

## 第二章 文獻探討

本章是研究所需的理論基礎以及相關的研究；包含測驗理論、電腦化測驗、空間能力之相關理論、網站互動式理論、系統開發程序之相關理論等。

### 第一節 測驗理論

測驗理論是一種解釋測驗資料間實證關係 (empirical relationships) 的理論學說 (Crocker & Algina, 1986; 余民寧, 民81)。測驗理論學者通常把測驗理論劃分為古典測驗理論及當代測驗理論兩大取向。古典測驗理論主要是以真實分數模式為主軸，亦即觀察分數等於真實分數與誤差分數之和，而當代測驗理論主要是以試題反應理論為架構；底下就古典測驗理論、試題反應理論、真分數等化分述如下。

#### 壹、古典測驗理論

古典測驗理論其內涵主要是以真實分數模式 (亦即，觀察分數等於真實分數與誤差分數之和) 為理論架構，此理論模式發展已久，其計算公式容易瞭解，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。

雖然古典測驗理論採用的計算公式簡單明瞭、淺顯易懂，但卻具有以下五項缺失 (Crocker & Algina, 1986; 吳裕益, 民81; 余民寧, 民81)：

1. 同一份試卷無法獲得一致的難度、鑑別度及信度；因為這三個測驗指標會受到受試者樣本不同而產生不同之結果。
2. 沒有考慮受試者的個別差異；因為傳統測驗理論以一個相同的測量標準誤來做為每位受試者的測量誤差指標，這樣的方法對高、低能力的受試者而言並不適當，無法測出受試者之實際能力。

3. 傳統測驗理論對於非複本（nonparallel form）但功能相同的測驗所測得之成績，無法提供有意義的比較。
4. 傳統測驗理論對於信度假設是建立在複本（parallel form）測量的概念假設上，然而這個假設往往不存在於實際測驗情境裡，因為不可能要求每位受試者接受同一複本測驗無數次，而依然假設每次測量之間的結果都不相關。
5. 傳統測驗理論假設得分相同的受試者，其能力必然相同，而事實上，得分相同的受試者，其能力估計值會因為反應組型之不同而可能不同。

## 貳、試題反應理論

一般說來，當代測驗理論的發展主要是為了克服古典測驗理論的缺失，其內涵是以試題反應理論為理論架構。關於試題反應理論的意義，學者Lord（1968）認為試題反應理論是個體的行為可以被證明其真實的程度；在定義人類特質後，依據這些特質，估計個體的能力，並以獲得之數值，在相關情境下預測或解釋個體的表現。而學者Hambleton & Swaminathan（1985）也認為試題反應理論是在測驗的情境下，定義受試者特徵及相關特質或能力，則受試者在測驗上的表現即可被預測或解釋，在這些特徵上估計受試者的得分，又稱為能力分數，並使用能力分數去預測或解釋試題或測驗表現。試題反應理論是建立在三個基本假定上（Hambleton & Swaminathan, 1985；余民寧，民81）：

1. 單向度（unidimensionality）：指測驗中每個試題都能測量到同一共同能力（abilities）或潛在特質（latent traits）。Hambleton & Swaminathan（1985）認為單向度是對受試者的測驗表現，唯一必須解釋或證明的能力或特質。
2. 局部獨立性（local independence）：即各試題之間無相關存在，也就是一個試題不能提供另一個試題線索（Crocker & Algina, 1986）。

3. 非速度測驗 (non speeded tests)：試題反應理論認為受試者作答不是一個測驗，是因受試者的能力問題所致，而不是由於作答時間不夠所致。

試題反應理論是依據受試者的測驗表現結果，經數學模式的運算，評估受試者能力和測驗反應間之關係，也就是以機率的概念來解釋受試者能力和試題反應間之關係，此數學模式稱之為試題特徵函數 (item characteristic function, ICF)，由於函數中採用參數個數不同，亦存在不同式子。以下就有關試題反應理論重要的課題做一概述，包括試題參數、試題特徵函數、能力參數估計方法及試題訊息函數。

#### 一、試題參數

試題反應理論將試題視為測量受試能力的基本單位，而試題的特質可以用三個試題參數來描述 (Hambleton, Ronald K., Swaminathan, H. & Rogers, Jane, 1991；陳麗如，民 86；陳英豪、吳裕益，民 89)。

1. 「鑑別度參數」，它的值愈大表示試題對受試者之鑑別力愈強；反之，則鑑別力愈弱。鑑別度參數對試題特徵曲線 (item characteristic curve, ICC) 的反映是在曲線的斜率，其理論值範圍介於 $-\infty$ 與 $+\infty$ 之間，在實際應用上常限定為 0 至+3。
2. 「難度參數」，它反映試題特徵曲線 (ICC) 位於能力量尺上之位置，因假定受試者在能力量尺之能力值範圍為 $-\infty$ 至 $+\infty$ ，所以難度參數的理論值範圍亦為 $-\infty$ 至 $+\infty$ ，不過在實際應用上常限定受試能力值介於-3 至+3 之間。
3. 「猜測度參數」，即代表猜對之機率，它的值愈大表示不論受試能力為高或低皆容易猜對，值愈小則不易被猜對。猜測度參數對試題特徵曲線 (ICC) 的反映是曲線的左下角漸近線縱軸的大小值，其理論值範圍介於

0.0與1.0之間，在實際應用上過高之猜測度參數常不被接受。

## 二、試題特徵函數

試題反應理論以試題特徵函數表達受試者能力和測驗反應間之關係，因函數中所採用的參數個數不同，可被區分為不同的模式，常用的數學模式有單參數、雙參數及三參數等三種，各模式之試題特徵函數如公式 1 至 3 所示 (Hambleton & Swaminathan, 1985)。

$$P_{ij}(\theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}} \dots\dots\dots \text{公式 1.}$$

$$P_{ij}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \dots\dots\dots \text{公式 2.}$$

$$P_{ij}(\theta_j) = C_i + (1 - C_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \dots\dots\dots \text{公式 3.}$$

其中  $D=1.702$ 。

$e$ ：自然對數。

$j$ ：受試者編號。

$\theta_j$ ：第  $j$  位受試者之能力值。

$i$ ：試題編號。

$b_i$ 、 $a_i$ 、 $c_i$ ：第  $i$  題的難度參數、鑑別度參數、猜測度參數。

$P_{ij}(\theta_j)$ ：能力  $\theta_j$  答對第  $i$  題的機率函數。

下圖即是一試題特徵函數所繪之試題特徵曲線，橫軸  $\theta$  表示受試者的能力分佈，縱軸  $P$  則為受試能力  $\theta$  答對此題的機率。在線上有三個點 A、B、C，分別代表三位受試者在此試題的答對機率，由圖可知能力值  $\theta$  愈高者，答對此試題的機率就愈大。

