第二章　文獻探討

本部分主要之目的在探討與本研究相關的理論與研究，共分爲五節。第一節針對項目反應理論之二元計分Rasch模式進行回顧，第二節瞭解結合潛在類別分析和Rasch模式的混合Rasch模式，第三、四節瞭解在中學階段之數學內容知識及數學認知能力之各大型評量機構之界定，第五節則介紹相關數學評量實證結果。

第一節　項目反應理論---二元計分Rasch模式(RM)

測驗理論(test theory)是一種解釋測驗資料間實證關係(empirical relationships)的有系統理論學說，測驗理論學者將其分爲兩部分：一爲古典測驗理論(classical test theory)，主要是以真分數模式(true score model)(Gullikson, 1987；Lord & Novick, 1968)爲骨幹；另一爲當代測驗理論(modern test theory)，主要是以項目反應理論(item response theory, IRT)(Hambleton & Swaminathan, 1985; Lord, 1980)、潛在特質理論(latent trait theory, LTT)爲相關名詞的統稱。

自1950年代，項目反應理論成爲教育的和心理學的測量之心理論基礎(Embretson & Reise, 2000)。在IRT的主要問題是，試題的正確反應機率與受試者心智特徵，如智力、學業熟練、態度和試題特質的關係，如試題難度、試題鑑別度等。

本節即介紹項目反應理論的定義、假定、特性，二元計分Rasch模式之定義、試題特徵曲線、試題參數及其估計方法。

一、項目反應理論之意義

Lord & Novick(1968)認爲項目反應理論是：在定義人類特質後，個體的行爲可以被證明其真實的程度，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。

項目反應理論包含一群數學模式，這些模式是論述觀察變項與其所建構之潛在假設特質間之關係。項目反應理論由以下要素構成：

(一)必須是呈現給個體的一個或一組刺激變項，例如：能力測驗、成就測驗、人格或態度量表的問題。

(二)呈現給個人之刺激，可引發個人特定反應以觀察和紀錄。

(三)假設觀察之反應和潛在特質模式之間的關係，可藉由數學方程式描述。

因此項目反應理論的核心爲：以方程式表示人對於刺激(如試題)所觀察之反應，及視潛在特質水準爲刺激變項之函數關係(Weiss, 1983)。

Hambleton & Swaminathan(1985)認爲項目反應理論是：在測驗情境下，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，又稱爲能力分數。並使用能力分數去預測或解釋項目或測驗表現。

二、項目反應的假定

項目反應理論建立在兩個基本假定上：

(一)單向度

單向度(unidimensionality)是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)認爲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。Crocker & Algina(1986)認爲單向度及各項目間爲統計相依(statistical dependence)，即對整體受試者，項目間互相關連才存在單一特質。

(二)局部獨立性

局部獨立性(local independence)是指當影響測驗表現的能力固定不變時，對某一受試者之能力而言爲統計獨立(statistical independence)。即各項目之間無相關存

在，也就是一個試題不能提供另一個試題線索(Crocker & Algina, 1986)。

三、項目反應理論的特性

(一)不變性(invariance)

1.項目參數不變性

項目參數不變性(invariance of item parameter)是指難度、鑑別度、猜測度三種項目參數，不因群體的不同而改變，也不因性別、種族、城鄉而改變；最重要的是它與單向度的假定互爲關連(如項目參數因次群體之不同而改變，表示同具某能力的受試者在不同團的答對率也不同，此違反單向度的假設) 。

2.能力參數不變性

能力參數不變性(invariance of ability parameter)是指受試者能力估計不因試題不同而改變。不同試題之樣本在項目參數上雖有不同，但若測同一種能力或特質時，其受試者能力的關係不會改變，改變的是不同樣本所使用的原點與單位的不同。

(二)未定性

在項目反應理論中，答對率$P(\theta)$爲$a(\theta - b)$之函數，a 爲鑑別度參數、b 爲難度參數、θ爲能力值，若在θ及b 各加一個常數，則$a(\theta - b)$不變，表示若能力與難度同步，能力量尺的原點(origin)是可以任定的；若在θ及b各乘一常數，a 除一常數，則$a(\theta - b)$也不變，表示能力量尺的單位也可任定。此性質即Lord(1980)所指的未定性(indeterminacy)。

四、項目反應理論之試題特徵曲線、項目參數及受試者能力估計法-----二元計分 Rasch模式

受試者表現情形與潛在特質之間關係，可經數學模式運算求得，其數學模式

稱爲試題特徵函數(item characteristic function, ICF)，在Rasch模式中只採用難度參數，以下就Rasch模式之定義、試題特徵曲線、項目參數及估計方法分別說明。

(一)二元計分Rasch模式

令$P(X_{vi})$表示受試者v在試題i的成功反應機率，二元計分的Rasch模式之主要概念是將$P(X_{vi})$分爲試題參數(難度)和受試者參數(能力)的線性組合，由於觀察變項只爲0和1，因而並非只以機率的方式表示，而以此機率的logit表示：

