那些，而最具代表性的因素為何。在實際的測驗情境裡，研究對象在測驗上的表現情形很少是純粹受到一種因素的影響，其他因素如：成就動機、考試技巧、學習態度、及人格特質等，也都都會影響到測驗的結果，因此單維度只是個假設，至於影響測驗表現的其他因素，則歸為測量誤差。其實，試題項目反應理論亦能處理多項度(multidimensionality)的問題，其模式較為複雜難懂 (鄭富森，民 83)。

### (二) 局部獨立性 (local independency)

當影響測驗表現僅考慮研究對象的能力因素後，研究對象在不同試題上的反應，在統計學上而言是獨立的，這意味著涵蓋在 IRT 模式的能力因素，才是唯一影響受試者在測驗試題上反應的因素。即一個試題不能為另一個試題提供線索，此意謂著涵蓋在試題反應模式裡的能力因素，才是唯一影響受試者在測驗試題上作答的因素。

局部獨立性假設的檢定通常與前述單向性假設的檢定相同，都利用因素分析

法來檢定，通常當單向度假設獲得成立時，局部獨立性假設也會獲得成立。換句話說，局部獨立性假設是單向性假設成立下的必然結果，其公式為(Lord, 1980):

$$\begin{aligned} P(U_1, U_2, U_3, \dots, U_n) \\ &= P(U_1) P(U_2) P(U_3) \dots P(U_n) \\ &= \prod P(U_i) \end{aligned}$$

$\mu$  : 受試者的能力

$U_i$  : 受試者在第  $i$  題作答情形

$P(U_i)$ : 代表能力值  $\mu$  之受試者在  $U_i$  題的答對機率

### (三) 等鑑別力 (equal discrimination)

通常使用單參數模式的研究者其重點在難易度參數，故而有等鑑別度的假設，因此使用單參數模式時各個試題鑑別度需大致相等，或選擇鑑別度大致在一個合理的範圍內的試題。在作等鑑別度假設的檢定時，若估計出來的鑑別度參數值落於 0.8 至 1.25 之間，便算符合等鑑別度的假設。

### (四) 最小猜測 (minimal guessing)

最小猜測是單參數模式及雙參數模式的假設之一，亦即模式的應用不能受測驗試題猜測因素的影響。因此在使用上述兩種模式時要進行誘答選項分析，以試題的誘答選項來讓猜測因素減到最小；或者刪除能力特低的受試者，通常這些受試者較可能產生猜測行為。

### (五) 非速度性 (nonspeedness)

測驗的實施不是在速度限制下完成的，研究對象在答題時沒有時間限制，其成績不理想是因為能力不足而不會答，不是由於時間不夠答完所致。因為如果受

到時間限制，答題的好壞普遍受到答題速度的影響，測驗成績會受到能力以外的因素所影響，則單向性的假設就會受到質疑。

### 三、試題項目反應理論的參數

試題項目反應理論就是以數學方式表示研究對象的能力與試題難度、鑑別力及猜測等參數間的關係，而試題項目反應理論的試題參數有難度參數、鑑別力參數及猜測參數。

#### (一) 試題特徵曲線

又稱為試題反應函數 (item response function, 簡稱 IRF), 而試題特徵曲線就是「試題答對機率」對「受試者能力」的迴歸線，該曲線是表示受試者試題表現情形與潛在特質的關係 (劉湘川, 民 83)。

根據 Allen & Yen (1979) 的解釋， $\theta$  為受試者之能力值， $P(\theta)$  為項目反應函數 (item response function)，表示具某能力  $\theta$  的受試者答對某題的機率  $P$ 。而將  $P(\theta)$  用平面座標圖表示出來，所產生之曲線為試題特徵曲線，是機率函數所產生之曲線，是以圖形的方式來表達。在某一試題中，研究對象在不同的成績分組裡，相對的答題通過率。用比較數學的說法，亦即 ICC 顯示某一試題的答對機率為該測量特質的函數，而圖形的位置、斜率及截距，分別代表著各個題目的難易度、鑑別度及猜測度。因此 ICC 可分為單參數 (難度) 雙參數 (難度、鑑別度) 與三參數 (難、鑑別度、猜測度) 三種 (何世欽, 民 89; 張育璋, 民 91)。

ICC 需符合下列重要基本假設：

1. 試題需為良好試題，即能力高者，答對率會高。

$$\theta_1 < \theta_2; P(\theta_1) < P(\theta_2)$$

2.  $0 < P(\theta) < 1$ ，即答對率在 0 和 1 之間。

3. 值可達 - 到 + ，並符合常態分配，平均數為 0，標準差為 1。

根據劉湘川等（民 87）指出，試題特徵曲線組合是指測驗資料中每條特徵曲線形狀的組合情況，多以鑑別力參數  $a$ （discrimination parameters）、難度參數  $b$ （difficulty parameters）、猜測參數  $c$ （guessing parameters）等來表示試題特徵曲線的組合。圖 1 是七個不同試題特徵曲線的例子。

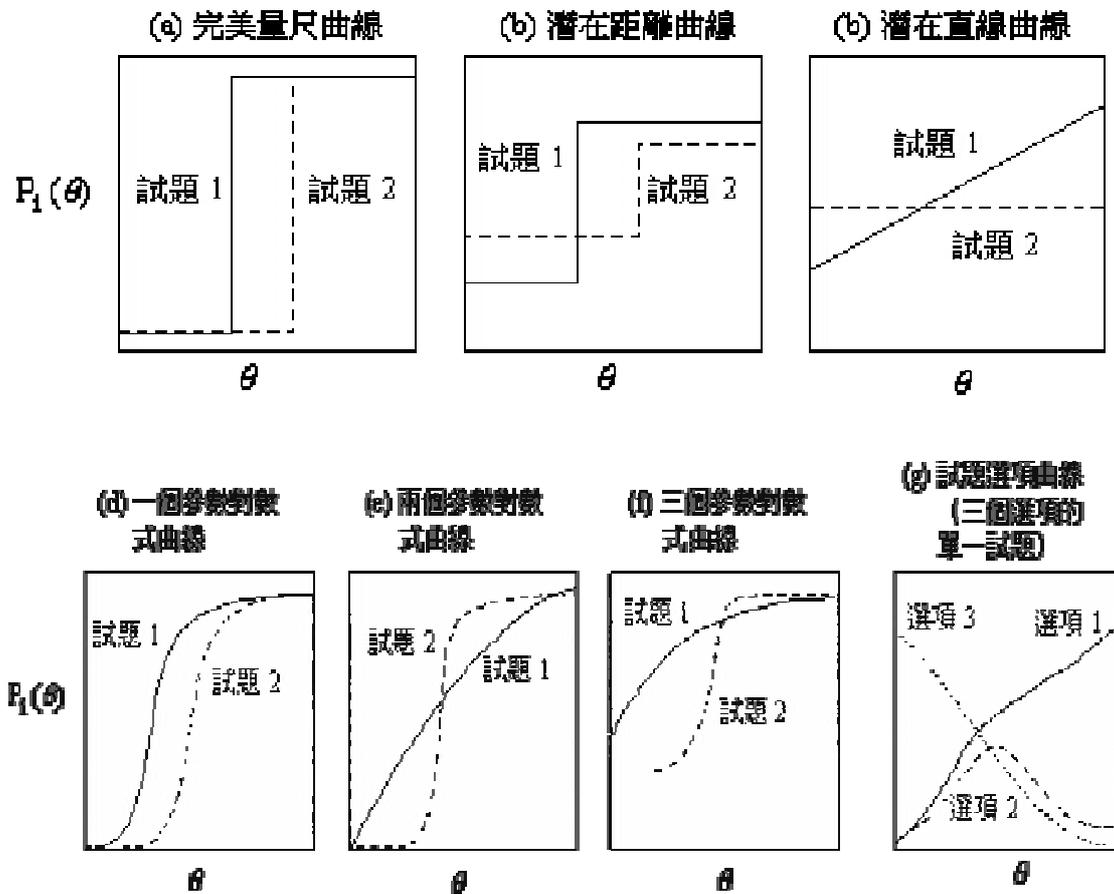


圖 1：七個不同試題特徵曲線的例子  
（取自余民寧，民 81a）

## (二) 難度參數

難度參數又稱為 b 參數，題目愈難，所需能力愈高，ICC 愈偏右，b 參數和能力量尺一樣，介於 - 到 + 之間，為方便計算，將其轉換成平均數 0，標準差為 1 的量尺，一般 b 參數大都介於 - 2 與 + 2 之間（吳孟勳，民 91）。