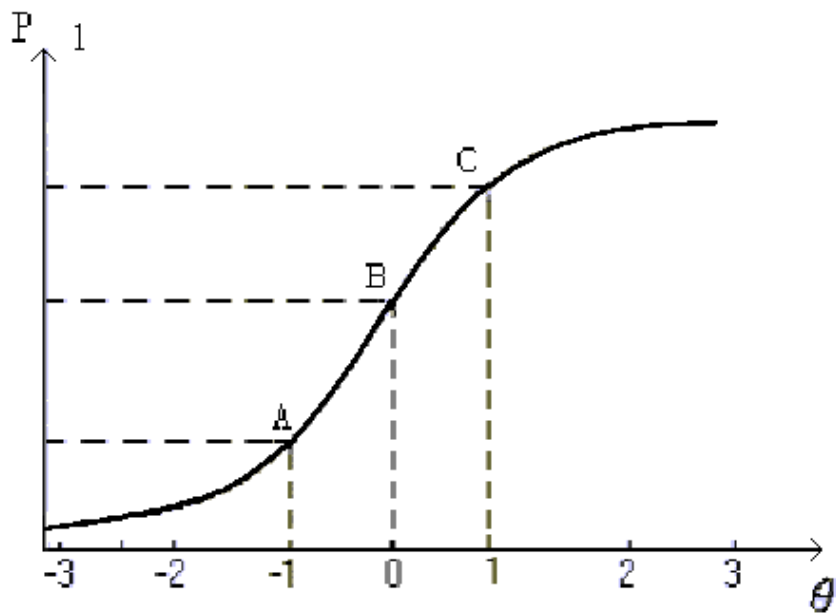


圖 2-1-1 試題特徵曲線 (ICC)

### 三、能力參數估計方法

電腦化適性測驗的能力估計法大部分是採用最大概似估計法 (Maximum Likelihood Estimator, MLE) 或者是貝氏估計法 (Bayesian estimation)。最大概似估計法假設受試者的反應結果與該受試者能力會形成一概率函數，如公式4 (Hambleton & Swaminathan, 1985)。

$$L(u_1, \dots, u_n | \theta) = \prod_{i=1}^n p_i^{u_i(\theta)} Q_i^{1-u_i(\theta)} \dots\dots\dots \text{公式 4.}$$

其中  $L$ ：概率函數值。

$u_i$ ：受試者在第  $i$  題的得分情形，如果反應正確，其值為 1；反應錯誤，其值為 0。

$n$ ：測驗的試題數。

$\theta$ ：受試者能力值。

$p_i(\theta)$ ：具能力值  $\theta$  之受試者答對第  $i$  題的機率。

$Q_i(\theta)$ ：具能力值 $\theta$ 之受試者答錯第 $i$ 題的機率。

雖然概率函數值隨著不同的 $\theta$ 值而異，但是概率函數只有一個最大值，這個最大值即為概率函數曲線的最高點，而它所對應的 $\theta$ 值就是受試者能力估計值。

貝氏估計法的估計程序為依據事先預設之受試能力值，代入答題結果以計算出後驗機率分配之概似函數最大值，此估計值即為受試者能力估計值。貝氏估計法的最大優點為可避免使用最大概似估計法在受試者全對或全錯的情況下，受試能力值的估計無法收斂的缺點（Mislevy, 1986；余民寧，民81），貝氏估計法在受試者全對或全錯的情況下，雖可計算出能力估計值，但會產生迴歸效應（Weiss, 1982），即受試者的能力估計值會有偏向先驗分配之平均數的趨勢，而造成估計的偏差，此情形在試題較少時特別明顯。

#### 四、試題訊息函數（item information function, IIF）

有關試題訊息量和受試能力間的關係則以試題訊息函數（IIF）來表示，如公式 5 所示（Hambleton & Swaminathan, 1985）。若將函數式以圖形表示則稱為試題訊息曲線（IIC）（Lord, 1985）。

$$I(\hat{\theta}) = D^2 \sum_{i=1} a_i^2 Q_i(\hat{\theta}) \left( P_i(\hat{\theta}) - c_i \right)^2 / (1 - c_i)^2 P_i(\hat{\theta}) \dots\dots\dots \text{公式 5.}$$

其中 $\hat{\theta}$ ：受試者能力估計值。

$I(\hat{\theta})$ ：試題 $i$ 對能力估計值為 $\hat{\theta}$ 之受試者所提供的試題訊息量。

$i$ ：是試題編號。

$a_i$ 、 $c_i$ ：分別代表第 $i$ 題的鑑別度參數、猜測度參數。

$P_i(\hat{\theta})$ ：具能力值估計值 $\hat{\theta}$ 之受試者答對第 $i$ 題的機率。

$Q_i(\hat{\theta})$ ：具能力值估計值 $\hat{\theta}$ 之受試者答錯第 $i$ 題的機率。

當將受試者能力估計值代入公式，算出題庫中所有試題的訊息量，而訊息量值最大的試題就用來施測。圖2-1-2為一三參數試題訊息曲線（IIC），橫軸 $\theta$ 表示受試者的能力分佈，縱軸 $I(\hat{\theta})$ 則為此試題對不同受試能力 $\theta$ 所提供之訊息量，圖中的試題訊息量的最大值=0.55，其相對的能力值 $\theta = -1$ ，表示此試題最適合受試能力 $\theta = -1$ 者施測。由於相同的試題對不同的受試者提供了不同的訊息量，所以依據試題訊息量的估計值，得以選擇最適合目前受試能力之試題。

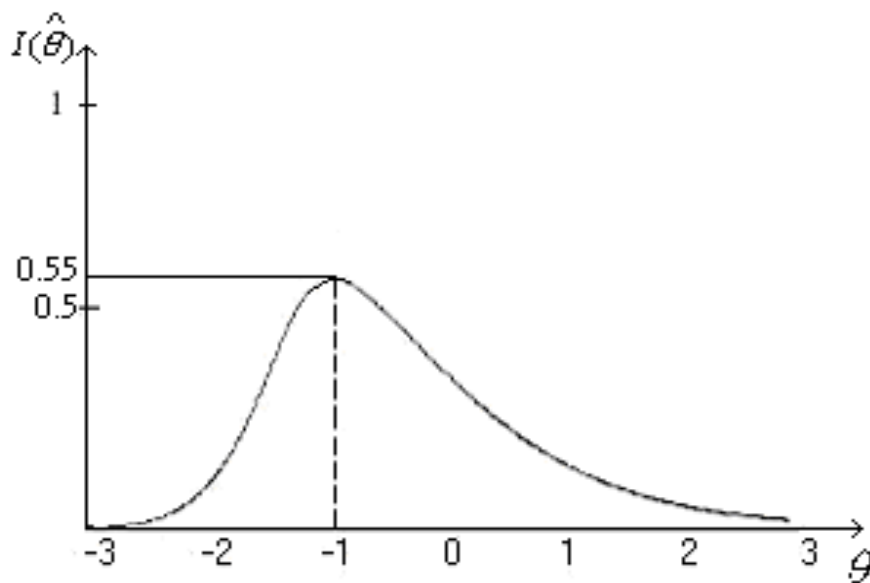


圖2-1-2 試題訊息曲線（IIC）

題目訊息函數一般是鐘形曲線，試題反應理論中不同參數的對數模式與試題訊息量的對應關係如表2-1-1（Hambleton & Swaminathan, 1985；Hambleton, Ronald K., Swaminathan, H. & Rogers, Jane., 1991）。