$$\ln\left(\frac{P_{vi}}{1-P_{vi}}\right)=\tau_v+\sigma_i$$

其中：τ_v是受試者能力

σ_i是試題參數

將上式轉換爲一般常知的Rasch模式的反應機率，可得：

$$P(X_{vi})=\frac{\exp(\tau_v+\sigma_i)}{1+\exp(\tau_v+\sigma_i)}$$

在潛在變項和反應之間的關係，通常以試題特徵曲線(item characteristic curve, ICC) (Lord, 1980)表示，如圖2-1：

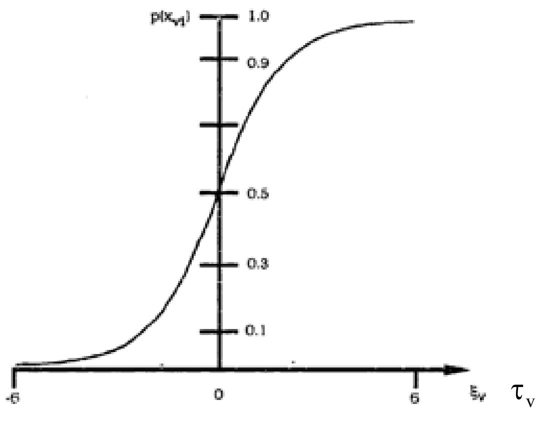

圖2-1：二元計分Rasch模式之試題特徵曲線

(二)項目參數

在Rasch模式中，其項目參數爲試題難度參數，會隨著不同試題呈現不同的難度。難度參數就是在能力量尺上的落點，落點在越左側，表示試題越容易；反之，試題越難。

(三)受試者能力估計

在Rasch模式裡，正確反應的機率端賴兩種因素：一爲受試者的能力參數、另一爲試題參數。在試題參數估計方面，項目反應理論將試題視爲測量受試能力的基本單位，並以試題參數來描述試題特徵，試題參數資料配合受試反應組型(response patterns)之函數運算後，即可推估出受試者之能力。試題參數的取得方式爲經傳統施測過程、收集受試者的反應組型、進行參數推估。若受試者能力爲已知時，則可使用最大概似估計法(maximum likelihood estimation, MLE)及貝氏估計法(Bayesian parameter estimation)(Owen, 1971)，直接估計試題的參數；當受試能力及試題參數皆未知時，則可使用聯合最大概似估計法(joint maximum likelihood estimation, JMLE)、邊際最大概似估計法(marginal maximum likelihood estimation, MMLE)、最大後驗估計(Bayesian model or maximum a posterior Estimation, MAP)、期望後驗估計(Bayesian mean or expected a posteriori estimation, EAP)等方法對受試能力及試題參數進行同時估計(Baker, 1992; Hambleton & Swaminathan, 1985)。

而在多種估計受試者能力之方法中，以最大概似估計法(maximum likelihood mstimation, MLE)及貝氏估計法(Bayesian procedure)，最爲研究者所常用(Lord, 1980)。

項目反應理論主要是探討測驗試題之特性，以試題來測量與個人屬性之間的機率關係。測驗試題包括多種方式，如能力測驗、態度測驗等；而個人屬性則例如精熟方面、態度建構等。本部份主要是從項目反應理論之定義、假定、特性瞭解其在能力測驗方面的分析，並針對Rasch模式說明其模式、項目參數及受試者能

力之估計方法。

第二節　潛在類別分析與混合**Rasch**模式

在心理學的研究中，我們常想要瞭解一些無法像物理測量一樣直接可測量的變項，例如：能力、態度、興趣等。我們經由蒐集及分析外在的觀察變項(observed variable)，以瞭解內在不可直接測量的變項，即潛在變項(latent variable)，觀察變項和潛在變項有時是連續變項、有時是類別變項，因而形成各種不同的分析，如表2-1(Cox ,Orme & Cuddeback, 2003)：

表2-1：不同變項型式之分析模式摘要表

變項型式 / 分析模式	觀察變項	潛在變項
驗證性和探索性因素分析	連續變項 或 類別變項	連續變項
潛在剖面分析	連續變項	類別變項
潛在特質分析(包含IRT)	類別變項	連續變項
潛在類別分析(LCA)	類別變項	類別變項

由表2-1，可明顯的看到項目反應理論與潛在類別分析最大之不同，項目反應理論在潛在變項的部分屬於連續變項，如受試者能力值；而潛在類別分析在潛在變項部分則屬於類別變項。在瞭解項目反應理論與潛在類別分析之主要差異後，本部分將介紹Rost(1990)提出之結合項目反應理論及潛在類別分析模式，此模式修正前兩種模式的缺點，並保留其優點，稱為混合Rasch模式，由於混合Rasch模式是潛在類別分析的一種擴充模式，以下將先介紹潛在類別分析，再介紹混合Rasch模式。

一、潛在類別模式

潛在類別分析是一種多變量類別資料的統計方法，主要是發現有關潛在類別

的次型式，潛在類別分析在醫學上應用廣泛，像是界定疾病的亞型或診斷的次類別，在其他領域的應用，則包含市場研究、調查研究、社會學、心理學和教育(Uebersax, 2003)。

潛在類別分析首先由Lazarsfeld(1950)提出，使用在基於二元變項的分析模式，大約在20年後，Goodman(1974)藉由發展運算法，以獲得模式的最大概似估計而使得LCA在實務上更爲適用，他同時提出針對多元觀察變項和多重潛在變項的擴充模式。在同一時期，Haberman(1979)指出針對有細格遺失值的列聯表，在潛在類別模式(latent class model) 和線性對數模式(log-linear model)之間的連結，自此很多古典的潛在類別模式的擴充模式陸續提出，例如模式包含共變數、局部相依性、次序變項、多個潛在變項和重複測量。Hagenaars(1990)提出對於間斷性潛在變項的類別資料分析，並由Vermunt(1979)擴充此模式(Vermunt & Magidson, 2003)。

Lazarsfeld(1950)所提出之潛在類別分析之原始概念如下：

在潛在類別分析中，母群體被視爲可區分爲若干個潛在類別，每一個潛在類別中，試題之間的答案是無關的，同時，同一試題在不同類別之答對率也不同，因此對於母群體而言，2 題或多題試題之答案將呈現有關。例子如下：

第一組試題反應

題 1 \ 題 2	答對	答錯	總和
答對	12	3	15
答錯	4	1	5
總和	16	4	20

+

第二組試題反應

題 1 \ 題 2	答對	答錯	總和
答對	1	5	6
答錯	3	15	18
總和	4	20	24

=

題 1 \ 題 2	答對	答錯	總和
答對	13	8	21
答錯	7	16	23
總和	20	24	44

在等式左側的兩組試題反應是無關的：就題 1 而言，不論題 2 答對與否，在第一組試題反應中，題 1 的答對率爲 $\frac{15}{20}=\frac{3}{4}$、答錯率爲 $\frac{5}{20}=\frac{1}{4}$；而在第二組試題

反應中，題 1 的答對率為$\frac{6}{24}=\frac{1}{4}$、答錯率為$\frac{18}{24}=\frac{3}{4}$。就題 2 而言，不論題 1 答對與否，在第一組試題反應中，題 1 的答對率為$\frac{16}{20}=\frac{4}{5}$、答錯率為$\frac{4}{20}=\frac{1}{5}$；而在第二組試題反應中，題 1 的答對率為$\frac{4}{24}=\frac{1}{6}$、答錯率為$\frac{20}{24}=\frac{5}{6}$。

若將等式左側的兩組試題反應就各細格分別相加，則可得到等式右側之試題反應，此試題反應中，呈現受試者在某一試題有較高的答對率時，在另一試題也會有相同的表現；相反的，我們可以由等式右側的試題反應，去尋求方法找出等式左側的各試題反應。

