

## 第五章 研究結果與討論

本章旨在分析依順序理論而建立的同分母分數減法階層之效率 (efficiency) 與精確性，以及國小學童使用線上形成性評量進行測驗後，同分母分數減法之學習成效，並進一步討論分析的結果。

### 第一節 同分母分數減法階層之效率與精確性

本研究預試中每一位學生皆進行所有題目的計算，因此以隨機抽樣的方式從預試試卷中抽取資料，計算抽樣取出的學生若以同分母分數減法階層進行測驗時，是否能提昇效率並達到一定的精確度。預試資料一共 483 筆，分別抽取 50、150、250、350 與 450 筆資料進行效率與精確性的計算，除此之外，進一步探討測驗人數、效率與精確性三者之關係。效率係指節省題數，也就是以總題數 14 題減掉每位學生實際有作答的題數，即該位學生的個人節省題數。以 50 筆資料為例，抽樣取出的 50 名學生平均節省題數為 11.78，接著除以總題數換算成百分比，得到 89.90%，表示當這 50 名學生以本研究的同分母分數減法階層做為測驗架構進行測驗時，可提升 89.90% 的測驗效率。

精確性是指某位學生未作答的題目符合順序階層預期的程度有多高，愈符合則精確性愈高。表 5-1-1 為 50 名學生的紙筆測驗成績記錄，本研究中困難度最高也最先呈現的試題為第七類型，以編號 2 號的學生來看，該學生第七類型答對，在同分母分數減法階層中，將預測剩下的試題共 13 題該學生會全部答對，但實際資料中發現 2 號學生在較簡單的第九類型反而答錯，與我們所預測的作答情形

不同，因此扣除預測錯誤的這一題，13 題中尚有 12 題的作答情形與預測相同，精確性為 12/13，換算成百分比為 92.3%，表示同分母分數減法階層對 2 號學生的作答情形有 92.3% 的精確性。一一計算其餘學生的精確性再平均後，得到這 50 名學生的平均精確性為 90%。

表 5-1-1 隨機抽樣之 50 名學生預試作答情形、效率與精確性

學生編號	一	二	三	四	五	六	七	八	九	十	十一	十二	十三	十四	節省題數	精確性 (換算百分比)
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
2	1	1	1	1	1	1	1	1	0	1	1	1	1	1	13	92.30%
3	1	1	1	1	0	1	1	1	1	1	1	1	1	1	13	92.30%
4	1	1	1	0	0	1	1	0	1	1	1	1	1	0	13	69.20%
5	1	1	1	1	0	1	1	1	1	1	1	1	1	0	13	84.60%
6	1	1	1	1	1	1	1	1	0	1	1	1	1	1	13	92.30%
7	1	1	1	1	1	0	0	1	1	1	1	1	1	1	12	91.70%
8	0	1	1	1	0	1	1	1	1	1	1	1	1	0	13	76.90%
9	1	1	1	1	1	1	1	1	1	1	1	1	1	0	13	92.30%
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
12	1	1	1	0	1	0	0	1	0	1	1	1	1	1	12	75.00%
13	1	1	1	0	0	0	0	1	1	0	1	1	1	1	12	66.70%
14	1	1	1	0	1	0	0	0	0	0	0	0	1	0	9	55.60%
15	1	1	1	1	1	1	1	1	1	1	1	1	1	0	13	92.30%
16	1	1	1	0	0	0	0	0	1	0	1	0	0	0	2	100.00%
17	1	1	1	1	1	1	0	1	1	1	1	1	1	1	12	100.00%
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
21	1	1	1	1	1	1	0	1	1	1	1	1	1	1	12	100.00%
22	1	1	1	0	1	0	0	1	0	1	1	1	1	1	12	83.30%
23	1	1	1	0	1	1	1	1	1	1	1	1	1	1	13	92.30%
24	1	1	1	1	0	0	1	0	1	1	1	0	1	1	13	69.20%
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
26	1	1	1	1	1	1	1	1	1	1	1	0	1	1	13	92.30%