表2-1-1 對數模式與試題訊息量對應的關係

有最大訊息量的能力點		最大訊息量
單參數 對數模式	$b_i$	$\frac{1}{4}D^2$
雙參數 對數模式	$b_i$	$\frac{1}{4}D^2 a_i^2$
三參數 對數模式	$b_i + \frac{1}{Da_i} \ln\left(\frac{1}{2} + \frac{1}{2}\sqrt{1+8c_i}\right)$	$\frac{D^2 a_i^2}{8(1-c_i)} \left[1 - 20c_i - 8c_i + (1+8c_i)^{3/2}\right]$

### 參、試題反應理論真分數的等化

假若有兩份測驗在經由試題反應理論估計出來的能力與試題參數並不是在同一量尺上，則要先將其轉換為相同的量尺，才能作等化的動作，其實， $\theta$ 量尺最主要的轉換用途是將它轉換成真實分數量尺（true-score scale）；因為真實分數量尺的範圍是由 0 到  $n$ ， $n$  為測驗的題數，而  $\theta$  量尺的範圍卻是介於正負無窮大之間（亦即  $-\infty < \theta < \infty$ ），若將  $\theta$  量尺轉換成真實分數量尺，不僅有助於我們陳報考生的能力高低，更有助於我們解釋測驗分數和做測驗等化（test equating）（余民寧，民 83）。

由於真實分數是測驗試題答對機率之和，其公式表示如下（Hambleton, Ronald K., Swaminathan, H. & Rogers, Jane., 1991）， $\tau$  為真實分數：

$$\tau = \sum_{i=1}^n p_i(\theta) \dots\dots\dots \text{公式 6.}$$

真實分數即是能力為  $\theta$  的受試者在一堆試題上的試題特徵曲線（ICC）之和。由此看來，真實分數其實就是考生在某一測驗上的測驗特徵曲線（TCC）。真實分數可以被看成是  $\theta$  的一種非直線轉換，因為  $\theta$  與  $\tau$  間具有一種依序遞增的函數關係。真實分數的等化是指給定受試者能力  $\theta$ ，則真實



分數  $\tau_x(\theta) = \tau_y(\theta)$  視為相等，因此我們可以將真實分數的等化歸納為三個步驟：

1. 給定一  $x$  測驗上的真實分數  $\tau_x$ 。
2. 找出  $\tau_x$  所對應的  $\theta$  值。
3. 求出  $y$  測驗上對應此  $\theta$  值的真實分數  $\tau_y$ 。

另一種常用的轉換，便是將  $\tau$  轉換成內容範圍分數 (domain score) 如下 (Hambleton, Ronald K., Swaminathan, H. & Rogers, Jane., 1991)：

$$\pi = \tau/n = \sum_{i=1}^n p_i(\theta)/n \dots\dots\dots \text{公式 7.}$$

可見  $\pi$  的值介於 0 與 1 之間，如同百分比一般介於 0% 到 100% 之間。在一個參數和二個參數模式下， $\pi$  的下限值为 0；而在三個參數模式時，由於  $\theta$  趨近於  $-\infty$ ，所以  $p_i(\theta)$  趨近於最低的漸近線  $c_i$ ，故  $\pi$  的下限值为  $\sum c_i/n$ ，與之相對的  $\tau$  的下限值則為  $\sum c_i$ 。

將  $\theta$  轉換成真實分數或內容範圍分數的好處是讓新量尺的範圍介於 0 與  $n$  之間 (或 0%~100%)，在分數的解釋上變得較為容易。

所以本研究首先依據試題反應理論的定義與公式，精確的估算出受試者在各個測驗單元測得的能力值，並將測得的能力值轉換為真實分數以及等化真實分數，做出受試者個人在空間能力測驗內三個向度測驗的評比，最後再與其他受試者的得分結果做出整體性的結果評比。

## 第二節 電腦化測驗

在教學活動中，評量的實施，不只要評定學生學習結果，也要能協助學生在學習歷程中有最好的學習，甚至是要預測學生未來可能的發展方向，所以教學與評量往往是相輔相成的，不可偏廢（陳英豪、吳裕益，民89）。傳統上有關評量的實施大多採用紙筆式的測驗，對受試者而言，通常只能得到一項總結性的結果訊息，而其學習過程卻是經常被忽略的，折損了測驗有助於正確學習方式的保留與錯誤觀念修正的功能。

隨著電腦科技及測驗技術的發展，現在電腦可應用在很多有關測驗的活動，因此配合電腦來發展與運用測驗，是不可抗拒也是不可避免的趨勢（何榮桂，民86）。自從電腦出現後，應用於測驗上主要可分為兩個方向：（1）建構測驗，也就是題庫的建立（2）實施測驗，也就是利用電腦作為考試的媒介（Alessi & Trollip, 1991）。電腦由於具有快速資料處理的優點，可以促進測驗記錄、統計、分析與測驗結果的工作自動化，因此電腦化測驗就被廣為研究。

早期電腦在測驗上的應用主要是用在題庫的建立及利用電腦出題來實施測驗，再根據測驗結果分析學習者學習成效或心理反應，此種測驗雖然能讓教師節省資料鍵入的時間及負擔，並讓學生於測驗結束後立即獲得回饋，但是卻無法根據學生的個別能力來出題，也無法針對評估的結果來修正教學內容或教學策略，而在加入適性測驗的理論之後，電腦化測驗已有了重大的突破。

從過去許多電腦化測驗的研究中顯示，電腦測驗可以減少施測時間、增加測驗的安全性、有立即呈現考試結果等優點（Bugbee, 1992）。而且，Recker與Greenwood（1996）曾針對WWW環境建置的測驗系統，研究其與紙筆測驗結果的差異，結果顯示不同的測驗媒體並不影響測驗結果（何榮桂、蘇建誠，民86）。因此，將傳統的紙筆測驗電腦化有很多的優點，劉亞平（民86）

將這些優點歸納如下：

1. 更真實的題目呈現：隨著多媒體的發展，電腦不再侷限於文字的方式呈現題目，其可以圖片、聲音、影像、動畫相結合的方式呈現於受試者的眼前，能將試題表現在一個較為真實的情境，呈現的刺激具有多樣性與動態性。
2. 立即回饋的提供：電腦可以針對受試者的答題，瞬間給予反應或計算得分，受試者可以得到立即的回饋。另外，施測者與受試者可在極短的時間內取得測驗分數並得到解釋，得知測驗的結果。
3. 施測情境的有效控制：因為電腦化測驗不以人擔任施測者，能精確地控制實施和計分的程序，可將測驗標準化，避免人為情境的干擾，施測者可賦予電腦相同的動作，使受試者在相同的情境下接受測驗。
4. 可減少測驗的時間：電腦化測驗和紙筆測驗，在平均數、標準差、信度、標準誤等方面所得一樣，但是電腦化測驗所需的時間較少（溫玲玉、洪銘建，民85）。
5. 額外訊息的獲得：利用電腦可以儲存一些測驗過程方面的訊息，能夠蒐集到較豐富的測驗資料，以提供改進測驗的依據，或幫助了解傳統測驗中受試者可能的反應，這是傳統紙筆測驗不容易做到的。
6. 工作效率的提昇：電腦化測驗能自動計分和分析結果，節省人力與時間而提高工作效率。另一方面可將傳統測驗所需的行政程序（命題→打字→印刷→分卷→測驗→閱卷→計分→統計）過程簡化為（命題→測驗→統計）三個步驟，將傳統測驗予以高度科技化與自動化。
7. 適性化的測驗：可以提供適合個別差異的試題，應試者不須像傳統測驗一樣作完所有題目，因此可以縮短應試時間，減輕考試負擔。