若要更清楚的了解潛在類別分析的試題反應組型機率，我們可令 X 表示潛在類別，Y_ℓ表示 L 個觀察變項之一個反應，其中$1 \le \ell \le L$，令 C 表示潛在類別的數目，D_ℓ表示Y_ℓ的水準數，向量 Y 和 y 視為受試者之反應組型。潛在類別分析的基本概念是獲得受試者反應組型之機率，此機率可以下式表示：

$$P(Y=y)=\sum_{x=1}^{C} P(X=x)P(Y=y \mid X=x)$$

其中$P(X=x)$是指受試者屬於某一潛在類別之比率，而$\sum_{x=1}^{C} P(X=x)=1$；

$P(Y=y \mid X=x)$是在潛在類別 X 中受試者成功答對的機率。

在潛在類別模式中其重要的假設為「局部獨立(local independence)」。一份測驗中的各試題，由於共同測量一個特質，因此彼此之間必定有相依的關聯，這就是一般所熟知的內部一致性或共同因素，若我們將共同特質的變異量部分去除，則會發現其實試題本身扣除了共同變異量後，彼此是獨立的，因此潛在類別分析所設定的前提是：在同一類別中，試題的反應互相獨立，因為共同變異的部分已歸屬於潛在類別(吳毓瑩、林原宏，民 85)。而觀察變項 L 可表示為：

$$P(Y=y \mid X=x)=\prod_{\ell=1}^{L} P(Y_\ell = y_\ell \mid X=x)$$

受試者之答對率可表示爲：

$$P(Y=y)=\sum_{x=1}^{C} P(X=x)\prod_{\ell=1}^{L} P(Y_\ell = y_\ell \mid X=x)$$

進行潛在類別分析時，首先根據受試者作答反應組型進行歸類，計算各潛在類別所佔全體受試者的比例(各潛在類別之比例爲：歸類到某一潛在類別的受試者數目，除以全體受試者數目)；其次類似於群集分析，潛在類別分析的目的之一爲，分派受試者至各潛在類別，此過程可藉由貝式規則(Bayes rules)獲得，式子如下：

$$P(X=x \mid Y=y)=\frac{p(X=x)P(Y=y \mid X=x)}{P(Y=y)}$$

而分派受試者至各潛在類別的規則，最常使用的是，根據貝式規則所求出受試者可能屬於某一潛在類別的最大事後機率 $P(X=x \mid Y=y)$ 進行分派。

由上述文獻可知，潛在類別分析的分析方法之基本原理爲：使用受試者的反應組型，對於觀察的類別變項進行分析，以界定潛在類別的數目，在同一個潛在特質之中，試題反應必須彼此局部獨立；因而可依照受試者的作答反應組型及模式與資料之適配度，根據事後機率、分派受試者至某一個潛在類別，尋找適合的組數與結構。若將潛在別分析應用在教育上，我們可分析所獲得估計之參數值，並可依受試者在各潛在類別之試題的答對率及試題特性，做認知結構上的分析，以對學生提供適當的協助。

二、混合Rasch模式之理論

混合Rasch模式結合了潛在類別分析與Rasch模式：在潛在類別分析部分，其分析的基本原理與前述所提之概念相同，而在Rasch模式部分，則擁有在不同的潛在類別，存在不同的試題參數之優點。本部分將省略混合Rasch模式中之潛在類別分

析之理論，只就混合Rasch模式理論進行介紹。

Rost(1990)提出結合Rasch模式和潛在類別分析的混合Rasch模式，以解決這兩種分析取向之缺點，並保留各自擁有的優點。Rasch模式描述一個潛在類別中所有人的反應行爲，但是在不同的潛在類別則存在不同的試題參數，由於類別事先是未知的，此種模式是以探索的方式來界定次群體，受試者所屬的次群體皆符合Rasch量尺。

混合Rasch模式結合了項目反應理論與潛在類別分析，模式的項目反應理論部分是屬於二元計分Rasch模式，只有一個特質水準描述受試者，其特質水準的意義是依據受試者所屬的類別而定；另外，在混合Rasch模式中，試題難度在不同類別之量尺是不能互相比較的(Embretson & Reise, 2000)。

以下介紹混合Rasch模式之理論：

令P_{vig}表示受試者v屬於某一潛在類別g中答對試題i的機率其二元Rasch模式之反應機率爲：

$$P_{vig} = \frac{\exp(\tau_{vg} + \sigma_{ig})}{1 + \exp(\tau_{vg} + \sigma_{ig})}$$

τ_{vg}是受試者的能力

σ_{ig}是試題參數，對於每一個潛在類別$\sum_i \sigma_{ig} = 0$

由於潛在類別的基本假設爲：各潛在類別之間是互斥的。因此受試者v答對某一試題之機率爲：

$$P_{vi} = \sum_g \pi_g P_{vig} = \sum_g \pi_g \frac{\exp(\tau_{vg} + \sigma_{ig})}{1+\exp(\tau_{vg} + \sigma_{ig})}$$

P_{vi} 是非條件的反應機率

π_g 是類別參數，即類別所佔之比率，其限制爲 $0 < \pi_g < 1$ 和 $\sum_g \pi_g = 1$

上述方程式並不能定義完整的模式，因爲對於受試者參數 τ_{vg} 並未明確說明，爲了要導出受試者反應機率的概式函數(likelihood function)，可令向量 $Y=(y_1,y_2,y_3,\ldots)$ 表示其反應組型機率P(Y)，其中 $y_i=0$ 或1，在此假設下，若存在若干潛在類別g，其反應組型機率可視爲 $P(Y) = \sum_g \pi_g P(Y \mid g)$。

另外，在Rasch模式下，答對反應的次數(即分數 $r = \sum_i x_i$)是 τ 的充分統計量，因此，所有具相同分數受試者，會得到相同的 τ 值，因此反應組型機率可視爲 P(Y|g)=P(Y|g,r)P(|g)。

在混合Rasch模式的估計部份，採用EM algorithm(Dempster, Laird & Rubin, 1977)。迭代(iterative)程序的E步驟(e step)中，每一個潛在類別的期望組型次數，是基於觀察次數和模式參數的初步估計，這些類別的組型次數，是藉由對觀察次數的比例加權而獲得；而在M步驟(m step)中，模式參數的最大概似估計，是基於類別中期望組型次數而求得，對於每一潛在類別試題參數和分數參數是分別被估計的，必須使得在一個潛在類別g中，資料的對數概似函數的值最大。