27	1	1	1	1	0	0	0	0	0	1	1	0	1	0	6	83.30%
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
29	1	1	1	0	0	0	0	0	1	0	1	0	1	0	2	100.00%
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
31	1	0	0	1	0	0	0	0	1	0	1	0	1	0	3	66.70%
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
35	1	1	1	1	1	1	1	0	1	1	1	1	1	1	13	92.30%
36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
38	1	1	0	1	1	0	1	0	1	0	1	1	1	1	13	69.20%
39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
40	1	1	1	1	1	1	1	0	1	1	1	0	1	1	13	84.60%
41	1	1	1	1	0	1	0	0	0	0	1	0	1	0	4	66.70%
42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
43	1	1	1	1	1	0	1	0	1	1	1	1	1	1	13	84.60%
44	0	1	1	1	0	1	1	1	1	1	1	0	1	1	13	76.90%
45	1	1	1	1	1	1	0	1	1	1	1	1	1	1	12	100.00%
46	1	1	1	1	1	1	0	1	1	1	1	1	1	0	12	91.70%
47	1	1	1	1	0	1	0	1	1	1	1	1	1	1	12	91.70%
48	1	1	1	1	1	1	1	1	1	1	1	1	1	0	13	92.30%
49	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	100.00%
50	1	1	1	1	1	1	1	1	1	0	0	1	1	0	13	76.92%
平均															11.78	89.90%

依上述方式分別抽取 50、150、250、350 與 450 筆資料進行效率與精確性的計算，由表 5-1-2 中得知，當樣本人數為 50 人時，節省題數、效率與精確性最高，

表 5-1-2 隨機抽樣樣本之節省題數、效率與精確性

人數	節省題數	效率	精確性
50 人	11.78	84.14%	89.90%
150 人	11.45	81.76%	88.25%
250 人	11.50	82.14%	88.17%
350 人	11.40	81.41%	87.47%
450 人	11.53	82.37%	88.17%
平均	11.53	82.36%	88.39%

然而整體看來，無論樣本人數為 50 人、150 人、250 人、350 人或 450 人，節省題數、效率與精確性三者之差異並不大，五個組別的節省題數大約為 11、12 題，效率約為 82.36%，精確性則介於 87%~90% 之間。茲將表 5-1-2 中，節省題數、效率與精確性的數值資料轉換為折線圖於下。