然而電腦化測驗並非全然都是優點，其本身也有一些問題存在，李連順（民89）歸納出可能的問題如下：

1. 電腦化測驗的題型多限於選擇題、填充題，因為電腦很難判斷開放式問題的答案。
2. 在測驗實施方面，施測者必須準備足夠的電腦供受試者使用，且必須考慮電腦當機等問題。
3. 電腦測驗較不適合呈現必須閱讀文字章節的題型，因為螢幕的解析度可能增加閱讀的困難性（Mazze & Harvey, 1988）。
4. 類似數學、理化等必須計算的試題，學生因為必須在草稿紙上抄下題目再計算，而可能影響答題的速度（Ager, 1993）。

近年來網路科技的蓬勃發展也讓電腦化的測驗方式擴展了其觸角。測驗由過去的單機施測演變至區域網路的施測環境，且現今更進一步的發展出以網際網路為骨幹的遠距測驗，其目的是希望結合網際網路的優點，提供超越時空、隨選隨測、更彈性的施測環境（何榮桂，民86），因此以網際網路為建構平台的電腦輔助測試環境也漸具規模（杜榮珠、王美智，民83）。將網路運用於電腦化測驗能提供的優點包括（周文正，民87）：

1. 提升施測與批閱給分效率，讓教師能有較多的時間在教學研究上。
2. 便於教師蒐集受試反應資訊，以協助教師做測驗難度與效度分析。
3. 提供具一致性的測驗準備與試題製作，使施測更趨於標準化。
4. 有利於發展網路多媒體測驗之呈現型態。
5. 有利於適性測驗的發展。
6. 減少紙上作業，降低人工作業出錯的機率。
7. 易於反覆練習，迅速提供施測結果與正確性的回饋導引資訊。
8. 個別化評量，不受時空、地點的限制。
9. 易於建立競爭式環境，提高學習動機與成效。

此外，測驗結果及成績統計存檔在磁碟中，可以隨時調閱；測驗試題同樣可存在電腦檔案中，教師能反覆參考使用（杜榮珠、王美智，民83）。

如果以發展的階段來分，我們可以將電腦化測驗的發展分成下列四代（Bunderson, Inouye & Olsen, 1989；林鴻源，民 88）：

1. 第一代電腦化測驗：以電腦呈現傳統的測驗，將傳統的紙筆測驗搬上電腦螢幕上，或藉著讀卡機來作評分的工作，電腦在測驗領域所提供的協助著重在統計方面及測驗結果的處理上。
2. 第二代電腦化測驗：是所謂的電腦化適性測驗（CAT），它是以 IRT 為基礎，在測驗的初始，先給考生平均難度的題目或是任一題目，依據考生的表現，調整下一個題目的難度、內容或呈現時間，以最適合考生能力的題目，來測驗考生的能力，最大的優點是能夠以較少的題數測得學習者的真正能力。
3. 第三代電腦化測驗：稱之為連續評量（continuous measurement, CM），是將具備量尺化題目參數的試題安排在課程中，以持續而不干擾的方式估計學生成就軌跡或剖面的動態改變，扮演著無聲、便利的教學諮詢及學習輔導的角色。
4. 第四代電腦化測驗：稱之為智能評量（Intelligent measurement, IM），根據知識基礎及推論程序也就是專家系統或人工智慧系統，提供智慧性的計分、個體剖面的解釋及對師生的建議。

在測驗的發展過程中，基於對學習不同的解釋，晚期的測驗有著與先前的測驗不同的面貌，分為下列三點來說明：

1. 前兩代測驗通常是靜態、單一時間點及考生成就狀態的描述；後兩代測驗則更偏重個體隨著學習經驗動態改變的記錄與解釋。
2. 將以往專業、嚴謹而針對機構所需求的測驗，轉向個體或少數人的需求設計。
3. 在作業形式上，早期是受科技的發展而產生的測驗形式，技術層次較低，即使沒有電腦化的版本，紙筆測驗依然能夠發揮測驗的功能，這樣的測

驗只能夠提供對測驗方面較少的協助。唯有將電腦化測驗發展到電腦化適性測驗時，才能顯現出測驗電腦化的真正意義。

隨著電腦化測驗的演進，國內多位學者也對電腦化測驗的提出了應該往技能、態度及輔導等方面來進展（莊謙本、李進寶，民 87；吳鐵雄、陳新豐、葉千綺，民 86），所以第三代電腦化測驗及第四代電腦化測驗的發展將漸為測驗研究者所重視。

一套實用的測驗系統應該開發具有多樣化的試題呈現模式、可不依序作答、可修改已做過的答案、測驗完畢立即顯示結果、儲存測驗結果供日後查詢列印、試題難度分析、試題難度分布顯示等功能，因此電腦資訊科技在測驗上的應用，可從兩方面來觀察：一是利用它來改進及提高原有測量工具、施測方法的特性和品質，另一為創造新的測量方法，如電腦化適性測驗。至於其應用範圍有（周文欽、歐滄和、許擇基，民 84）：

1. 評分、計算與分數報告。
2. 分析試題。
3. 執行考試。
4. 運用於適性測驗與診斷性測驗。
5. 編製題庫與測驗。
6. 撰寫試題。
7. 評閱主觀題。
8. 人工智慧與應用。

Mark (1997) 亦認為有兩個技術之運用，對未來電腦化測驗的發展可能有明顯的影響：

1. 多媒體電腦技術的運用：它可使測驗藉由電腦模擬的方式，呈現更接近真實生活的問題。
2. 人工智慧的應用：它可以讓電腦會表徵測驗所欲測量的知識，與技巧的

建構受試者的狀態。

本研究是屬於第二代電腦化測驗，並架構在Internet上，除了能更有效的運用社會現有之電腦資源，讓更多的受試者能更方便的來接受測驗，提高測驗之使用率，達到超越時空、隨選隨測彈性的施測環境外，亦能提供遠距教學在學後評量上，精確評估學習者能力，建立最適當模的一個良好工具。

### 第三節 空間能力之相關理論

空間的概念可分為大空間 (large-scale space) 的環境認知能力與小空間 (small-scale space) 模型，其中大空間是指個人必須本身在環境中的運動知覺才能瞭解到整個環境的架構，而小空間則是對物件在三度空間中做心理旋轉和方向位移的操弄 (張麗芬，民78、陳采穗，民87)。

空間能力之研究工作最早起源於本世紀初，當時研究的重點在於探討學習者的個別差異。至第一次世界大戰，更將此研究領域擴展至機械性向的預測，並將空間能力視為“非文字的智力”。第二次世界大戰後，開始研究空間能力與數學間的關係，甚至研究空間能力與音樂的關係 (Smith, 1964；戴文雄，民85)。