由上述模式，我們可以瞭解混合Rasch模式可解釋潛在類別中所有人的反應行爲，但是不同的試題參數之組合，提供了不同的潛在類別。

在本研究中，研究者將此模型應用於成就測驗結果的分析，一方面可以瞭解受試者的不同組別之存在，另一方面也可瞭解其不同組別中受試者的反應行爲之特徵，在混合Rasch模式針對基本學力測驗數學科學生作答反應進行分析的情形

下，可了解國中生在數學內容知識與數學認知能力是否存在若干組別、各組別間在這兩方面之結構與程度上的差異爲何，以及在學生在數學內容知識所表現之優劣部分。

第三節　數學內容知識

本部份將回顧國內外各大型評量測驗、以及教育相關單位所提出之數學內容知識的理論。有關數學內容知識之理論，了解(NAEP, 2005)、TIMSS(2003)兩個大型評量測驗所界定之數學內容知識、美國數學教師協會(National Council of Teachers of Mathematics, NCTM, 2000)所界定之數學內容知識，以及國內教育部所提出之九年一貫課程(2003)在數學內容知識的看法。以下針對上述大型評量測驗之及相關教育單位對數學內容知識之定義進行回顧，分述如下：

一、NAEP所界定之數學內容知識

2005年NAEP數學架構爲數學內容知識與認知能力，在數學內容知識部分，以五個領域來描述，此區分並非將數學分爲零散的元素，而是提供一個可幫助分類的基模，以描述數學內容知識，此五個領域爲：

(一)數的特質與運算

1.數的表徵

2.數的次序

3.數的計算

4.對給定的情境正確的估計

5.使用比率和比例的推理

6.應用數的特質和運算解決真實世界之數學問題

(二)測量

1.了解測量的屬性

2.選擇適當的單位和工具進行測量

3.解決有關測量的應用問題

(三)幾何

1.界定幾何圖形

2.呈現圖形之對稱和轉換的知識

3.界定翻轉、旋轉的圖像

(四)資料分析與機率

1.資料呈現

2.資料集的特徵

3.實驗與抽樣

4.機率

(五)代數

1.組型、關係和函數

2.代數表徵

3.變項、式子和運算

4.方程式和不等式

二、TIMSS所界定之數學內容知識

TIMSS(2003)在數學內容知識部分定義特定的數學主題，主要區分為五個主題，各主題下存在若干細目：

(一)數

1.數感、數的特質和數的了解

2.表現數的方式

3.數之間的關係

4.數系

(二)代數

1.數量的組型和關係

2.使用代數符號表徵數學情境

3.發展產生等價式之流暢性

4.解決線性問題

(三)測量

1.瞭解測量的屬性

2.呈現單位的相似性

3.測量多重屬性時說明單位和過程的相似性

(四)幾何

1.瞭解座標的特徵

2.在二維和三維的形體及表徵中使用空間想像的技巧

3.學生應有能力使用對稱和應用轉換去分析數學情境

(五)資料

1.瞭解如何蒐集資料

2.瞭解如何組織資料

3.以圖表方式呈現資料

三、NCTM所界定之數學內容知識

NCTM(2000)所界定之數學內容知識，共分爲五大類，陳述如下：

(一)數與運算

1.了解數、表徵數的方式、數之間的關係和數系

2.了解運算的意義及運算之間的關係爲何

3.熟練的計算及合理估算

(二)代數

1.了解組型、關係及函數

2.使用代數符號表示及分析數學的情境與結構

3.使用數學模式以表示和了解數量的關係

4.分析在各種脈絡中的變化

(三)幾何

1.分析平面或立體圖形的特徵與性質，發展幾何關係的數學論證

2.藉座標幾何與其他表徵系統，描述位置與空間關係

3.應用轉換與使用對稱，以分析數學的情境

4.使用視覺、空間推理和幾何模式，以解決問題

(四)測量

1.了解物件、單位、系統的測量屬性和測量過程

2.應用適當的技術、工具與公式進行測量

(五)資料分析與機率

1.將資料形成問題，並收集、組織與呈現相關資料回答該問題

2.選擇與使用適當的統計方法以分析資料

3.發展與評估資料中顯示的推論與預測

4.了解與應用機率的基本概念

四、九年一貫課程所界定之數學內容知識

九年一貫課程在數學內容知識部分，區分爲五個主題，各主題下存在能力指標(教育部, 2003)：

(一)數與量(包括測量)

1.整數

2.量與實測

3.有理數

4.估算

5.方根的認識及根式的四則運算

6.等差數列與級數的理解

(二)幾何

1.幾何形體的認識、探索與操作

2.學習運用幾何形體的構成要素（如角、邊、面）及其數量性質（如角度、邊長、面積）

3.了解形體的性質與幾何量的計算及非形式化推理

4.學會推理幾何證明

(三)代數

1.能理解常用算術符號的使用方式，並用來列出日常問題的算式，以進行解題

2.四則混合計算

3.理解等量公理

4.以數學式描述問題中有關之數量關係，並理解數學式中定數與變數之差異

5.方程式與不等式的解法，及其解之合理性

6.一維數線、二維平面直角座標系相關定義及內容

7.函數的概念，以及線型函數、二次函數的樣式及其圖形

8.以乘法公式進行多項式因式的分解

9.商高定理及其應用

(四)資料分析、統計與機率

1.建立資料的整理與分組的概念，進而練習報讀與說明資料

2.認識長條圖、折線圖與圓形圖

3.認識累計次數、累計相對次數、百分位數、中位數、全距、四分位距

4.機率

(五)連結

數學內部的連結，可以貫穿前四個主題，強調的是解題能力的培養；數學外部的聯結，則強調生活及其他領域中數學問題的察覺、轉化、解題、溝通、評析諸能力的培養。

研究者針對上述各研究機構對數學內容知識之分類，進行相關的比較，茲以九年一貫課程(教育部，2003)所界定之數學內容知識的分類爲比較的參照依據，見

表2-2：

表2-2：數學內容知識比較表

	九年一貫課程	NAEP	TIMSS	NCTM
數與量	○	○	○	○
幾何	○	○	○	○
代數	○	○	○	○
資料分析、統計與機率	○	○	○	○
連結	○	×	×	×
測量	×	○	○	○