### （三）鑑別力參數

鑑別力參數又稱為 a 參數，是能力變化時，答對機率的變化程度，a 參數表示 ICC 陡峭的程度，愈陡，鑑別力愈高；愈平，鑑別力愈低。a 介於 0 與 + 之間，一般以 0.2 最多，但以 0.8 a 1.25 最有效（吳孟勳，民 91）。

### （四）猜測參數

猜測參數又稱為 c 參數，指能力極低時仍能答對的機率。c 值愈小題目愈有效，一般以 c 值在 0.3 以下為選題重點，超過 0.3 的試題不是理想的試題。若 m 為選項數目，c 參數值應小於 1/m 為宜（吳孟勳，民 91）。

## 四、參數模式

IRT 可分為參數模式與無參數模式，前者因參數的數目多寡可以分為單參數與多參數模式。前者包含單參數模式（one-parameter logistic model）、單參數常態肩型模式（one-parameter normal ogive model）、完全量尺模式（perfect scale model）等，而後者可分為雙參數模式（two-parameter logistic model）、雙參數常態肩型模式（two-parameter normal ogive model）、三參數常態肩型模式（three-parameter logistic model）、四參數常態肩型模式（four-parameter logistic model）等。

以下簡介幾種 IRT 常見的模式（余民寧，民 81b）：

### （一）單參數模式（簡稱 1PL，又稱 Rasch Model）

$$P_i(\theta_s) = \frac{e^{D(\theta_s - b_i)}}{1 + e^{D(\theta_s - b_i)}}$$

$\theta_s$  : 第s位受試者的能力參數。

$P_i(\theta_s)$ : 表示能力參數為  $\theta_s$  的受試者s, 答對試題i, 或在試題i上正確反應的機率。

D : 表示常數為1.7。

$b_i$  : 試題難易度的參數。

單參數模式係指模式中只有一個參數, 即難易度。所謂難易度, 通常用  $b_i$  來表示, 即是試題困難的程度。當第i題的難易度參數  $b_i$  落在試題特徵曲線上答對機率  $P_i(\theta_s)$  為 .5 的點時, 試題的難易度參數會等於受試者能力值  $\theta_s$ , 換句話說, 當研究對象的能力值和試題難易度相等時, 答對第i題的機率  $P_i(\theta_s)$  是0.5。若能力值小於試題難易度, 受試者答對第i題的機率  $P_i(\theta_s)$  低於0.5。亦即答對第i題的機率  $P_i(\theta_s)$  要高於0.5, 必須要有更高的能力。

理論上, 難易度值介於 - 2 到 + 2 之間, 但實際應用上, 通常只取 - 2 與 + 2 之間的範圍。相對於古典測驗理論的難易度指數, 其所指的是試題真正的「難」度, 而不是古典測驗理論所指的「易」度, 而且古典測驗理論的難易度指數是一種樣本依賴 (sample dependent) 的指標, 其值受到研究對象樣本的影響很大 (楊明宗, 民90)。

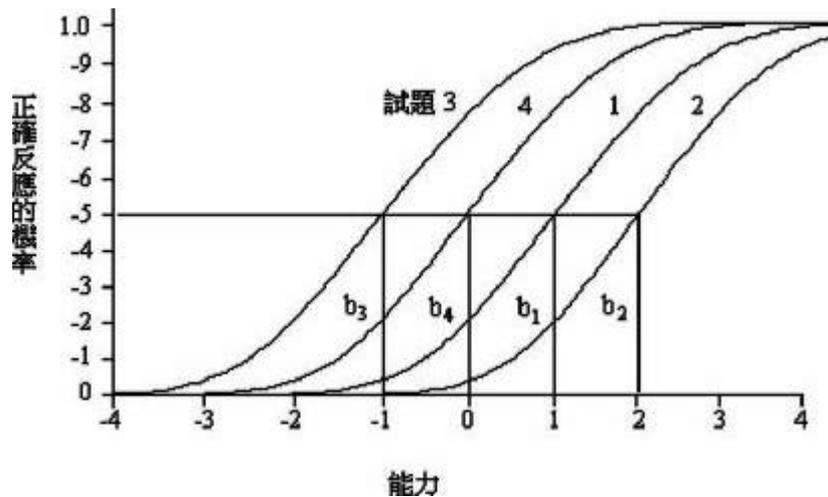


圖 2：四條典型的單參數試題特徵曲線  
(取自余民寧，民81b；林育柔，民90，p38)

## (二) 雙參數模式 (簡稱 2PL)

$$P_i(s) = \frac{e^{D a_i (\theta_s - b_i)}}{1 + e^{D a_i (\theta_s - b_i)}}$$

$\theta_s$ ：第s位受試者的能力參數。

$P_i(s)$ ：表示能力參數為  $\theta_s$  的受試者s，答對試題i，或在試題i上正確反應的機率。

D：表示常數為1.7。

$a_i$ ：試題鑑別度的參數。

$b_i$ ：試題難易度的參數。

雙參數模式係指模式中有兩個參數，即難易度和鑑別度。難易度參數與上述單參數模式中的相同。所謂鑑別度（通常用  $a_i$  來表示）是指試題對不同能力的受試者是否能反應出其答題的差異，也就是說鑑別度大的試題，對於能力高的受試者而言，其答對率高；對能力低的受試者而言，其答對率低。從 ICC 來看，鑑別度是指能力  $\theta_s$  變化時，受試者答對機率  $P_i(s)$  變化的程度（斜率），因此 ICC 愈陡的試題，其鑑別度愈大，ICC 愈平緩的試題，其鑑別度愈小。理論上，

鑑別度參數的值介於- 到 + 之間，但學者們通常捨棄負的鑑別度值，因為該試題反向區別不同能力水準的受試者，從 ICC 看，能力愈高的受試者答對某試題的機率愈低，所以在實際應用上，鑑別度值通常只取 0 到 + 2 之間的範圍。相對於試題的難度，IRT 所指的試題鑑別度與古典測驗理論所指的鑑別度之意義是相同的，不過古典測驗理論的鑑別度就像難度指數一樣，是一種樣本依賴 (sample dependent) 的指標，其值受到研究對象樣本的影響很大 (林育柔，民 90 )。

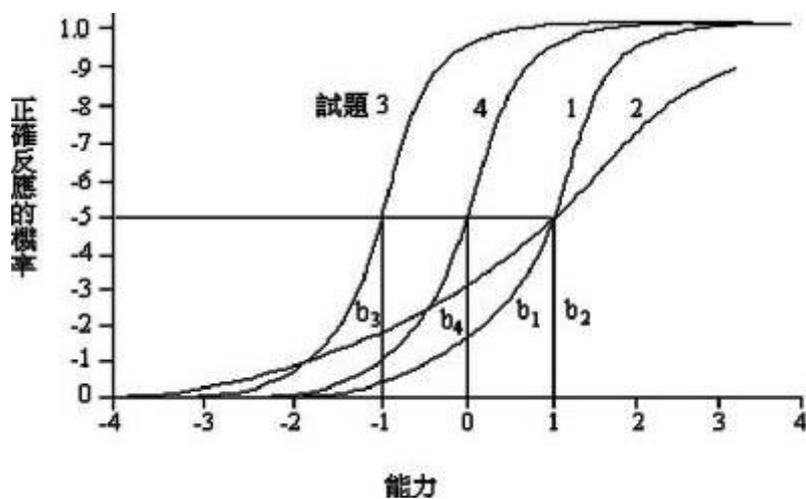


圖 3：四條典型的雙個參數試題特徵曲線  
(取自余民寧，民81b；林育柔，民90，p39)