關於空間能力的定義，有眾多的說法，即使是名稱相同，也可能代表著不同的意義。下表所示為學者們對於空間能力的定義 (梁勇能，民89；簡慶郎，民90)：

表2-3-1 學者對空間能力的定義

學 者	空 間 能 力 的 定 義
French (1951)	空間能力具有：(1) 空間因素 (spatial factor)：能夠正確察覺空間的模型並彼此互相比較的能力。(2) 空間方位 (spatial orientation) 因素：當空間模型被改變方位之後，依然可以保持不備混淆的能力。(3) 視覺化 (visualization)：能夠理解三度空間中的心像移動或是在想像中操作物體的能力。
Michael et al.(1957)	(1) 以受試者本身為參考軸，來瞭解視覺刺激物內部元素安排的本質。(2) 心理操作一個高度複雜的刺激模型。(3) 能有同感地將整個圖形轉向或是向左(右)扭轉的一個移動。
Smith(1964)	(1) 操作相同圖形的能力。(2) 察覺，保留和確認整個圖形的能力。



表2-1-1 (續)

McGee(1979)	(1) 對視覺刺激形式內部安排的理解，即不受方位改變的混淆能力。(2) 能在心理操作、旋轉或扭轉以視覺呈現的刺激物的能力。
Priby & Bordner (1985)	(1) 心理操作動態圖形化刺激之認知、保留與回憶的能力。(2) 個體對不斷改變方向之物件或圖形保持清晰不亂的能力。
Tartre (1990)	受試者心理上去重新調整自己的視野，以便和視覺呈現的物件表徵變得一致。
丁振豐 (民 83)	空間能力是對心像操弄的能力。
戴文雄 (民 87)	個體因人而異之揣想或思考三度空間以及依據圖解或圖案在心理上想像或作物體旋轉或移動及改變方向和位置的抽象能力。
洪蘭 (民 89)	能夠將東西轉化成圖形顯現在腦海裏，正確的將形狀、位置、大小比例變換出來。
康鳳梅 (民 90)	空間能力是涵蓋認知能力，超越了如記憶、複製或配對等認知能力，它含有感觀、記憶、邏輯思考以及創造性空間思考的能力。

從以上的資料中，可以發現到學者們對於空間能力解釋都不盡相同，但是，如果由空間能力的構成因素來看，則可以看出學者們較認同空間能力是「由外在刺激產生內在的心像，並且能夠操作這些心像的能力」。

空間能力的增強對人腦資訊處理之幫助有下列三點 (戴文雄，民87)：

(1) 當資訊輸入時，協助學習者編譯大量的資訊，結果更多的資訊可以儲存在暫存區並加以處理。(2) 幫助保留旋轉中的影像或加速影像的旋轉。(3) 空間能力之增強亦將改變解決問題的整體策略。故瞭解學習者的空間能力，將有助於瞭解其學習過程中處理資訊的方式，進而發展適合其處理模式的學習材料與學習系統，以提昇學習成效 (戴文雄，民87)。

相關之研究顯示個體的過去經驗會影響空間能力，或經由規劃設計之策略或方法能改進空間能力 (Wavering, 1986；Lord, 1987；McCormack, 1988；

戴文雄，民87）。

Moses（1982）認為讓學生經由實物之接觸後自描述此物，為一有效增進空間能力的方法。

Cohen（1983）研究亦顯示讓學習者從不同的角度來觀察操作中的物體，也能增加學習者的邏輯空間能力；所以提供各種實體或模型，供學習者接觸、描繪、比較大小、形狀、方向和位置都可增進學生的空間能力。

Holley & Dnserean（1984）認為經由適當之空間教學增進策略可以增進空間能力。

Billy（1988）也發現接受電腦作為教學工具的學生，其空間測驗成績比紙、筆為教學工具者來得高。

McCormack（1988）建議增強空間能力之方法有三：（1）鼓勵學生於學習中儘可能將其觀點以概念圖表示。（2）當腦力激盪時，任何觀點儘可能使用圖形表現而不要用文字表達。（3）儘可能從不同之角度去思考影像或圖形。

戴文雄（民87）認為電腦模擬更是提昇空間能力之最有效工具，因電腦具彩色、移動與旋轉、動畫及重複顯現等功能，可以快速顯現各種以前可能在現實環境中無法複製的二度或三度空間物體或狀況。

戴文雄（民83）並依McCormack對空間能力的分類，提出增進空間能力的具體方法為：（1）經由比較實物與照片或描繪之各項細節，將可增進空間感觀能力。（2）經由分解或組裝各項模型或機具設備、描繪以前所見之幻燈片、投影片或其它圖畫，將可增強空間記憶能力。（3）摺紙、展開、預想物體之斷面圖、或從物體之投影中預想物體之實體細節，將可增強邏輯空間能力。而人腦之資訊處理模式，亦值得探討（戴文雄，民87），當人腦輸入資訊時特徵區別與模式辨認乃是首務，根據學習模式，當資訊進入人腦暫存區時，已經是一種空間視覺的型式，隨之轉入長期儲存區或自腦中消

失，資訊進入長期儲存時是以網狀結構型式呈現，節點表示資訊之類別，諸如人、事、物等，而鏈結則以示節點間之關係（Lakin, et al., 1980；Stewart & Atkin, 1982）。

因此，本研究希望藉由學者對於空間能力的定義來瞭解受試者在面對空間視圖時其可能的處理資訊的思考方式，以配合網頁動態畫面呈現、視覺化的操作處理來設計最佳的評量工具。

## 第四節 網站互動式理論

### 一、互動的意義

Jonassen (1988) 認為互動是指兩個有機體的交互活動，而在電腦教學的活動中，乃指學習者與電腦間的互動關係。Price (1991) 則認為互動是人與機器之間訊息、反應與回饋的交換，以達到個別化與適應化。戴文雄 (民87) 對互動的定義則是應建立於學習者是否能夠融入所提供的環境中而成為環境的一份子，同時，系統也應能依學習者不同的需求調整互動，並提供建議。Schwier & Misanchuk (1993) 認為互動的程度是由學習者與電腦系統的品質來決定。

互動為現今多媒體的網際網路所標榜的主要功能之一，應用在提升學生學習能力的電腦化教學上更是一項不可或缺的利器；互動即是在學習歷程透過問題、反應、回饋之步驟來幫助學習者達到更高的認知與技能，甚至提昇學生的學習情境，加深學習者對學習課程的了解。

透過回饋，學習者可以明白自己在學習上所得的實際成果與預期成果之間的差距，也能改變成堅持自己對學習主題的認知度 (Mory, 1992; 趙寧, 民85)。因此回饋乃是在電腦輔助教學或電腦輔助測驗中，相當重要的一個環節。國外學者Cohen (1985) 則認為回饋在電腦教學上，即為學習者的反應後所出現的訊息。Clariana (1991) 認為回饋是藉由提供額外資訊而增進學習成效的方法。而國內學者黃清雲 (民84) 則認為回饋在電腦學習上的定義是教材內容對學生輸入訊息的反應。湯清二 (民85) 認為回饋功能主要是提昇學習效果，並提供學習回答時有益的資訊。