分析九年一貫課程、NAEP、TIMSS、NCTM對數學內容知識之界定可以看到：在「數與量」這個學習主題所著重的為數感、數的特質、數的運算等；在「幾何」這個學習主題所著重的為座標平面、立體圖形、對稱及轉換等；在「代數」這個學習主題所著重的為數量組型及關係、函數、使用代數號進行表徵等；在「資料分析、統計與機率」這個學習主題所著重的為蒐集、組織、圖表呈現、以及機率等。

經由表2-2，我們可以看到各大型評量測驗與教育相關單位對於數學內容知識在「數與量」、「幾何」、「代數」、「資料分析、統計與機率」，都視為是學生必須學習的主題，而略有差異的主題則為「連結」與「測量」。

在測量這個主題，NAEP、TIMSS、NCTM皆將其獨立為一個學習主題，而在九年一貫的課程架構下則隸屬於「數與量」之下。研究者認為，測量這個學習主題其實與日常生活息息相關，我們應該多注重此主題的學習及在生活上的應用，經由生活實務讓學生能更有效的學習。

在連結主題，九年一貫的課程架構主要是希望能在「數與量」、「幾何」、「代數」、「資料分析、統計與機率」之間提供一個橫向的關係，能夠將這四個主題有

關部份連結，並能應用於日常生活中。這個部分在NAEP、TIMSS、NCTM並未特別提出，但在圖2-2之NAEP的數學能力評量架構圖中，可以看到NAEP在數學領域的測量，除了測量學科內容知識及認知能力外，還測量「數學力(mathematical power)」，包含推理(reasoning)、連結(connections)和溝通(communication)；而研究者在NAEP、TIMSS所對外公開的測驗試題中，也可以看到某些試題的設計是連結上述四個主要學習主題。因此除了各大型評量測驗與教育相關單位所注重的主要四個學習主題，如何將這四個學習主題進行連結，並且應用於日常生活中，也是學習數學很重要的一環。

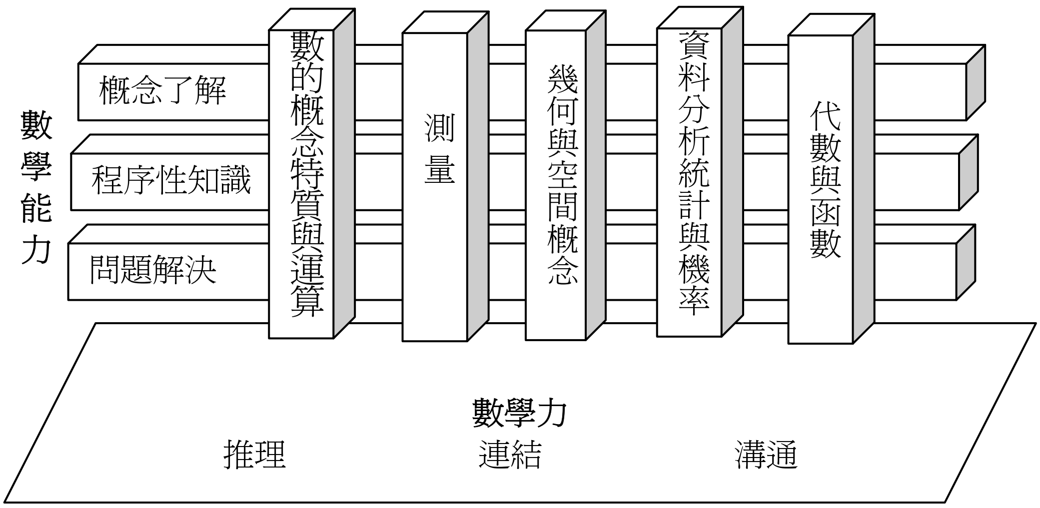

圖2-2：NAEP數學能力評量架構圖(NAEP, 2005)

經由，比較分析九年一貫課程、NAEP、TIMSS、NCTM之數學內容知識，本研究將數學內容知識界定為四類：

1.數與量(包括測量)

2.幾何與空間概念

3.代數

4.資料分析、統計與機率

第四節　數學認知能力

本部份將回顧國內外各大型評量測驗、以及教育相關單位所提出之數學認知能力的理論。有關數學認知能力之理論，了解(NAEP, 2005)、TIMSS(2003)兩個大型評量測驗所界定之數學內容知識、同時也提出NCTM(2000)所界定之數學認知能力，以及國內BCTEST(2004) 所界定之數學認知能力。以下針對上述大型評量測驗之及相關教育單位對數學認知能力之定義進行回顧，分述如下：

一、NAEP所界定之數學認知能力

NAEP(2005)在數學認知能力方面提出「數學複雜性(mathematical complexity)」，試圖聚焦於認知能力，數學複雜性區分為低、中或高，每一個複雜性的水準包括知道(knowing)和做(doing)數學，如推理、表現程序、了解概念或解決問題。

複雜性建立在數學能力(mathematical ability)(概念了解、程序性知識問題解決)和數學力(mathematical power)(推理、連結和溝通)上，以下介紹NAEP所界定數學能力：

(一)概念瞭解

1.提供學生可識別、標籤和產生概念的例子之證據

2.使用及轉換相關的模式、圖表、操作和變化的表現

3.識別和應用原理

4.知道和應用事實和定義

5.比較、對照和整合相關概念和原則

6.識別、解釋和應用記號、符號、術語以表現概念

(二)程序知識

1.正確的選擇與應用適合的程序

2.藉由具體的模式或象徵性的方法，檢驗、判斷程序的正確性

3.擴充或修正程序，以處理問題中原有的要素

(三)問題解決

1.問題的辨識與形成

2.決定資料的一致性

3.使用策略、資料、模式

4.產生、擴充與修正程序

5.在新情境使用堆裡

6.判斷解答的合理性與正確性

二、TIMSS所界定之數學認知能力

TIMSS(2003)在數學認知能力方面是定義學生在數學內容之預期行爲，主要區分爲四個認知能力，各認知能力下存在若干細目：

(一)知道事實與程序(knowing facts and procedures)

1.回憶(recall)

2.識別/辨識(recognize/identity)

3.計算(compute)

4.工具使用(use tools)

(二)使用概念(using concepts)

1.知道(know)

2.分類(classify)

3.表徵(represent)

4.形成公式(formulate)

5.區別(distinguish)

(三)解決例行性問題(solving routine problems)

1.選擇(select)

2.模式化(model)

3.解釋(interpret)

4.應用(apply)

5.證明/檢查(verify/check)

(四)推理(reasoning)