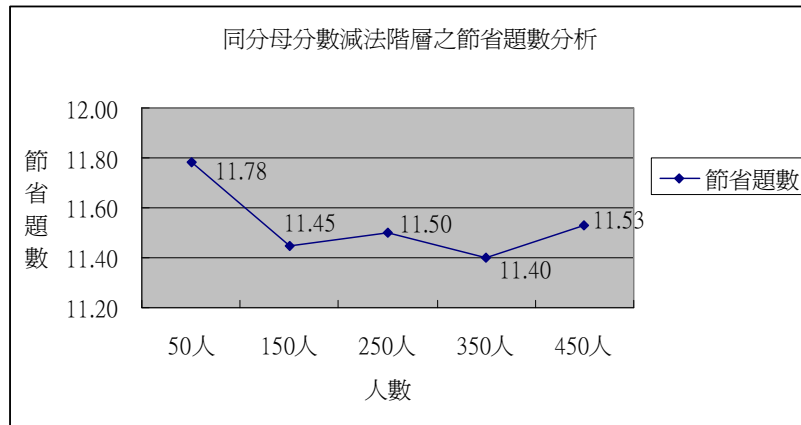


圖 5-1-1 同分母分數減法階層之節省題數分析

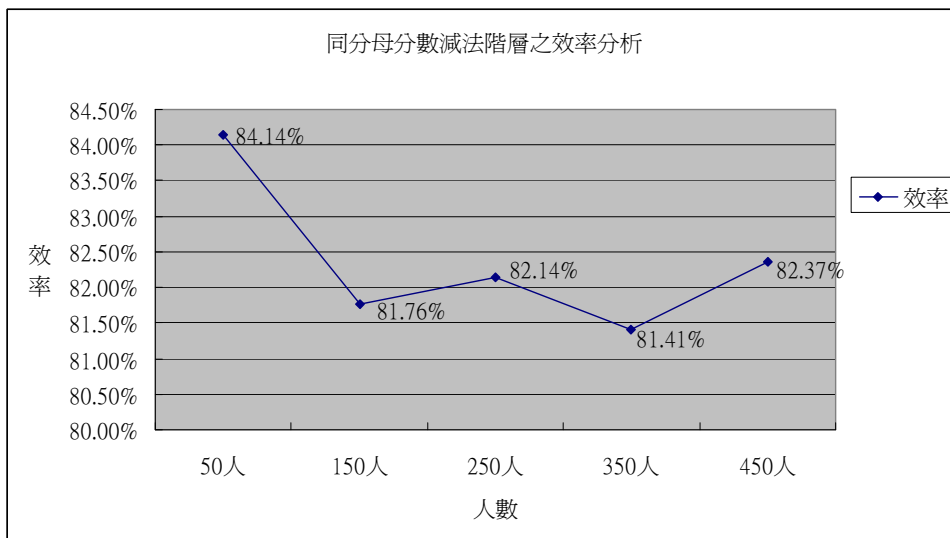


圖 5-1-2 同分母分數減法階層之效率分析

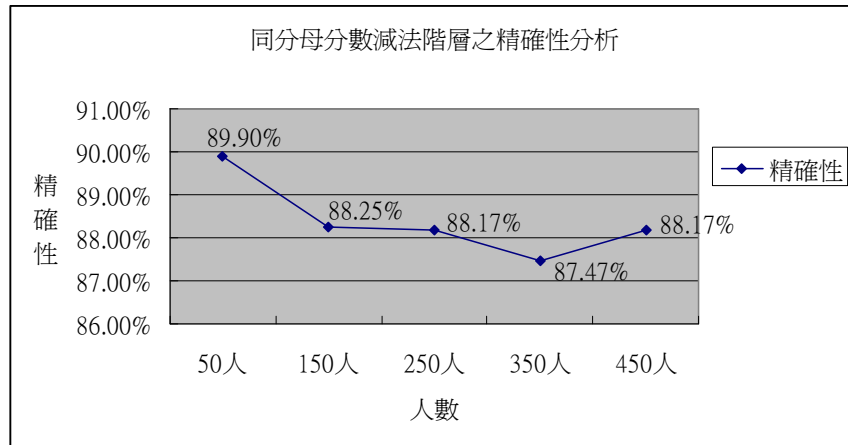


圖 5-1-3 同分母分數減法階層之精確性分析

最後進一步比較控制組與實驗組的平均測驗效率。控制組前測平均作答題數為 14 題，實驗組前測平均作答題數為 2.06 題，實驗組在前測的平均作答題數少控制組 11.94 題，因此實驗組的平均測驗效率高於控制組 85%。控制組後測平均作答題數仍為 14 題，實驗組後測平均作答題數為 1.42 題，實驗組在後測的平均作答題數少控制組 12.58 題，因此實驗組的平均測驗效率高於控制組 90%。實驗組與控制組之詳細資料分析，將於下一節以統計方法加以檢定並說明。

## 第二節 同分母分數減法測驗結果分析

本研究實驗組與控制組的學生於實驗期間進行前後測，試題共十四題，控制組每題以一分計算，往上累加；實驗組每題亦以一分計算，但由於是從難度最高的試題進行測驗，因此以滿分十四分為起始分數，答錯則往下扣減。蒐集兩組前後測資料，分析其樣本數、平均數與標準差的結果如表 5-2-1。

表 5-2-1 同分母分數減法前後測分數摘要表

組別	樣本數	前測		後測	
		平均數	標準差	平均數	標準差
實驗組	50	13.26	1.38	13.80	0.67
控制組	49	12.47	3.06	13.20	2.05