回饋用以指出學習結果的正確與否，且依照設計及運用程度的不同，一般歸納為 (Sales, 1988; 黃清雲, 民84; 湯清二, 民85)：

1. 無回饋：允許學生在整個學習活動中，沒有得到任何的指示與問題。在許多研究中，無回饋設計，學生學習成就較低。

2. 知識反應型之回饋：當學生回答時，僅告知學生答對或答錯了。
3. 改正型知識反應之回饋：該回饋在提出問題後，不論學生作答正確與否，均給予正確的答案，這種方式設計很簡單，效果也較佳。
4. 再試一次回饋：當學生回答錯誤時，准許學生重做一次或多次。
5. 精緻型回饋：即針對學生作答時，給予正確的解說，為什麼答錯或答對，而且准許學生重新去查看相關資料再回答，或查詢正確答案之相關知識。

回饋對智力愈高者而言，並不需要太多回饋訊息，反而是訊息愈簡單愈好，否則太多回饋訊息對其造成一種無謂的浪費。因此回饋的良劣對學習者的學習效果是會依學習者的年齡與智力而有所不同，但為兼顧大部分的學生學習進度，最好的方式，應採用精緻型回饋，因為該回饋方式可依學生自己的學習進度來自行調整。（戴文雄，民87）回饋是電腦輔助學習最常用的策略，且回饋技巧更是電腦輔助學習成效中影響最深遠的，所以回饋必須被小心的處理。

## 二、互動的層次

Hoffman & Novak (1995) 針對以電腦為仲介的學習環境，提出機器互動與人際互動二種互動形式。機械互動指是使用者對電腦進行存取動作；人際互動則是使用者彼此之間透過電腦為媒介進行溝通。Nassey & Levy (1999) 將線上媒體的互動性，分為兩個面向：內容互動與人際互動。內容互動指的是使用者對線上媒體內容所能參與的程度，人際互動則是指使用者可以在其上透過電腦中介進行交談，因為使用者主要透過人機介面設計進行瀏覽網站動作，而網站使用者可在其上進行的人際互動則有電子郵件、討論區、聊天室等，透過這些互動性的功能設計，使用者可向線上媒體進行意見回饋，使用者之間也可進行意見交流。李依倩 (1999) 則認為人際互動與人

機介面不可混為一談，人機互動的使用者的控制權是有限制的，只能在設計者所規劃的範圍內進行有限度的互動行為，但人際互動中訊息交換是互相依賴的，並且具有自由創新的人類思考模式，與人機互動中封閉系統並不相同。

Schwier & Misanchuk (1993) 根據電腦軟體性質，將互動性分為三個層次；反應式、主動式、雙向式。互動式互動意指一種被動式的反應。使用者僅能根據系統所提供的固定刺激進行有限度的反應；主動式互動強調使用者自主建構與整合資訊的活動，使用者不僅在現有的資訊架構中選擇與回應，還能主動地反應個人的思考與需求；雙向式互是最高層次的互動，強調使用者與系統之間相互因應改變，近似於人類之間的真實溝通行為，如虛擬實境和人工智慧等。

### 三、互動性的衡量層面

互動性是一個和訊息傳播技術息息相關的多層次概念，因此互動性可從不同的層面來以衡量，包括傳播科技、商業網站、電腦教學等層面，蔡淑如（民89）的研究論文中，歸納出幾個衡量互動性的層次，分就媒體與使用者兩個主要層次加以統整，做為考衡互動性的依據。

（一）就系統本身而言，有下列幾個衡量的層次：

1. 系統提供選擇的複雜性：指系統所能提供給使用者的選擇有多少，訊息來源提供不同的系統也會以不同的行為去取得資訊，如此一來提供使用者的選擇性越多，越有利於使用者進行回饋，也就越能滿足使用者不同的傳播需求。因此，在選擇的層次中，包括資訊內容的選擇性，也包括使用者行為的選擇性（黃含綿，民85）。
2. 系統回應能力的程度：就系統本身而言，當使用者提供回饋時，系統的回應能力越強，越能滿足使用者的傳播需求，互動性也就越高。Heeter（1989）認為，系統回應能力的程度，應取決於系統本身所能模擬人類

對話的程度，系統能如人類般進行回應時，即達到最高層次的模擬成熟境界，當系統與使用者之間能如人類般互相回應時，彼此之間傳送的訊息就能互相因應且具有相關性，如Selnow（1988）指出，訊息必須是針對明確的對象而發，因此訊息來源若能因應不同的交談者而傳送不同的訊息，則回應能力越強。

3. 系統回應速度：指系統能接收與處理的速度。Borsook & Higginbotham（1991）提到，立即回應是互動性的主要成分，因此，最高的互動性傳播應能達到即時接收、處理、回應的速度。
4. 系統監測能力的程度：隨著新科技的發展，系統可以潛在的監視系統持續性地瞭解使用者的使用情形，進而根據使用者的需求，提供更符合使用者興趣的資訊內容（Heeter, 1989）。Ha & James（1998）提出類似資訊蒐集概念，藉由使用者註冊、網站到訪人數計數器、cookies等機制，來蒐集使用者的相關資訊，瞭解使用者的使用情形，以利系統依據使用者的需求進行回應。
5. 促進人際溝通能力：Heeter（1989）指出，新傳播科技如視訊會議可以讓使用者之間進行，面對面地溝通對話，亦可隔一段時間彼此非同步地傳送訊息，促進人際溝通能力概念。Steuer（1992）提出相符相聯概念，相符相聯是指使用者在中介環境的行為之間，存有最高度的關聯性，使用者可以最自然的方式在此中介環境中進行活動，便是最理想的互動。Rogers（1986）所提出的近似個人參與對話的能力亦表達了理想的中介式傳播過程，若能趨近人際傳播過程則越理想。
6. 樂趣性：Ha & James（1998）認為，使用者藉由點選動作與網站互動過程中充滿了樂趣，藉由一些線上遊戲或是激勵好奇心機制的設立，使用者可在遊玩的情境中，滿足與自我溝通的需求，並從中獲得成就感，因此線上遊戲與問題機制都可視為是具有互動性的設計，此處所指的線上

遊戲是以遊戲規則為基礎，參與者依據個人表現和技巧的好壞，以獲得獎品或高分，而激勵好奇心機制則是指能夠吸引使用者參與其中的機制，問答題即是其中的一種形式，例如網站可以設計一些問題，讓使用者來填答，增加參與感。

(二) 就使用者而言，有下列幾個衡量的變項：

1. 使用者付出的努力的程度：指系統中的使用者為了取得資訊必須付出的努力程度 (Heeter, 1989)。當資訊選擇性增多的情況下，使用者必須主動去加以判斷並取得最適用於自己的，使用者參與度與自主性較高，互動性也就較高，此衡量面向與前述系統提供選擇的複雜性是相關的概念，系統提供選擇的複雜性越高時，使用者越能從中選擇出最適合自己的資訊，以利進行回饋。
2. 使用者非線性的資訊處理：當使用者有任何問題時，可以隨時反應，暫時中斷資訊的線性流通過程，也就是所謂的非線性資訊處理，也使得互動的雙方可以視互動對象的不同，採取不同的因應對話方式，亦即所謂的適應性 (Borsook & Higgin botham, 1991)。
3. 資訊增加的容易度：隨著新科技的發展，使用者的地位改變，不再只是一個接受訊息的被動者，而可以成為向媒介系統主動提供資訊的來源，化被動為主動，使得傳播者和接受者之間的界線模糊，使用者與系統的互動性更為顯著 (Heeter, 1989)，亦即黃含綿 (民85) 所言互動性指參與者在傳播過程中擁有控制對談，並且交換角色的程度。播播訊息內容越容易修改，能改變得範圍越大，則雙方所能得到彼此的回饋將越多，增加更多互動的機會。