1.假設/推測/預測(hypothesize/conjecture/predict)

2.分析(analyze)

3.評估(evaluation)

4.一般化(generalize)

5.連結(connect)

6.綜合/整合(synthesize/integrate)

7.解決非例行問題(solve non-routine problems)

8.證明(justify/prove)

三、NCTM所界定之數學認知能力

NCTM(2000)在數學認知能力方面，主要區分爲四個認知能力，各認知能力下存在若干細目：

(一)問題解決

1.由問題解決建立新的數學內容知識

2.解決數學和其他脈絡的問題

3.應用各種不同但合適的策略，以解決問題

4.監控和反應數學問題的解題過程

(二)推理與證明

1.辨識數學的基本理論與證明

2.形成和建議數學推測

3.發展與評估數學的論證與證明

4.選擇與使用各種推理形式和證明方法

(三)溝通

1.藉由溝通，組織與強化數學思維

2.能一致的、清楚的與同儕、老師及其他人溝通他們的數學思維

3.分析和評估他人的數學思維與策略

4.使用數學語言，以精確地表示數學概念

(四)連結

1.辨識和使用數學概念間的連結

2.了解如何形成數學概念內部連結和建立具一致性之整體概念

3.辨識與應用數學於一般日常生活情境

(五)表徵

1.創造與使用表徵，以組織、紀錄與溝通數學的概念

2.選擇應用轉換數學表徵，以解決問題

3.使用表徵，以模式化和解釋物理的、社會的與數學的現象

四、國中基本學力測驗將數學認知能力界定爲

國中基本學力測驗將數學認知能力界定爲三大類，各類之下存在若干細目(林世華, 2004)：

(一)概念理解

1.知道與記憶數學的知識與概念

2.用不同的表徵方式表示相關數學內容知識與概念

3.能舉例說明或分類區別相關的數學概念

4.解釋與理解數學相關概念或其交互關係（原理、原則、性質、定理）

(二)程序執行

1.選擇、執行正確的程序或方法

2.使用工具(電算機、圓規、尺、量角器、…)

3.製作、使用圖形或表格

4.檢驗、判斷程序的正確性

(三)解題與思考

1.將問題轉化成幾何圖形或數學式子

2.學的知識與技能的應用

3.題過程（答案）之合理性或正確性的判斷

4.數學思維將問題一般化

5.命題真偽的證明

6.數學的模式與決策的評析

7.推論與猜想

在瞭解各研究機構對數學認知能力之定義後，研究者進行相關的比較，以瞭解各研究機構所重視的數學認知能力，茲以國中基本學力測驗(BCTEST)所界定之數學認知能力定義為比較的參照依據，見表2-3：

表2-3：數學認知能力比較表

	BCTEST	NAEP	TIMSS	NCTM
概念理解				
知道與記憶數學的知識與概念	○	○	○	○
用不同的表徵方式表示相關數學內容知識與概念	○	○	○	○
能舉例說明或分類區別相關的數學概念	○	○	○	×
解釋與理解數學相關概念或其交互關係	○	○	○	○
程序執行				
選擇、執行正確的程序或方法	○	○	○	○
使用工具	○	○	○	○
製作、使用圖形或表格	○	○	○	○
檢驗、判斷程序的正確性	○	×	×	×
解題與思考				
將問題轉化成幾何圖形或數學式子	○	○	○	○
數學的知識與技能的應用	○	○	○	○
解題過程（答案）之合理性或正確性的判斷	○	○	○	○
以數學思維將問題一般化	○	×	○	×
命題真僞的證明	○	○	○	○
數學的模式與決策的評析	○	×	○	×
推論與猜想	○	×	○	○

經由上表之比較發現，雖然對於數學認知能力之定義各有不同，但主要可區分爲概念理解、程序知識及執行、及問題解決三大類，各類下又可細分若干項目：概念理解能力主要著重於記憶、表徵、區別數學概念及其交互關係；程序知識與執行主要著重於選擇及執行正確的程序或方法、工具使用、使用圖形及表格，其中工具使用的這項能力，研究者認爲很重要，因爲在日常生活中的很多情境會與

數學有關，因此我們必須要教會學生正確的使用工具以執行某項任務；在問題解決上主要著重於將問題轉化爲幾何圖形或數學式子、並且能應用數學的知識與技能、同時要能判斷過程或答案之正確性或合理性等。

上述比較中，最大之不同點在於BCTEST所提出之「檢驗、判斷程序的正確性」，在NAEP、TIMSS、NCTM中皆未特別提出；另外在問題解決中的「以數學思維將問題一般化」、「數學的模式與決策的評析」，其中只有BCTEST、TIMSS有提到，而在以美國爲主要評量及教學的NAEP、NCTM則皆未提到這兩項數學認知能力。由此，我們可以看到美國在問題解決的能力方面，較著重於如何將習得的能力應用於生活情境中。

經由分析、比較數學認知能力，本研究將數學認知能力界定爲三大類：

1.概念理解

2.程序知識及執行

3.問題解決

第五節　數學評量實證研究

經由文獻回顧，對於各大型評量測驗及教育單位所提出之數學內容知識及數學認知能力已有所了解，由於本研究之目的爲了解我國國中生在基本學力測驗數學科之表現，因此，了解先前專家學者所進行之數學評量實證研究是相當重要的，本部份將針對數學評量之實證研究結果進行分析討論。以下說明國內八年級學生參加TIMSS於1999年及2003年所進行之國際數學科成就調查的結果，以及林世華、盧雪梅、陳伯熹(2005)針對國中畢業生進行數學能力研究之結果。

一、TIMSS-1999

我國八年級國中生在1999年有5772位學生，TIMSS1999將數學內容領域，區分爲分數與數感、測量、資料表徵分析和機率、幾何、代數五大類；在數學預期

表現，區分爲知道、使用例行性程序、使用複雜程序、調查和解決問題、溝通和推理五大類。以下分析討論我國八年級國中生在TIMSS-1999的表現(陳竹村，2003)：

(一)總體表現

1.就平均量尺分數來看

TIMSS在各國總體表現上，以「平均量尺分數」的高低來表示各國之成就表現，同時也指出各國的平均數是否顯著的高於或低於國際平均數，或與國際平均數未達顯著差異，TIMSS-1999年數學科的國際平均數爲487，是38個國家的個別平均數的總平均數。