(三) 三參數模式 (簡稱 3PL),

$$P_i ( s ) = c_i + ( 1 - c_i ) \frac{e^{D a_i ( s - b_i )}}{1 + e^{D a_i ( s - b_i )}}$$

$s$  : 第s位受試者的能力參數。

$P_i ( s )$ : 表示能力參數為  $s$  的受試者s，答對試題i，或在試題i上正確反應的機率。

$D$  : 表示常數為1.7。

$a_i$  : 試題鑑別度的參數。

$b_i$ ：試題難易度的參數。

$c_i$ ：試題猜測度的參數。

三參數模式係指模式中有三個參數，即難度、鑑別度和猜測度。其中難度、鑑別度與單參數或雙參數模式所指的相同。所謂猜測度（通常用  $c_i$  來表示）是指將能力極低（或能力參數值為 0）的研究對象考慮到模式裡，計算出此類受試者答對試題的機率。從 ICC 來看，它是位於該曲線的左下漸近線。 $c_i$  值愈小，表示猜測的因素愈小，試題愈有效，最理想的  $c_i$  值是 0，表示試題完全不受猜測影響，但是只要測驗的型式是選擇題型，通常很難避免研究對象的猜測行為。

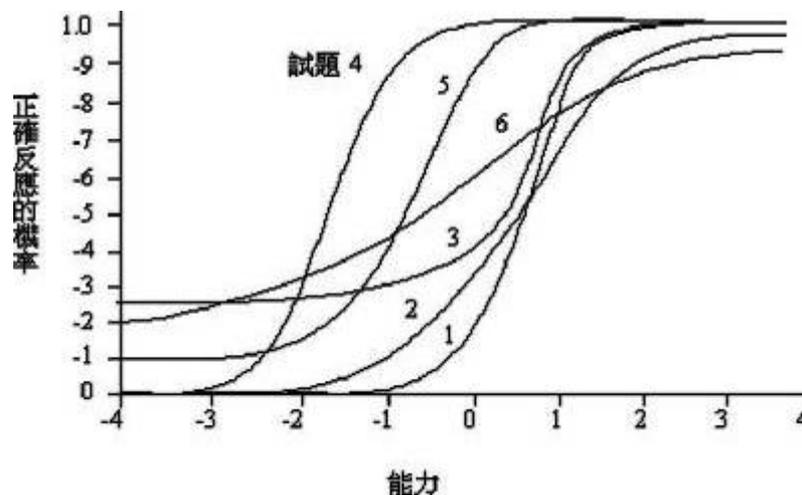


圖 4：六條典型的三個參數試題特徵曲線  
(取自余民寧，民81b；林育柔，民90，p40)

#### (四) 無參數模式 (non-parameter item response model)

無參數試題項目反應理論 (non-parameter item response theory) 是一種將傳統的古典測驗理論結合現代試題項目反應理論所整合出來的一種新的測驗理論。該模式所採用的參數估計法非一般 IRT 所用的最大可能性估計法 (maximum likelihood estimator, 簡稱 MLE); 以最小方差法 (least squares estimator, 簡稱

LSE), 來估計所有未知參數 (何世欽, 民 90)。

鄭富森 (民 83) 所提出的無參數試題項目反應理論, 拋開參數迴歸的模式, 認為研究對象在每個測驗試題上的反應, 主要取決於「受試的特性」、「試題的特性」和「評分的特性」三方面因素。其已撰寫成 PC 軟體名稱為 ICCNP。無參數試題項目反應理論具有參數型 IRT 所沒有的優點如下 (劉湘川, 民 84; 林育柔, 民 90):

1. 無參數 IRT 不需「局部獨立」的假設。該理論採最小平方法來估算, 因此不受限於局部獨立性的假設。

2. 無參數 IRT 其能力參數估計值為一直接解 (closed-form solution), 較其它理論更加精確。

3. 無參數 IRT 的試題特徵曲線不似參數型 IRT 受到 logistic 函數的限制, 可呈現更多樣性的特徵曲線, 使得利用試題特徵曲線來進行概念診斷更加可行 (簡茂發、劉湘川、許天維、郭伯臣, 民 82)。因為一般常用之試題項目反應理論軟體皆有試題參數, 依據參數來繪製試題特徵曲線, 卻限制了該曲線的真實性。透過試題選項特徵曲線 (item option characteristic curve, 簡稱 IOCC) 的方式真實呈現研究對象作答反應情形, 與誘答選項的作答情形。

4. 一般試題項目反應理論的軟體, 其分析資料的樣本需要相當大, 而 ICCNP 並不需要非常大的樣本。ICCNP 亦可針對多元計分的試題進行分析, 不似一般試題項目反應理論軟體只適用於固定某種計分的情境。

5. 在命題技術上可透過試題選項特徵曲線, 找出具有良好鑑別能力的選項, 以保留良好的試題選項, 刪除題意不明之選項, 可控制試題的品質。

## 五、優點與限制

當代測驗理論是為改進古典測驗理論的缺失而來, 它具有下列幾項特點, 而

這些特點正是古典測驗理論所無法具備的 (Hambleton, 1989; Lord, 1980; 余民寧, 民 86):

1.當代測驗理論所採用的試題參數 (item parameters) (如: 難度、鑑別度、猜測度等), 是一種不受樣本影響 (sample-free) 的指標; 也就是說, 這些參數的獲得, 不會因為所選出接受測驗的受試者樣本的不同而不同。

2.當代測驗理論能夠針對每位受試者, 提供個別差異的測量誤差指標, 而非單一相同的測量標準誤, 因此能夠精確推估受試者的能力估計值。

3.當代測驗理論可經由適用的同質性試題組成的分測驗, 測量估計出受試者個人的能力, 不受測驗的影響 (test-free), 並且對於不同受試者間的分數, 亦可進行有意義的比較。

4.當代測驗理論提出以試題訊息量 (item information) 及測驗訊息量 (test information) 的概念, 來作為評定某個試題或整份試卷的測量準確性, 倒有取代古典測驗理論的「信度」, 作為評定試卷內部一致性指標之勢。

5.當代測驗理論同時考慮受試者的反應組型與試題參數等特性, 因此在估計個人能力時, 除了能夠提供一個較精確的估計值外, 對於原始得分相同的受試者, 也往往給予不同的能力估計值。

6.當代測驗理論所採用的適合度考驗值, 可以提供考驗模式與資料間之適合度、受試者的反應是否為非尋常等參考指標。

綜合上述, 當代測驗理論似乎是絕對優於古典測驗理論, 但是事實上, 當代測驗理論被採用於解決真實測驗資料者, 比起古典測驗理論廣泛地被應用的情形而言, 尚屬少數, 微不足道。其主要原因有下列諸點:

1.當代測驗理論係建立在理論假設嚴謹的數理統計學機率模式上, 是一種複雜深奧、艱澀難懂的測驗理論, 這對於在數學方面訓練有限的教育與心理學界學

者而言，無非是一大挑戰。閱讀有關此理論之數學方面的研究報告與專書，已頗感困難，實在更難以深入將之發揚光大。

2.多數當代測驗理論學者都是出身自數學界或曾是數學主修者，或至少在數理統計學上訓練有素者，他們偏愛對理論模式的探討，遠勝於對實際應用的推廣工作。

3.過去，電腦科技的進步有限，沒有電腦套裝軟體程式的即時配合，當代測驗理論中對模式參數的估計，難以用手算或小型計算機順利進行，因此，在應用上更受限制。

4.有些古典測驗理論的擁護者，對當代測驗理論的研究與發展，所能獲致之成效與應用性深表懷疑。為了證明與解釋疑惑，當代測驗理論學派的支持者，便更朝理論模式的量化技術方面探討，致使當代測驗理論的發展愈趨數學化、數量化、與電腦化。