#### 四、互動性的設計原則

Kristof & Satran (1995) 認為，就電腦產品設計而言，互動界面設計的



意涵主要是以使用者為中心的設計，使用者擁有相當的控制權，包括使用者可以決定瀏覽的速度、方向以及想要看或不想看的內容，因此如何設計一個好的互動性介面，必須考慮使用者的經驗。

張大鈞（民89）綜合Kristof & Satran（1995）及Mok（1996）對互動性介面設計提出的原則，以及全球資訊網的特性，統整網站互動性設計的六個原則：

1. 指引方向與導航：網站互動性設計必須要能夠指引使用者明確的方向與位置，例如首頁提供足夠的資訊，清楚告訴使用者這是哪裡？以及將會經驗到什麼，並藉由簡單的圖形或標題的呈現加以指引，對於複雜的內容則採用層級方式，讓使用者自行選擇。導航的目的在創造一個幫助使用者知道現在位置、目的地以及到目的地的方法。一個好的導引方式必須讓使用者能走最少的路徑即能到達目的地，必須切實掌握三次點選原則，所謂三次點選原則，是指使用者進行瀏覽任務時，最好能在三個視窗畫面之內即可到達目的地，否則使用者將很容易迷失在其中。且為了避免使用者在瀏覽進行中迷失方向，不同的網頁必須保有原來的特色或要件，如固定框架的設計。
2. 圖像與暗喻的使用：圖像的功能在於以圖形方式呈現網站內容，讓使用者更為方便，而暗喻則是圖像的一種，藉由其與真實世界有意義的關聯，讓介面的使用更為容易。
3. 超文字特性：全球資訊網以超文字為主要的架構，使用者可任意選擇節點與鏈結進行非線性的閱讀，由自身來決定瀏覽的順序，同樣屬於非線性的資訊處理過程。在此過程中，使用者可以主動選擇自己所要的資訊，以及閱讀的順序，具有更大的自主權和主動性。
4. 系統回應：系統回應指的是介面本身對於使用者所執行的每一動作，是否能給予明確的回應，讓使用者可以知道每個行動會產生何種效果。系

統回應可分為三種層次：

- (1) 針對使用者執行一個行動或指令，系統透過介面都必須有所回應。
- (2) 網站對使用者的回應，必須以使用者可以預測到的方式來加以回應，亦即互動介面設計必須有一致性，讓使用者知道其行動後會出現什麼樣的反應結果。
- (3) 當使用者指令錯誤時，系統必須能夠避免錯誤，自動糾正。例如，可出現訊息方塊，告知使用者錯誤所在並加以指正。

5. 可用性：可用性包含下列五個因素。

- (1) 學習能力：網站互動介面設計必須簡單明瞭，讓使用者能夠很快學習上手。
- (2) 效率：系統的功能與介面的設計必須是顯而易見且簡單容易使用，讓使用者可以有效率地加以使用，以減少使用者搜尋目標與學習適應的時間。
- (3) 記憶力：系統功能必須容易被使用者所記憶，經過一段時間再回來使用時，才不需重頭學習。
- (4) 錯誤率：系統必須減少錯誤率，使用者在使用時才能減少錯誤發生。
- (5) 滿意度：系統功能的設計必須讓使用者覺得喜歡，增加使用滿意度。

6. 功能性：提供多樣性的功能設計，並讓使用者知道如何執行功能，而且這些功能必須是容易學習的，依賴直覺即可加以使用，倘若網站提供的功能越多，網站的吸引力也就跟著提高，當瀏覽人數眾多時，亦會激發網站功能的提升或擴充，而使用者就能有愈多的選擇。

互動式理論已是一相當完整的理論，透過該理論瞭解在規劃互動式系統需衡量的層面、與互動性的設計原則，以利本研究規劃出系統所應具備的功能，如圖3-1-1所示。本研究為了達到受試者最後結果評比的公正性與準確

性，避免受試者有記憶答案的可能，將採用互動理論之知識反應型回饋的方式，建立本系統的測驗機制。所以結合IRT 與互動式理論，將可提供使用者一個有效的互動式測驗環境。

## 第五節 系統開發程序之相關理論

一般資訊系統的建立是必須耗費大量的人力、時間與物力，且要有專業系統分析與設計人才，因為在系統的建立過程之中，分析與設計的前置工作若沒有確實，可能導致整個系統的失敗，所以必須存在一個方法，作為系統發展的依據（許元，民87）。發展資訊系統由起始到完成的步驟與歷程，一般稱作為系統發展生命週期（system development life cycle, SDLC），由於每個資訊系統皆有其不同的特性，因此有關係統生命週期的劃分與步驟，並沒有固定的標準模式。底下就瀑布式系統發展法、雛型式系統發展法、重覆加強式系統發展法分別探討如下。

### 一、瀑布式系統發展法

它為1970年為Royce所提出，此發展法以系統層次為起點，經由分析、設計、程式撰寫、測試及維護等階段進行軟體發展程序，請參考圖2-5-1，瀑布式系統發展法為最古老而且至今仍為使用最廣泛的方法（陳世安，民88）。其各階段的工作重點如下所述（許元，民87；張大鈞，民90）：

1. 初步分析：進行蒐集工作及定義系統任務的目標，找出所有可行方案，進行可行性研究，找出最佳方案，並為最佳方案做成本效益分析。
2. 需求分析：是針對初步分析的需求做更深入的研究，一般而言，需求分析的重點在於瞭解系統需要什麼？通常必須確認系統的功能需求、資料需求、人機界面需求、性能需求、軟硬體需求、安全及控制需求及系統配置需求等。
3. 軟體設計：主要的目的在於需求分析階段所確認的各種功能需求，進一步擬出一份解決這個問題的方案。一個系統由模組所組成，模組本身提供了某項特殊功能，藉由模組間的控制關係，來完成使用者所要求的功能。

4. 程式撰寫：主要工作為選擇合適的程式語言，利用所挑選的程式語言轉換成真正的程式碼。
5. 測試：測試工作主要針對前述的程式碼，進行一連串的驗證及證實的工作，確認是否符合系統之需求。
6. 維護：維護即當系統正式運作後，可能由於錯誤、功能修改、軟硬體의更新，而對於原有的程式及文件進行修改。它可能進行更正性維護、適應性維護、完善性維護及預防性維護等。