台灣八年級國中生在TIMSS-1999數學成就平均量尺分數爲585，標準誤爲4.0，高於國際平均數487，而最低分數爲南非275，由此可顯示台灣在數學方面的表現確實優於各國平均的表現。

2.就國際排名來看

表2-4：我國八年級國中生在TIMSS-1999成就表現

各領域	分數與數感	測量	幾何	資料表徵 分析和機率	代數	整個 數學領域
排名	3	4	4	3	1	3
平均量尺分數	576	566	557	559	586	585

在TIMSS-1999，我國八年級學生在有參與調查研究的38個國家中，數學內容領域的總題排名爲第三，但數學的前四個國家(新加坡、韓國、台灣、香港)並無統計上顯著之差異。另外，在各數學內容領域之成就表現，我國八年級學生在代數的表現是38個國家中表現最好的，而其他數學內容領域，其排名也分居第三或第

四名，表現堪稱優異。

3.就百分位量尺分數來看

表2-5：各國數學成就各百分位量尺分數

國別	第5百分位	第25百分位	第50百分位	第75百分位	第95百分位
台灣	396(6.4)	524(6.1)	595(4.4)	656(3.5)	739(5.1)
新加坡	464(7.5)	555(5.7)	608(6.6)	658(8.9)	728(6.5)
南韓	448(5.8)	538(3.6)	592(2.3)	640(2.0)	710(3.2)
日本	441(2.8)	529(2.4)	583(1.8)	633(1.9)	702(4.1)
香港	456(11.0)	538(5.8)	587(3.7)	632(3.8)	693(4.1)
南非	113(9.6)	200(5.5)	263(6.6)	337(9.9)	485(11.1)

表2-5，為各國數學成就之百分位量尺分數，其中括號內的數字為標準誤。就第5百分位數，新加坡的量尺分數為464，比台灣396高出甚多，而南韓448、日本441、香港456，也都領先台灣甚多，這些平均成就領先世界的五個參與國中，台灣明顯的落後，這樣的訊息顯示台灣在成就最低的5%，明顯與其他四國有相當差異。

就第95百分位數，台灣的量尺分數為739，明顯高於其他四國，這樣的資訊呈現的是，台灣當前的數學教育模式(包括課程與教學法)，是較利於資優生的。

就第5百分位數及第95百分位數之間的區間，台灣分佈的寬度為343(僅次於南非的372)，也就是說，不考慮前後的各5%學生，其他中間的90%學生之優劣，台灣分散程度僅次於南非，若與新加坡、香港、南韓、日本相比，可明顯看到台灣的中上程度學生表現較這些國家為好，中下程度表現則較這些國家為差。

(二)數學內容知識方面

TIMSS-1999將數學內容領域區分為分數與數感、測量、資料表徵分析和機率、幾何、代數五大類，其各內容領域之題數及答對率，如表2-6(吳琪玉，2004)：

表2-6：我國八年級國中生在TIMSS-1999數學內容領域之試題答對率分析表

內容領域 / TIMSS1999	分數與數感	測量	幾何	資料表徵分析和機率	代數
題數	61	24	22	21	35
答對率(%)	73.01	67.68	70.14	78.32	73.28

由表2-6，可明顯看到：在1999年我國八年級國中生在「測量」表現相對較差，答對率低於0.7，在其他數學內容領域之答對率，則皆高於0.7，其中「資料表徵分析和機率」表現最好。

(三)數學預期表現方面

TIMSS-1999將數學預期表現區分為知道、使用例行性程序、使用複雜程序、調查和問題解決、溝通和推理，其各預期表現之題數及答對率如表2-7(吳琪玉，2004)：

表2-7：我國八年級國中生在TIMSS-1999數學預期表現之試題答對率分析表

預期表現 / TIMSS1999	知道	使用例行性程序	使用複雜程序	調查和解決問題	溝通和推理
題數	21	15	18	25	3
答對率(%)	75.45	75.94	77.43	69.90	64.33

由表2-7，可明顯看到：在1999年我國八年級國中生在「調查和解決問題」「溝通和推理」表現相對較差，答對率低於0.7，在其他數學預期表現之答對率，則皆高於0.7，其中「使用複雜程序」表現最好。

二、TIMSS-2003

我國八年級國中生在2003年有5379位學生參與TIMSS的研究調查。在TIMSS-2003將數學內容領域區分爲數、代數、測量、幾何及資料五大類；同時將數學認知領域區分爲知道事實和程序、使用概念、解決例行性問題、推理四大類。以下分析討論我國八年級國中生在TIMSS-2003的表現(曹博盛，2005)：

(一)總體表現

1.就平均量尺分數來看

TIMSS在各國總體表現上，以「平均量尺分數」的高低來表示各國之成就表現，同時也指出各國的平均數是否顯著的高於或低於國際平均數，或與國際平均數未達顯著差異，TIMSS-2003年數學科的國際平均數爲467，是所有國家的個別平均數的總平均數。

台灣八年級國中生在TIMSS-2003數學成就平均量尺分數爲585，標準誤爲4.6，高於國際平均數467，而最低分數爲南非264，由此可顯示台灣在數學方面的表現確實優於各國平均的表現。

比較TIMSS-1999與TIMSS2003整體數學成就，台灣八年級學生的量尺分數皆爲585，而在TIMSS-2003，新加坡的量尺分數爲605、南韓爲589、香港爲586，這些國家之量尺分數皆優於TIMSS-1999，雖然國際平均量尺分數由TIMSS-1999的487下降至TIMSS-2003的467，但在其他領先的幾個國家都在進步的情形下，我國學生的表現就需要特別注意。

2.就國際排名來看

表2-8：我國八年級國中生在TIMSS-2003成就表現

各領域	數	測量	幾何	資料	代數	整個數學領域
排名	4	4	2	4	3	4
平均量尺分數	585	574	588	568	585	585

在TIMSS-2003，我國八年級學生在有參與調查研究的大約50個國家中，數學內容領域的總題排名爲第四，。另外，在各數學內容領域之成就表現，我國八年級學生在各數學內容領域之排名介於第二到第四名。

3.就數學成就之國際基準點來看

TIMSS-2003八年級數學成就的國際基準點(international benchmark)，共分爲四個等級：

(1)優級(advance)國際基準點---625分

(2)高級(high)國際基準點---550分

(3)中級(intermediate)國際基準點---475分

(4)初級(low)國際基準點---400分

八年級學生數學成就達到各等級國際基準點人數百分比，如表2-9：

表2-9：領先群各國達到各等級國際基準點人數百分比

等級 / 國家	優級(625)	高級(550)	中級(475)	初級(400)	未達400
國際平均	7	16	26	25	26
台灣	38	28	19	11	4
新加坡	44	33	16	6	1
南韓	35	35	20	8	2
日本	24	38	26	10	2
香港	31	42	20	5	2