5.礙於嚴苛的基本假設，當代測驗理論所能適用的教育與心理測驗資料有限，並且需要大樣本的配合，因此使得它的應用性大打折扣，未獲一般測驗使用者的全力擁護。

由此可知，古典測驗理論雖然不夠嚴謹，但理論淺顯易懂，便於在實際測驗情境（尤其是小規模資料）實施；當代測驗理論雖然嚴謹，但理論艱深難懂，僅適用於大樣本測驗資料的分析。所以，這兩派測驗理論各有所長，在應用上也各有其限制。

## 六、IRT 電腦應用軟體

IRT 電腦常見的應用軟體有：BILOG、BILOG-MG、MULTILOG 以及 PARSCALE 等四種，目前均有視窗版本，由於本研究採用 MULTILOG 7.03，故

以下加以簡單介紹：

MULTILOG 是由 Thissen 於 1991 年所出版，適合二分化計分和多元化計分的測驗資料。MULTILOG 所根據的理論主要有：單參數對數模式、雙參數對數模式、三參數對數模式、等級反應模式 (Samejima, 2000)、名義模式 (nominal model; Bock, 1972) 等。為了分析多元化計分的測驗資料，MULTILOG 可以應用在一些特定的 IRT 模式，在此簡介其中兩種 (Thissen David, 1991; 劉湘川, 民 87)：

(一) 等級模式，其基本假設簡述如下：

1. 假設研究對象在某一試題可能有  $m$  種不同的反應  $x$ ，而且這些反應之間存在有某種次序關係。

2. 假設研究對象的能力參數為  $\theta$ ，則各種反應發生的機率定義如下列公式：

$$\text{假設 } p^i(k) = \frac{e^{a(\theta - b_{k-1})}}{1 + e^{a(\theta - b_{k-1})}} \quad k = 2, \dots, m$$

而且  $p^i(1) = 1$  和  $p^i(m+1) = 0$

$$\text{再假設 } p(x=k) = \frac{e^{a(\theta - b_{k-1})}}{1 + e^{a(\theta - b_{k-1})}} - \frac{e^{a(\theta - b_k)}}{1 + e^{a(\theta - b_k)}} = p^i(k) - p^i(k+1)$$

3. 由上式可知， $P(x=k)$  表示研究對象發生第  $k$  種反應的機率，而  $P^i(k)$  表示研究對象發生第 1, 2, ...,  $k$  種反應的機率總和。

4. 等級模式適用場合的範例：有一張考卷內有 20 個問答題，每題的計分方式為依照學生答案的正確性和完整性給分，最低為 0 分，最高為 5 分。如此學生在每個試題上的得分可能是 0、1、2、3、4、5 等六種可能的結果。

(二) 名義模式，其基本假設簡述如下：

1. 假設研究對象對某一試題可能有  $m$  種不同的反應  $x$ ，而且這些反應只代表不同的反應，它們之間沒有任何次序關係或其他關係。

2. 假設研究對象的能力參數為  $\theta$ ，則各種反應發生的機率定義如下列公式：

$$\text{假設 } P(x = k) = \frac{e^{a_k + c_k}}{\sum_{i=1}^m e^{a_i + c_i}} \quad k = 2, \dots, m$$

3. 名義模式主要用於處理採用名義量尺所產生的測驗或問卷，舉例說明其適用場合如下：某一份問卷有一個問題，其內容列舉五台不同廠牌與款式的汽車，並且詢問研究對象，喜歡其中哪一台？如此研究對象可能的反應有五種（如果不考慮「拒絕回答」），而且每種反應之間沒有一定的次序、大小或前後關係。

表 3：目前常見的試題項目反應理論參數估計的電腦程式

程式名稱	來源	適用模式	估計方法	優點(+)、缺點(-)、特性( )
LOGIST	Wingersky (1983)	一個參數 二個參數 三個參數	非條件化最大近似值	(+) LOGIST 5 提供標準誤 (+) 具彈性，選擇多 (+) 允許未完成 / 空白未答的反應 (-) 資料輸入繁瑣 (-) 設定許多限制，以便獲得收斂的參數估計值
ASCAL	評量系統公司 (1988)	一個參數 二個參數 三個參數	修改過的貝氏估計法	(+) 包括適合度分析 (+) 併入 Micro CAT 套裝程式裡 ( ) 使用貝氏估計法
RASCAL	評量系統公司 (1988)	一個參數	非條件化最大近似值	(+) 包括適合度分析 ( ) 併入 Micro CAT 套裝程式裡
BILOG	Mislevy & Bock (1990)	一個參數 二個參數 三個參數	邊緣的最大近似值	(+) 選擇性貝氏估計值 (+) 可避免極值的估計值出現 (-) 錯誤的前置項假設會導致錯誤的估計值
BILOG-MG 3	Zimowski, Muraki, Mislevy & Bock (1996)	一個參數 二個參數 三個參數		(+) 允許不同群組有不同的難度參數 (+) 可以做等組的試題等化、不等組的試題等化、試題等化、試題功能區分、試題參數演變、兩階段測驗 ( ) 允許在測驗資料中加入有關受試者背景資料，增加各項估計結果的精準度
PARSCALE 4	Eiji Muraki & Darrell Bock (1996)	一個參數 二個參數 三個參數 等級模式		(+) 適合二元計分和多元計分 (-) 非常態分配資料估計值差
MULTILOG 7	David Thissen, Wen-Hung Chen, & Darrell Bock (2003)	多元類別		(+) 適合二元計分和多元計分 ( ) 把 BILOG 程式擴展成能夠處理多元類別資料的程式

( 改自余民寧，民 81b；<http://www.ssicentral.com/> )

## 七、相關研究

試題項目反應理論的相關研究在國內正蓬勃發展，表 4 列舉國內相關論文及期刊研究的簡單摘要。

表 4：國內試題項目反應理論相關論文及期刊研究摘要

研究者	論文或研究篇名	摘要
林世華 (民 72)	潛在特質理論與其應用於適性測驗之評估研究	本研究主要目的在於探討：潛在特質理論與傳統測驗理論二者之間之關係；潛在特質理論應用於適性測驗時，對於心理測量所產生的助益。結果發現：二者之基本觀念上是相通一致的；將潛在特質理論應用於適性測驗是可行的，而且可以改善測量工作。
劉湘川、簡茂發、林原宏 (民 83)	試題關聯結構與試題反應理論之聯合分析研究-以乘法概念之暗隱模式為探討基礎	本研究以試題關聯結構及無參數試題反應理論為分析工具，應用於國民小學數學教育領域，探討學生在乘除方面之概念。根據研究結果，大致能印證「暗隱模式」理論。
郭伯臣、施淑娟 (民 84)	從試題特徵曲線到概念影響曲線	結合試題的知識結構分析及試題特徵曲線二者，發展出所謂的概念影響曲線來說明，學生作答「分數加法」測驗時，各項主要概念影響正確作答的程度，並配合學生作答所產生的錯誤類型加以分析。
劉湘川、林原宏 (民 84)	從試題特徵曲線進行認知診斷之理論與應用	根據試題反應理論之資料分析，從理論與驗證的角度，針對試題特徵曲線，進行認知診斷。從資料的分析結果發現，此認知診斷方法支持數學上的乘除迷思概念理論。
劉湘川、許天維、鄭富森 (民 87)	IRT 軟體估計精確度之比較	本研究同時採用電腦模擬資料與真實測驗資料，利用參數對數模式製造二元計分資料，測量 BILOG-MG、ICCNP、MULTILOG PARSCALE 等四種 IRT 計分軟體實際應用情形與其參數估計精準度。