以下針對兩組之測驗成績進行分析，分別採卡方獨立性檢定與二因子混合設計變異數分析，並進一步以數學能力區分高低分組詳細探討。

### 一、卡方獨立性檢定

為了解實驗組與控制組的學生測驗方式與進步狀況(進步、退步、不變)是否互為獨立，因此採用卡方獨立性檢定了解兩者之間的關係，兩個變項即為測驗方式與進步狀況。

表 5-2-2 同分母分數減法進步情況表

組別	樣本數	進步情況		
		退步	不變	進步
實驗組	50	2人	36人	12人
		2.0%	36.4%	12.1%
控制組	49	5人	33人	11人
		5.1%	33.3%	11.1%

實驗組加上控制組的學生共 99 人，由表 5-2-2 可看出其中 69 人的進步情況是不變的，比例高達 69.7%。表 5-2-3 為同分母分數減法之卡方獨立性檢定結果，

$df=2$  時， $\chi^2$  值=1.45， $p=.48$ ，接受虛無假設，表示實驗組別與學生進步情況之間並無顯著相關存在。

表 5-2-3 卡方獨立性檢定摘要表

	Value	df	Asymp. Sig. (2-sided)
皮爾遜卡方	1.45	2	0.48
概似比卡方檢定	1.49	2	0.47
線與線關聯關聯卡方檢定值	0.53	1	0.47
樣本數	99		

## 二、二因子混合設計變異數分析(未區分高低分組)

由於卡方獨立性檢定結果為不顯著，因此進一步試以二因子混合設計變異數分析進行分析，其中兩個自變項分別為實驗組別(A 因子)與前後測之測驗成績(B 因子)，表 5-2-4 為二因子混合設計變異數分析之結果，顯示測驗成績與實驗組別兩個自變項的交互作用為不顯著( $F=0.22$ ， $p=.64$ )。若以實驗組別為自變項，以受試者兩次測驗成績的得分總合為依變項，進行單因子獨立樣本變異數分析，為 A 因子(實驗組別)主要效果的考驗。結果顯示仍實驗組別在依變項(測驗成績)上並無顯著差異( $F=4.12$ ， $p=.45$ )。

表 5-2-4 二因子混合設計變異數分析摘要表(未區分高低分組)

變異來源	SS	df	MS	F	Sig.
受試者間					
組別	23.79	1	23.79	4.12	0.45
誤差	559.80	97	5.77		
受試者內					
測驗成績	20.11	1	20.11	9.38	0.03
交互作用(測驗成績*組別)	0.47	1	0.47	0.22	0.64
誤差(測驗成績)	207.99	97	2.14		

由於交互作用不顯著，因此 B 因子(測驗成績)在結果解釋上具有實質意義

存在。該結果顯示 B 因子（測驗成績）達到顯著，意即就全體受試者而言，後測成績顯著高於前測成績( $F=9.38$ ， $p=.03<.05$ )。

### 三、二因子混合設計變異數分析(區分高低分組)

為徹底了解實驗組別與學生前後測成績之關係，本研究試以受試者第一次段考數學成績將學生區分為高分組與低分組，探討不同數學能力的學生於不同實驗組別，前後測成績是否具有差異。實驗組第一次段考數學成績平均為 81.32 分，控制組第一次段考數學成績平均為 82.47 分，以此為標準篩選出各組高於平均者前 15 人，及低於平均者後 15 人，進行二因子混合設計變異數分析，其中兩個自變項分別為實驗組別(A 因子)與前後測之測驗成績(B 因子)。

#### 1. 高分組

由表 5-2-5 可發現，測驗成績與實驗組別兩個自變項的交互作用不顯著( $F=1.17$ ， $p=.29$ )，B 因子(測驗成績)雖在結果解釋上具有實質意義存在，但其主要效果項亦未達到顯著( $F=0.13$ ， $p=.72$ )。該結果除了顯示高分組的學生在測驗成績與實驗組別兩個自變項的交互作用不顯著外，也顯示高分組的學生學習能力較高，在前測時已發揮良好的學習成效，後測進步空間有限，因此前後測的成績無顯著差異。