若就達到優級的人數百分比來看，台灣排名第二，表現算是相當優異；但若就達到高級的人數百分比來看，台灣的名次滑落至第四；若就達到中級的人數百分比來看，台灣的名次則滑落至第五；若就達到初級的人數百分比來看，台灣的名次滑落至第七。亦即台灣學生仍有相當比例的學生是存在學習困難的。

另外，台灣在中級以下(含中級、初級、未達400)的人數百分比爲34%，比新加坡的23%、南韓的30%、香港的27%還差，只比日本的38%好一點；而若只就初級(含初級、未達400)的人數百分比來看，台灣則有15%、新加坡有7%、南韓有10%、香港有7%、日本有12%，相較於其他領先國家，顯示台灣國中二年級學生有一群是拒絕學習或學習遲緩的學生。

(二)數學內容知識方面

TIMSS-2003將數學內容領域區分爲數、測量、資料、幾何、代數五大類，其各內容領域之題數及答對率，如表2-10(吳琪玉，2004)：

表2-10：我國八年級國中生在TIMSS-2003數學內容領域之試題答對率分析表

內容領域 / TIMSS-2003	數	測量	幾何	資料	代數
題數	113	57	57	52	88
答對率(%)	69.35	64.76	68.90	65.40	66.39

由上表可明顯看到：在 2003 年我國八年級國中生之試題答對率約在 0.65-0.7 之間，表現相對較差的數學內容領域仍是「測量」，而表現相對較好的數學內容領域則爲「數」及「幾何」。

(三)數學認知能力方面

TIMSS-2003將數學認知能力區分爲知道事實和程序、使用概念、解決例行性問題、推理，其各認知能力之題數及答對率，如表2-11(吳琪玉，2004)：

表2-11：我國八年級國中生在TIMSS-2003數學認知領域之試題答對率分析表

認知領域 / TIMSS-2003	知道事實和程序	使用概念	解決例行性問題	推理
題數	26	21	34	18
答對率(%)	72.94	72.91	64.99	57.87

由上表可明顯看到：在 2003 年我國八年級國中生，表現相對較差的數學認知領域爲「推理」，而表現相對較好的數學內容領域則爲「知道事實和程序」及「使用概念」。

小結：

就整體表現來看，台灣八年級國中生在 TIMSS-1999 及 TIMSS-2003 之數學成就平均量尺分數，皆優於國際平均數甚多，且與新加坡、南韓、香港、日本等國，名列世界前幾名。在 TIMSS-1999，就各百分位量尺分數，可看到在最高的 5%，台灣學生之表現明顯優於上述國家，但在最低的 5%，台灣學生則明顯表現較差，即台灣中上程度學生較上述國家表現較好，中下程度則表現較差；在 TIMSS-2003，台灣達到數學成就國際基準點的人數百分比，也呈現同樣結果。因此，在我們高興國內學生之數學刻在國際優異表現的同時，也應該多花心力照顧成績較爲弱勢的學生。

三、國中畢業生能力研究

林世華、盧雪梅、陳伯熹(2005)採用 2001-1005 年基本學力測驗國文、英文、數學三科之學生作答反應，欲了解我國國中畢業生(即九年級生)之能力表現(即基本學科能力)。其研究步驟爲：首先，依據能力分析架構分類各科試題；其次，分別以全體樣本與高低分組樣本，計算各類試題平均通過率(難度等級)、及各類試題通過率過 50%試題百分比(能力等級)。由於 2005 年爲九年一貫課程實施後，國中畢業生第一次參與基本學力測驗，因此在進行能力分析時，將 2001-2004 年與 2005 年之學生作答反應，分開計算上述數據。全體樣本數學科能力表現，見表 2-12，高低分組樣本數學科能力表現，見表 2-13：

表 2-12：2001-2004 年與 2005 年全體樣本數學科能力表現

數學能力	試題通過率過50%試題百分比		數學內容知識	試題通過率過50%試題百分比	
	2001-2004	2005		2001-2004	2005
計算能力	●	●	數與量	●	●
理解數學概念與程序	◆	●	代數	◆	●
解釋資料與推論	◆	●	幾何	◆	◆
問題解決	◆	◆	機率與統計	●	●

表 2-12 中，●表示全體樣本數學科試題通過率過 50%試題百分比高於 0.8，◆表示全體樣本數學科試題通過率過 50%試題百分比低於 0.8。就全體樣本來看，國中畢業生不論在九年一貫課程實施前或實施後，計算能力的表現皆相當好；而理解數學概念與程序、解釋資料與推論，則在九年一貫課程實施後的表現較好；但在問題解決部分，則不論是九年一貫課程實施前或實施後，表現皆較差。

表 2-13：2001-2004 年與 2005 年高低分組樣本數學科能力表現

數學能力		試題通過率過 50%試題百分比		數學內容知識		試題通過率過 50%試題百分比	
		2001-2004	2005			2001-2004	2005
計算能力	高分組	●	●	數與量	高分組	●	●
	低分組	◆	◆		低分組	◆	◆
理解數學概念與程序	高分組	●	●	幾何	高分組	●	●
	低分組	◆	◆		低分組	◆	◆
解釋資料與推論	高分組	●	●	代數	高分組	●	●
	低分組	◆	◆		低分組	◆	◆
問題解決	高分組	●	●	機率與統計	高分組	●	●
	低分組	◆	◆		低分組	◆	◆

表 2-13 中，●表示樣本數學科試題通過率過 50%試題百分比高於 0.8，◆表示樣本數學科試題通過率過 50%試題百分比低於 0.8，◆表示樣本試題通過率過 50%試題百分比雖低於 0.8，但非常接近 0.8.。就高分組樣本來看，國中畢業生不論在九年一貫課程實施前或實施後，在所有的部分表現試題通過率過 50%試題百分比，皆高於 0.8，表現較好；就低分組樣本來看，國中畢業生不論在九年一貫課程實施前或實施後，除了在機率與統計的試題通過率過 50%試題百分比，接近 0.8，其他所有的部分皆低於 0.8，表現明顯較差。