劉湘川 ( 民 89 )	點二系列相關試題鑑別指數之值譜分析及其在 IRT 上之應用	本文以光譜學之值譜分析概念，證明 50% 高低鑑別試題指數 $D_{50}$ ，不僅在實用上受試者人數限制較少，且鑑別力亦較佳。
劉湘川 ( 民 89 )	多元計分三參數試題選項分析固定效應模式	本研究在證明 Thissen & Strinberg( 1984 ) 之 BS( 多元計分三參數試題選項分析固定效應模式 ) 模式，非二元計分三參數對數型試題分析之擴充模式，並提出後者之真正擴充模式。
何世欽 ( 民 90 )	數學科新式選擇題型之試題分析研究—以「高職免試登記入學方案」彰化區為例	本研究顯示新式題型的優點是比傳統題型來得多，可以在以後大型的升學考試中採用。優點有：1.除可以用統計的方式來分析外，透過雙向細目表分析發現可以測驗受試者較高階的數學理解與應用層次的的能力。2.給予中、低數學程度的學生有所表現的機會。3.在試題統計分析上有傑出表現，建立題庫是可信且可行的。4.學生對於新試題型的適應情形相當良好。
林育柔 ( 民 90 )	試題選項特徵曲線分析法-在「國小中年級面積概念」的應用	本研究旨在以 OCC2001 的試題分析應用軟體，配合 TestGraf98 繪圖功能，繪製試題選項特徵曲線，以診斷自編「國小中年級面積概念測驗」試題的良窳，據以分析國小中年級學生對面積概念的了解情形。
郭舒華 ( 民 90 )	多重相關加權核平滑多元計分試題選項特徵曲線之研究	以 Matlab 為工具進行實例應用，畫出試題選項特徵曲線，估算受試者能力值，以 logit 轉換高低鑑別指數作為受試者加權總分排序時之加權函數，稱為「核平滑化無參數試題選項特徵曲線估計法」。
李亭儀 ( 民 91 )	核平滑化試題反應理論與試題選項關聯結構分析法之整合模式	本研究之目的在結合試題選項分析與 IORS 整合模式以發展出的電腦應用軟體，應用於國小六年級數學科「分數除以整數」的單元，分析其不同能力值受試者的試題選項關聯結構圖並比較不同臨界值的試題選項關聯結構圖。

吳孟勳 (民 91)	應用模糊集合理論與試題反應理論於學習評量之研究	電腦適性測驗採用試題反應理論，依答題結果來估算能力，容易忽略受試者在評量過程中之知識狀態，以致產生相當的模糊性。本論文提出一應用模糊集合理論與試題反應理論之學習評量模式，整合主觀與客觀的受試者評量資訊，藉以改善一般電腦適性測驗系統未能考量知識狀態差異之缺失，做為教學者評量受試者學習成效之工具。
姚漢禱 (民 91)	用試題反應理論估計運動項目的成績表現排名	以試題反應理論之 FACETS 程式，將順序資料轉換為近似等距數線（連續性數線）的量尺，可以提供更多測驗訊息。
張育璋 (民 91)	以知識地圖為基礎發展的遠距測驗選題策略	基於知識地圖的觀念，以概念階層為主的概念式選題策略，同時並使用隱含概念矩陣與雙向式累進標的概念向量來計算所選試題組合之內容效度，最後利用基因演算法來完成選題策略中試題組合最佳化的問題。透過與迷思概念診斷式測驗系統之整合，將可完成概念診斷式電腦化適性測驗之實現。
楊明宗 (民 91)	數學試題分析模式的建制	利用試題檢核表、雙向細目表、內容效度檢核表作質的分析，利用試題反應理論對試題進行量的分析，進行 IRT 假設的檢定及模式適合度檢定。
郭伯臣、許天維、黃志傑、曾玉琳 (民 92)	定錨試題分佈對測驗等化效果之影響	探討三參數 logistic 模式中，定錨試題分佈對於測驗等化的影響。研究發現最佳選擇會因定錨題數及受試者能力分佈不同而異，若定錨題數愈多且分佈適當，則等化估計愈精準。
陳怡如、吳慧黃碧雲 (民 93)	電腦化適性診斷測驗之研究	本研究將資料依據順序理論及試題關聯結構分析法進行整理，進而發展一電腦化之適性診斷測驗，以節省試題和預測精準度。

(資料來源：研究者參酌文獻整理)

## 第四節 壓力的理論與相關研究

### 一、壓力的定義

壓力 (stress) 一詞，在日常生活中使用甚廣，最早源自於物理上。此一議題源自於 Cannon 於 1929 年描述身體面對緊急狀況時會產生「戰或逃」(fight or flight) 的反應 (王文豐，民 89)。繼之 Selye 於 1956 年首先將壓力的概念引進社會科學的領域，開啟研究壓力的先河。

研究壓力的相關文獻探討相當多，但研究者因所持觀點的不同而其涵義亦甚分歧，至今尚無明確一致的看法。表 5 列舉國內、外部分研究者對壓力所做的釋義 (蔡璧煌，民 78；張春興，民 86)。

表 5：國內、外研究者對壓力釋義彙整一覽表

研究者	壓力的釋義
周天賜 (民 80)	壓力是指外在刺激超過個體容忍度，造成個體的焦慮或逃避
王以仁 (民 81)	壓力是個體預期未來可能發生的不安或威脅有所知覺，因而對有機體產生刺激、警告或使其活動
吳英璋 (民 83)	個人的因應能力與環境要求呈現不平衡時的反應
馮觀富 (民 83)	壓力是指發生於當事人在認知經驗上受到威脅時的一種反應狀態，引發內在身體上、認知上、行為上與情緒上的連鎖反應
葉重新 (民 84)	壓力是衝擊個人內外的一種刺激，是個人對內外環境事件產生心理上或生理上的反應，也是個人與環境交互作用的結果
江鴻鈞 (民 84)	壓力不外是個人生理和心理反應的一種刺激，更是個人與環境交互作用的結果
程一民 (民 85)	當個體面對刺激事件時，個人評估此一關係對他造成負荷或個人資源無法應付，且危及個人福祉、身心健康時

	的反應，此反應以個人特質、經驗與認知歷程為媒介
邱憲義（民 85）	壓力是知覺情境有礙其表現，所產生的負向情感，而影響其生理與心理狀況
林杏霞（民 86）	壓力為任何會使個體身心產生不舒服感覺的情境或反應；亦即任何會打破個體身心恆常狀態，引起身心不平衡的情境反應
林純文（民 86）	當個體面對外界要求或特定事件的刺激時所做身心適應的反應，而此種反應是以個人過去的經驗、人格特質、經驗與認知歷程為媒介
呂秀華（民 86）	壓力是個人所知覺的由外在刺激所引起的緊張狀態
張春興（民 86）	是指個體生理或心理上感受到威脅時的一種緊張狀態。此種緊張狀態，使人在情緒上產生不愉快甚至痛苦的感受。
李玉惠（民 87）	個體在面對具有威脅性的刺激情境時，因一時無法消除威脅、脫離困境而產生一種被壓迫的感受或緊張狀態，且此種狀態在生活上或工作上持續存在
葉照祺（民 88）	壓力是個體面對環境要求時，造成環境、心理、生理三方面的失衡狀態
陳青勇（民 90）	壓力是個人與環境互動的產物，惟個體感受到威脅時，所引起生理與心理的變化
藍采風（民 90）	壓力是指在某種情境下，使個人覺得像受到某種程度或種類的威脅，使人感覺到必須付出額外的精力以保持身心平衡。不適的程度會因個人的體質與身體健康狀況、文化、社會價值與規範而有差異。
陳淑芬（民 91）	個體對於事件之刺激主觀感受到的程度
Selye（1956）	壓力是任何環境要求下，在身體內部所產生的一種非獨特性之連鎖反應
Miller（1976）	「壓力是外在環境向系統輸入過多的訊息，以致於耗損了系統的能量，而使其系統無法負荷時所產生的狀態
Ivancevich & Matteson（1980）	壓力是一種適應性的反應，視外界的動作、情況或事件對個人產生特殊之生理或心理要求而造成的結果。且此種反應是以個人特質或心裡力成為中介
Statt（1981）	壓力是個體長期處於身體及心理的緊張狀態，而言重影響到期處理特定情況的應變能力
Martin &	壓力是個人對環境中的要求、限制與機會所產生的主觀