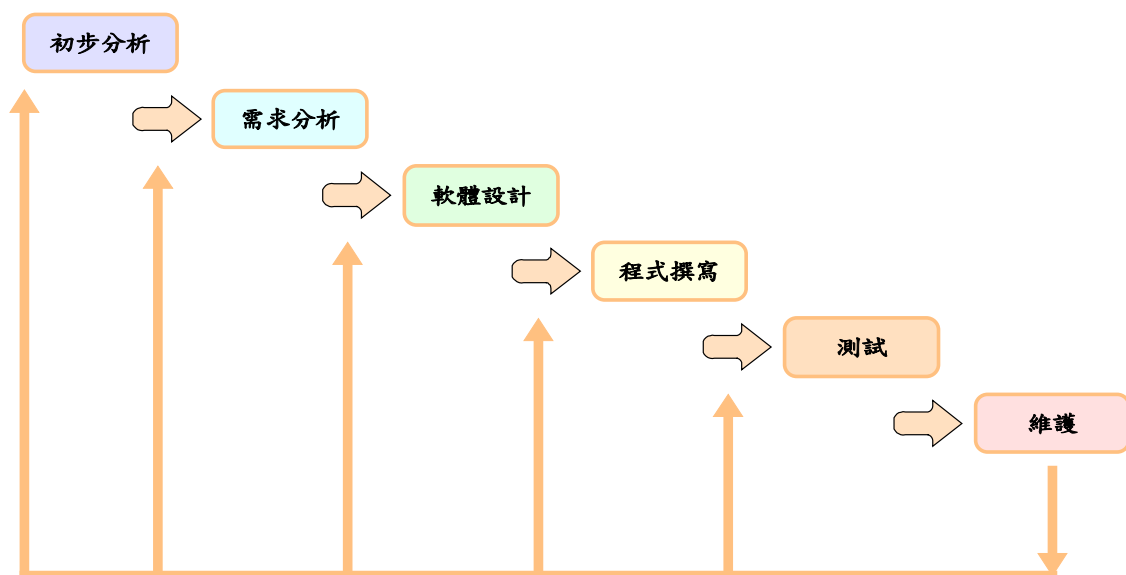


圖2-5-1 瀑布式系統發展法

## 二、雛型式系統發展法

雛型式系統發展法於1984年由Boar首先提出，如圖2-5-2所示，此種方法按照使用者的初步需求，儘速完成一個雛型系統供使用者使用，之後再不斷地根據使用者的意見進行系統的修改至使用者滿意為止之軟體系統開發方法（Boar, 1984；吳琮璠、謝清佳，民81）。由於在新系統發展時，使用者通常只能提出部分的需求，所以若採用雛型法則可以很快來瞭解使用者的需求，並能隨時與系統發展者做適當的溝通，此種方法亦可以用來辨別使

用者的動態功能需求，或者將工作雛型評估修正成最終的軟體產品。

雛型系統發展法過程可分為需求分析、快速規劃、快速製作、建立雛型、評估修正及完成系統等階段，各階段的工作重點如下所述（張大鈞，民90）：

1. 需求搜集：開發者與使用者需一起討論並定義軟體的目標及所有的功能需求，同時列出必須做深入定義的地方。
2. 快速設計：著重在使用者可看到的介面，如輸入輸出之畫面及報表等，開發者須使用現有的程式模版或應用相關工具，如報表及格式產生器來達成快速設計的目的。
3. 建構一套軟體雛型：完成一可執行的軟體雛型產品。
4. 評估需求：透過使用者的實際操作與需求評估，提出需求的調整。
5. 循環：透過上述的程序週而復始的執行，直到滿足使用者之需求，才繼續進行後續分析，設計...等階段，因此開發者在此階段即能充分了解那些功能是必須設計。

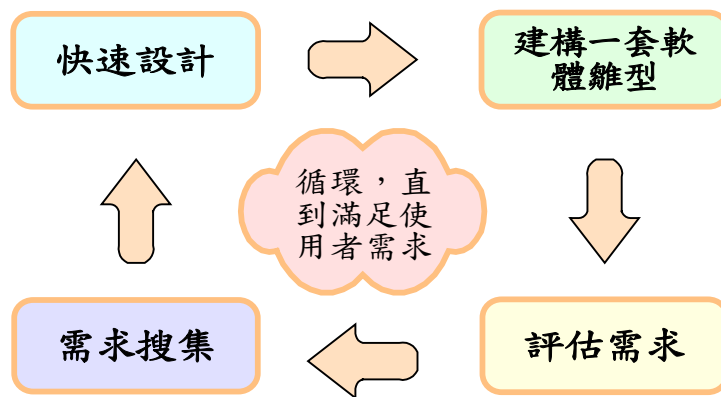


圖2-5-2 雛型式系統發展法

### 三、重覆加強式系統發展法

重覆加強式系統發展法又被稱遞增式系統發展法，它是採用步階增量的方式來發展，額外的子系統功能是在每次增加的過程中被加入至軟體系統，

並重新進行設計、製作及雛型修正的工作，如此重覆增量直到所有子系統功能被加入系統為止，如圖2-5-3所示。

其工作步驟有三項：1.建立處理控制清單、2.增量步驟、3.判斷處理控制清單是否空白。根據許元（民87）的解釋：首先必需建立處理控制清單，它可定義整個系統必須具備的所有應用系統功能，也可以瞭解到距離完成的系統還有多遠，且由增量步驟中自處理控制清單尾端取一項分析工作，快速建立雛型，由學習者的評估結果獲得修正的意見，它包含設計、製作及分析，其中設計部份可自選設計方法，製作方面則需包括程式的撰寫及測試，分析則是經由使用者線上使用，並提供意見，經過評估後，將所需的修正，放進處理清單的尾端。最後判斷處理控制清單是否空白，若是則整個系統完成，若為否則重覆進行增量步驟。透過此種過程重覆不斷地實行之後，即可將我們所需的功能附加於該系統中。

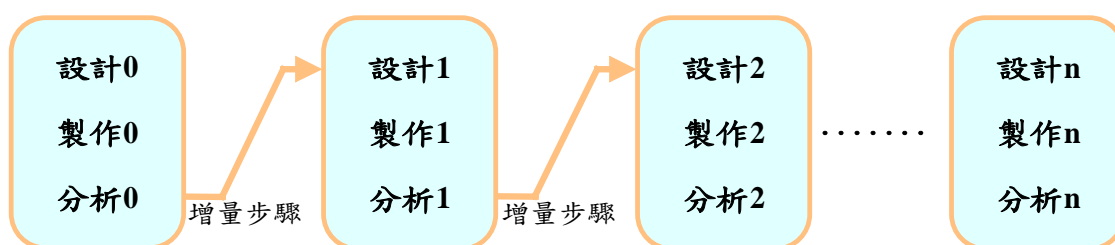


圖2-5-3 重覆加強式系統發展法

表2-5-1 系統發展方法的比較

系統發展方法	花費時間	優點	缺點
瀑布式	較長	循序漸進，容易掌握整個系統架構。	若有錯誤，可能要花相當多的時間來修正。
雛型式	較短	快速地製做出雛型，再依使用者需求，不斷地修正，避免大幅錯誤的修改。	需要不斷測試，所花費的測試人力相對增加。
重覆加強式	較短	可以瞭解距離系統完成還有多遠。	需要不斷測試，所花費的測試人力相對增加。

從表2-5-1 的比較中，可以知道三種系統發展方法各有其優缺點。而本研究限於時間的問題，無法採用瀑布式的系統發展法，且在人力及希望系統在完成後能符合使用者需求的考量下，本研究的系統開發方法，將依雛型式系統發展法，來發展「網路適性測驗系統」。