Schermerharm( 1983 )	知覺，而發生正向或負向的反應狀態
Lazarus & Folkman ( 1984 )	壓力是當個人評鑑這個事件，對它造成負荷或不是其資源所能應付，且危及身心健康時所產生的壓力感
Hobfoll ( 1995 )	壓力是個體面臨下列三種情境時的反應：1.覺察到失去網絡的恐懼；2.真正失去資源網絡；3 缺乏資源。所謂資源是指個體認為有價值或有意義的人物、情況、人格特質或能量
Greenberg & Baron ( 1997 )	壓力是指個體面對外界壓力源時，情緒、認知和生理三方面交互作用所產生的一種反應模式

( 資料來源：研究者參酌文獻整理 )

## 二、壓力的理論和取向

有關壓力理論模式的研究，大多是以生物學和社會心理學為主要觀點，只是這些理論模式可幫助我們瞭解壓力，但卻無法說明壓力的過程。大致可分五種( 吳英璋等，民 81；沈翠吟，民 92；陳佩汝，民 90 )。

( 一 ) 生物心理社會模式 ( biopsychosocial model ) : Engel 在 1977 年所提出的生物心理社會模式，是健康心理學領域裡最常被運用的理論，他認為一個健康的人是生物 ( 例如：病毒和傷害 ) 心理 ( 例如：態度、信念、心理需求、人格、過去經驗 ) 和社會 ( 例如：階級、職業、支持系統 ) 等因素交互作用的結果。此模式說明壓力乃源於人與環境間的某種特殊關係，也就是說壓力是一種個別性的現象，某種情境是否對人造成威脅，需視個人與環境間的關係而定 ( 陳綠蓉，民 85 )。

( 二 ) 個人 - 環境適配理論模式 ( person-environment fit model ) : French 與 Kahn ( 1962 ) 所提出，以「適配」( fit ) 的觀點來討論人與情境的互動關係。所謂「適配」分兩種，一是個人需求與環境供給的適配程度；另一是個人能力與環

境需求的配合程度。因此，王秀枝（民 75）指出個人與情境間包含兩種適配方式：客觀的適配，即客觀的個人（獨立於知覺外的個體）與客觀的環境相配合的情形。另一種是主觀的適配，即主觀的個人（自我知覺的個體）與工作環境知覺間的配合情形。故依據此理論，當情境的要求遠超過個人知覺滿足這些要求的能力時，或是無法滿足這些要求卻會被視為有嚴重的後果時，都會產生壓力。也就是說，壓力的大小是取決於個人與環境相適的程度

（三）壓力循環理論模式（stress cycle model）：Gmelch（1988）歸納過去學者之壓力理論研究而提出，包括四個階段：

第一個階段是壓力來源，即壓力因素的刺激，其結合了特定的壓力源（stressors），或加諸於個人的特殊需求而導致。

第二個階段是個人知覺，就是個人對壓力的知覺，包括生理與心理的反應，且兩者相互影響。

第三個階段是因應反應。即當個人生理或心理無法負荷時，就認定這些要求是壓力，面對壓力的因應方式，各有不同，其因應方式有：智能的、人際的、身體的、管理的、環境的、態度等。因應方式需彈性且須考慮社會、文化、心理的因素。

第四個階段是結果，就是個人長期承受強烈且負面的壓力，而造成生理或心理上的疾病。

（四）Jacobson 的壓力模式：Jacobson（1986）由支持概念發展出來的，有三種最受注意：

1. 模式需求（need）：個人經由各種社會關係的互動而達到需要的滿足感，若無法滿足則會產生焦慮。

2.交互模式 ( transaction)：當察覺到的外在要求超過知覺到的資源，且對個人在幸福感方面有負面影響，則會產生壓力。

3.變遷模式 ( transitions )：壓力的來源是來自於「變遷」，其中包括獲得、失落或者是心理社會的變遷。

(五) Ivancevich 及 Matteson 的工作壓力模式：Ivancevich 及 Matteson( 1980 ) 提出，強調「個人知覺」是社會壓力源與壓力導致的結果之間的中介歷程，因此個人的特質與認知成為決定壓力產生的重要關鍵。被一些學者認為是傳統的工作壓力模式，此模式包括五種變項，如圖 5 ( 黃義良，民 88 )：

- 1.是來自社會的壓力源。
- 2.是個人對於壓力的知覺。
- 3.是個人對於壓力所產生的反應，包含生理、情感與行為三方面。
- 4.是長期壓力反應下所產生的身心症狀。
- 5.是屬於個人或情境的制約變項。此模式。

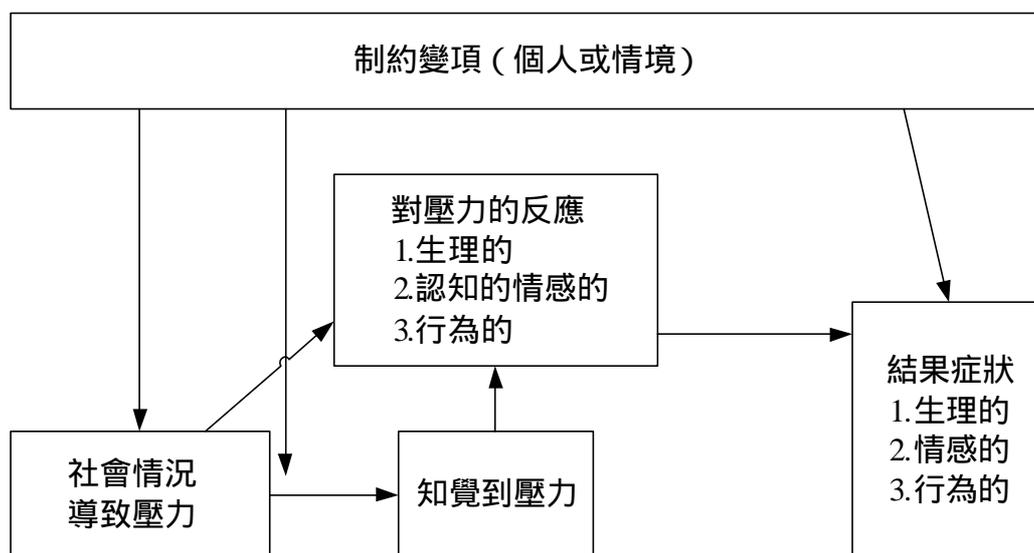


圖 5：Ivancevich 及 Matteson 的工作壓力模式  
( 改自黃義良，民 88，頁 25 )

Lazarus 與 Folkman ( 1984 ) 綜合各家看法，將壓力的定義歸納為以下三個取向：

#### (一) 刺激取向

刺激取向的概念著重在環境方面，將壓力視為環境中客觀存在的威脅性刺激。主要以 Holmes 與 Rahe ( 1967 ) 最具代表性。他們是最早研究生活壓力測量的學者，視壓力是遇到外界事件而失去生活平衡時，個體為了恢復原有狀態所花費之精神與體力的總量，並以個人客觀的「生活變動值」( Life Chang Unit, 簡稱 LCU ) 做為壓力的指標。

此取向強調壓力是外在的事件或歸因於環境因素，這些被認為具有威脅性或傷害性的刺激，被稱為壓力源，這些壓力源大致可分為三類：

1. 災難事件，例如颱風或地震。
2. 重大生活事件，如失去所愛的人或工作。
3. 長期的生活問題，如環境太吵雜。

#### (二) 反應取向

反應取向則將壓力視為一種個人對壓力源的反應，以 Selye 為代表。Selye 主張，當有機體面對壓力源時，將促使他自己採取行動 ( action )，是身體為了保護自己而產生的一般性反應，這種動員一般稱為一般性適應症候群 ( general adaptation syndrome, 簡稱 GAS )，為個體在環境刺激的要求下，所產生的「非特定反應」( non-specific response )。大致上可以分為三個階段 ( Selye, 1974 ; 黃寶園, 民 93 ; 藍采風, 民 92 ):

第一階段為警覺反應階段 ( alarm reaction stage ): 在警覺期，身體透過刺激交感神經系統，分泌荷爾蒙對抗壓力源，激起身體系統產生最大的力量和準備產

生「戰或逃」的反應。釋出的腎上腺素會增加心跳速率和血壓、呼吸變快、轉移內在器官的血液流到骨骼肌上、汗腺分泌加快、胃腸的蠕動減緩。此為緊急情境的短期反應，是為了適應目前的壓力情境。

第二階段為抗拒階段（resistance stage）：抗拒期是指對壓力的一種相對反應，它代表身體對壓力的有效應變，使壓力降低或至少使個人有能力與壓力相抗衡。這個階段，有機體處在壓力源之下，若個體能夠適應這種壓力，則抗拒階段會持續很長一段時間。在這個階段，個體的外在表現正常，但身體的內在功能則否，持續的壓力會導致神經和荷爾蒙的改變。Selye 認為，長期對壓力的抗拒結果，會造成消化性潰瘍和結腸炎、高血壓和心血管疾病、甲狀腺機能亢進和支氣管氣喘。

第三階段為耗竭階段（exhaustion stage）：當個人長期與壓力相抗衡，到了某階段便會逐漸感到精力耗盡，而且對壓力的抗拒力低於正常程度，最後個體能力衰竭，以致崩潰。這個階段副交感神經的功能出現不正常的現象。耗竭階段的徵候有所謂衰老現象，個體出現憂鬱的傾向，或個人受疾病打擊，甚至導致死亡。

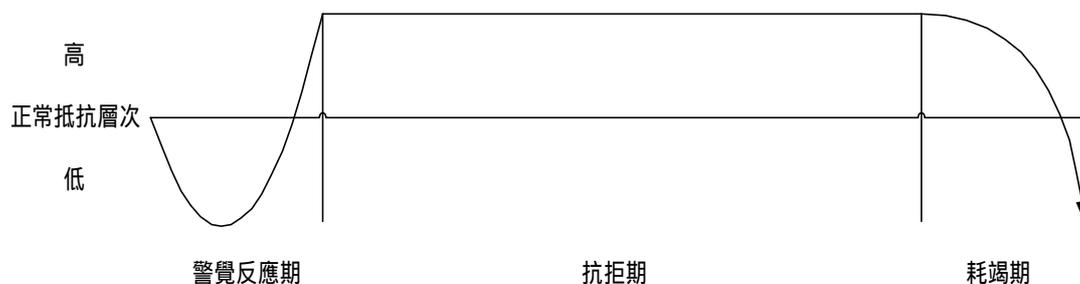


圖 6：Selye 對壓力反應之 GAS 理論

（改自陳筱瑄，民 91，p19）

### （三）互動取向

互動取向把壓力視為一種歷程，為刺激與反應間的交互作用，個體受到環境

的刺激，透過其主觀的認知過程而產生心理威脅和生理衝動的現象，是個體和環境互動的結果，意即並非所有個體在相同刺激下，都會產生相同的壓力反應。

最具代表性為 Lazarus 與 Folkman (1984) 的研究，他們提出的互動模式除了強調壓力的互動本質外，也著重個體的認知和心理層面，強調壓力是個人與環境間互動的結果，壓力不是刺激事件，也不是個體的反應狀態，而是個體在環境變動與因應資源間所做的「認知評估」的結果，注重認知在壓力形成歷程中所扮演的角色。他們認為壓力是個人與環境間的特殊關係，個人評估此一關係是對他造成負荷的或是超出他的資源所能應付的，而且危及個人的幸福及健康。因此個人透過此認知評估來衡量個人能力與環境要求是否能達到平衡，若不能達到平衡時，壓力便會產生。故將其分為初級評估 ( primary appraisal ) 和次級評估 ( secondary appraisal )。

Bandura 則認為環境和行為及個人是相互影響的，他指出個人的認知評估會受環境影響，而產生不同的行為；同樣的，個人行為會因個人評估及環境的不同而不同，總之，人、環境和行為之間是相互關聯的 ( 陳筱瑄，民 91 )。

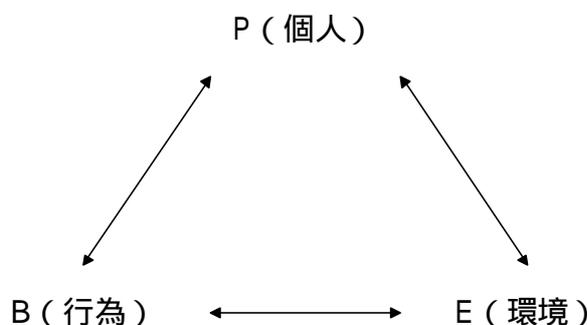


圖 7：Bandura 的環境互動模式

( 取自陳筱瑄，民 91，p20 )

Cox (1978) 提出了更複雜的交流模式，主張壓力是個人與環境間存在的某種特殊關係，是人與環境不斷互動交流的系統，存在於個體所知覺的環境需求

與所知覺的反應能力之間的不平衡所致。在壓力過程中，自我調節的認知、行為與情緒的因應策略會影響壓力源對個人的衝擊。其中，在壓力因應過程中造成反應差異的因素可分為內在及外在兩大類。內在因素包含個人變項及人格特質；外在變項則來自於情境及環境，包括壓力源的出現和持續時間，社會網絡等。

藍采風（民 90）提到，人們如何反應外在事件的刺激，才是壓力或壓力經驗，而個人對壓力的反應會對生理、心理和社會產生全面性的影響。因此：

「壓力來源 + 個人對壓力源的反應 = 壓力經驗」

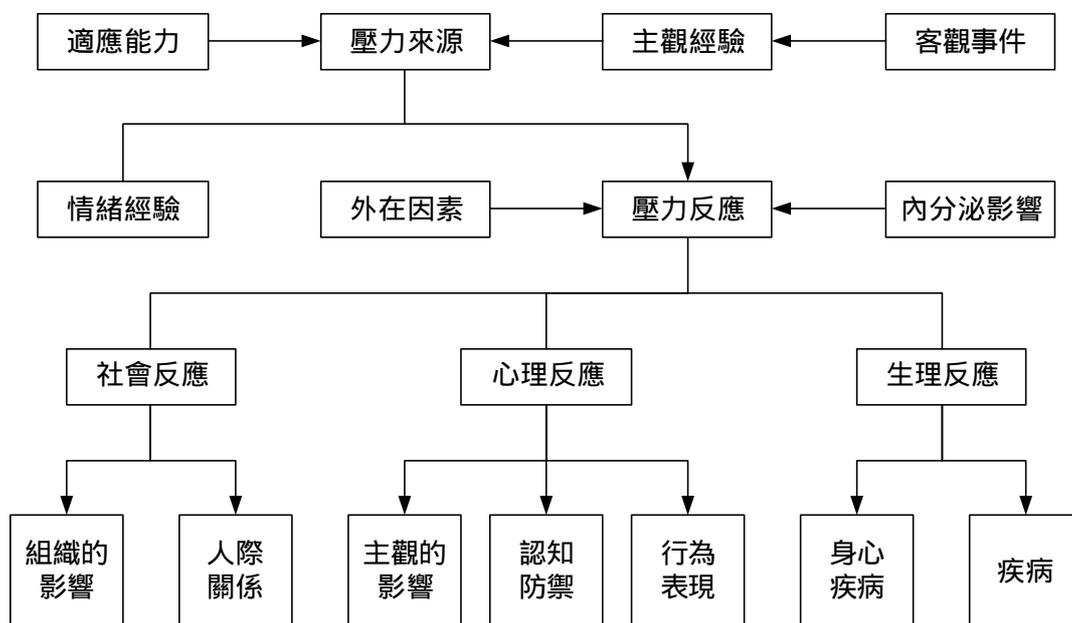


圖 8：壓力反應之生理、心理與社會面  
（取自藍采風，民 90，p52）

表 6：三種壓力取向的比較

壓力取向	刺激取向	反應取向	互動取向
個體、環境關係	單向	單向	雙向，互為因果
個體角色	被動	被動	主動
壓力測量	客觀生活壓力	主觀生活壓力	主觀生活壓力
優點	客觀、容易測量	重是個體的主觀感受及反應	強調個體知覺的重要性，並評估個體能否運用資源加以調適
缺點	1.忽略個體主觀認知的感受 2.混淆壓力與身心疾病的因果關係	1.忽略壓力來源和個體認知因素 2.不確定壓力源，難以提出確定診斷	難以證實壓力與身心疾病的因果關係
適用領域	心理、教育	生物醫學	心理、教育

(整理自李欣瑩，2001；陳柏齡；1999)

綜上所述可知：壓力刺激是多層面存在的，是許多因素交互作用形成。刺激取向的觀點，雖然著重以客觀標準來測量壓力，容易找出壓力源，但卻忽略了個體對壓力主觀感受認知的重要性。若以反應為取向，則忽略了同樣的生理反應，可能源自不同的壓力事件，忽略了壓力源與個人的認知層面的關係。若以互動為取向，則可看出壓力是一個過程，即個體「認知」到環境的需求，且評估超過個人的「因應」能力所產生的。由於互動取向兼採刺激與反應兩方面，並提出兩者交互作用時「認知評估」的重要性，可以算是較為周延的觀點，因此本研究以互動取向理論為依據，依據生活型態評價問卷（Lifestyle Appraisal Questionnaire，LAQ）第二部分，以個人對需求的認知性評價來評估壓力程度，意即以壓力的主觀認知視為個人壓力感受度。

### 三、壓力的測量

黃寶園（民 93）指出，過去有關壓力的測量大致可分為表現測驗、生理測量和自陳量法等三種。下面分別依序簡介：

#### （一）表現測驗

壓力都會影響表現，因此表現測驗就是去測量人們暴露於壓力源下的副作用以及表現能力。

#### （二）生理測量

生理測量法是測量各種生理的指標例如血壓、心跳、呼吸速率，和生物化學的測量包括糖皮質固醇和兒茶酚胺的分泌。這個方法的優點是直接、可靠性高和容易量化，但缺點是電子儀器和門診環境本身就會產生壓力。

#### （三）自陳量法

自陳量法包括測量生活事件或日常瑣事，在健康心理學中最重視自陳法。

##### 1.生活事件量表：

最早而且最著名的是 Holmes 和 Rahe 於 1967 年所發展的「社會再適應量表」(social readjustment rating scale, 簡稱 SRRS)。個體對於壓力的感受程度有個別差異，事件本身是中性的，只有當個體將事件評估為壓力源時壓力才存在，所以 SRRS 測量個人在過去特定時間內所遭遇的生活壓力大小，量表裡每一項生活事件均有一個壓力分數，代表此事件的生活變動值，亦即將過去特定時間內所遭遇的生活事件勾選出來，加總所有生活事件的生活變動值，即可得到個人在特定時間內所遭遇的生活壓力大小（黃寶園，民 93）。

表 7 列舉出所選 43 項的生活重大事件，並採用大小估計法，評量出各生活事件的平均值，視為該事件的壓力值。此量表的使用是評量受試者在過去一段特定的時間內（通常是一年或一年半內），個人所遭遇到的事件壓力值總和，這個

值稱為生活改變值 ( life change units , 簡稱 LCU)( 陳筱瑄 , 民 91 )。

表 7：社會再適應量表 ( SRRS 中 43 項生活重大事件 )

等級	生活事件	平均值	等級	生活事件	平均值
1	配偶死亡	100	23	子女離家	29
2	離婚	73	24	法律問題困擾	29
3	夫妻分居	65	25	個人傑出成就	28
4	牢獄之災	63	26	太太外出工作或離家	26
5	親密家族成員死亡	63	27	入學或畢業	26
6	嚴重個人受傷或疾病	53	28	生活環境重大變化	25
7	結婚	50	29	個人習慣調整	24
8	被解雇	47	30	與上司相處差	23
9	夫妻復和	45	31	改變工作時間或狀況	20
10	退休	45	32	遷居	20
11	家屬健康變化	44	33	轉學	20
12	懷孕	40	34	休閒方式改變	19
13	性生活障礙	39	35	宗教活動改變	19
14	增添家族成員	39	36	社交活動改變	18
15	工作整頓	39	37	小額貸款	17
16	重大財物變化	38	38	睡眠習慣改變	16
17	密友死亡	37	39	家人團聚次數改變	15
18	換工作	36	40	飲食習慣改變	15
19	與配偶吵架	35	41	度假	13
20	鉅額貸款	31	42	聖誕節	12
21	貸款抵押物被沒收	30	43	輕微違法	11
22	職務改變	29			

( 取自藍采風 , 民 90 , p127-128 )

Lazarus 與 Folkman ( 1984 )、陳淑芬 ( 民 92 )、陳筱瑄 ( 民 91 ) 均指出 SRSS 有下列缺失：

(1) 許多壓力的產生是不需經歷生活改變的。

(2) 忽略個人對生活事件的主觀感受、價值信念和人格特質的差異，而以生活事件的平均值代表壓力大小，假定個人 LCU 超過 150 即有危機，而忽略壓力容忍度個別差異的可能性。

(3) 將心理壓力當作造成疾病的主要因素，忽略生活狀況和環境條件的重要性、未考慮到壓力事件是有正負向。

(4) 主觀壓力評量方式較客觀壓力評量對心理不健康有高的解釋力（彭秀玲，民 75；蔡嘉慧，民 87）。

(5) 這個測量方法與身心健康的相關很低，平均在 0.12 左右，很少超過 0.30（Rabkin & Struening，1976）。因此可知，事件本身是否帶給個體心理及情緒上的衝擊，取決於個體對該事件的主觀知覺評估，事件本身是中性的，只有當事人將其評估為壓力來源時，該事件才會具有壓力性（李欣瑩，民 89）。

(6) 關於 SRRS 中的生活事件，Lazarus（1985）指出：不同組群的人有其特殊的生活事件，SRRS 無法涵蓋所有年齡層和社經地位的人。

(7) 根據 Rahe 等人（1974）指出：生活事件有群集（cluster）的現象，意某一生活事件發生與否，會影響及另一些事件的發生。SRRS 中的許多項目有四個主要的群集：工作、個人或社會的生活、婚姻關係、最近受罰的行為。Lei 和 Skinner（1980）以因素分析發現，SRRS 的所有生活事件可得到下列六個因素：個人和社會性的活動、工作的改變、婚姻的問題、居住環境的改變、家庭問題及學校生活的改變。

## 2. 日常瑣事量表

Lazarus（1984）強調從日常瑣事，而不是重大生活事件的取向測量壓力。因此，Lazarus 和他的同事便發展出日常瑣事量表，由個人認為是麻煩的、惱人

的或是挫折的 117 個項目組成。這份量表首先要求填答者檢視過去一個月內發生的瑣事，接著在三點量表中評量每件事情的嚴重度（李新鏘，民 86）。

Kanner, A. D., Coyne, J. C., Schaeffer, C., & Lazarus, R. S. (1981) 發現瑣事與生活事件只有中度相關，但瑣事量表比生活事件量表能更精確地做為心理健康的預測。

#### 四、生活型態評價問卷（Lifestyle Appraisal Questionnaire，簡稱 LAQ）

Craig, A., Hancock, K. 與 Craig, M. 以多元文化的觀點於 1996 年發展出來的問卷，用以評估生活型態，藉著評量過去八週累積危險的因子和壓力知覺來評估生活型態。LAQ-Part I 有 21 題，項目包含吸菸行為、血壓、身體質量指數、攝取酒精飲料數量、藥物使用、家族病史、運動休閒、飲食型態、社會支持、生活壓力事件數、慢性疾病、身體症狀和睡眠，用來評估生活型態疾病的危險因子；LAQ-Part II 有 25 題，藉由壓力和需求的認知性評價來評估壓力程度（黃惠貞，民 91）。

由於本研究是在測量壓力知覺的程度，且 LAQ-Part I 和 LAQ-Part II 並無關聯，故只使用 LAQ-Part II 來測量研究對象的壓力知覺。