表 5-2-5 二因子混合設計變異數分析摘要表(高分組)

變異來源	SS	df	MS	F	Sig.
受試者間					
組別	0.27	1	0.27	0.61	0.44
誤差	12.33	28	0.44		
受試者內					
測驗成績	0.07	1	0.07	0.13	0.72
交互作用(測驗成績*組別)	0.60	1	0.60	1.17	0.29
誤差(測驗成績)	14.33	28	0.512		

## 2. 低分組

低分組方面，由表 5-2-6 可發現，測驗成績與實驗組別兩個自變項的交互作用雖不顯著( $F=1.07$ ， $p=.31$ )，但 B 因子（測驗成績）的主要效果項達到顯著( $F=6.52$ ， $p=.02<.05$ )。該結果顯示相對於高分組的學生，低分組的學生後測進步空間較大，前後測成績達到顯著。

表 5-2-6 二因子混合設計變異數分析摘要表(低分組)

變異來源	SS	df	MS	F	Sig.
受試者間					
組別	70.42	1	70.42	5.83	0.23
誤差	338.07	28	12.07		
受試者內					
測驗成績	22.82	1	22.82	6.52	0.02
交互作用(測驗成績*組別)	3.75	1	3.75	1.07	0.31
誤差(測驗成績)	97.93	28	3.50		

### 第三節 討論

本研究以同分母分數減法階層為測驗架構，該階層的效率與精確度經由計算後得知，無論樣本人數為 50 人、150 人、250 人、350 人或 450 人，節省題數、效率與精確性三者間的差異並不大，五個組別的節省題數大約為 11、12 題，效率約為 80%~85%，精確性則介於 87%~90%之間，顯示同分母分數減法階層架構樣本數約在 500 人之內時，效率與精確性都十分穩定，因此可以說，容忍水準設為 0.12 且樣本數約在 500 人之內時，同分母分數減法階層的效率可達 80%以上，精確性則可達 87%以上。

容忍水準為影響階層架構的重要關鍵，在郭伯臣等(2005)的研究中，進一步探討容忍水準、樣本大小與精確性之間的關係，由於本研究並未探討在不同的容忍水準之下，同分母分數減法階層之架構、效率與精確性的變化，這將是未來繼續研究的項目之一。

而同分母分數減法計算的十四個類型中，由統計分析結果得知，實驗組與控制組的學習成效沒有顯著差異，進一步分析發現，在滿分 14 分的題目中，實驗組在前測的平均分數為 13.26 分，而控制組為 12.47 分，均已接近滿分，導致進步空間太少而造成學習成效的不顯著，可能為「天花板效應」(ceiling effect) (Reber, 2003)。前後測成績相比較後，雖發現進步幅度小，但若屏除天花板效應的因素，補救教學的效益應該更大，學生的進步分數應該更高，因此，雖實驗組與控制組間無顯著差異，但兩組後測成績皆高於前測成績，仍證明經由教師之補救教學，

能提高學生的學習成效。因此我們可以說線上形成性評量與傳統紙筆測驗皆能輔助教師掌握學生學習困難，進而實施補救教學。

以數學能力區分高低分組後發現，無論是實驗組或控制組，高分組的學生前後測成績無顯著差異，低分組的學生則是因進步空間較大而在前後測成績上有顯著差異。在張燕滿（2004）為期兩週實施補救教學的研究結果中發現，中低成就學生的進步幅度顯著高於高成就的學生，這與本研究結果類似，顯示中低成就的學生比高成就的學生更需要補教學。林岑、葉道明與楊漢勇（2005）的研究對象皆為低成就的學生，也就是「計算機概論」學期成績不及格之學生。將研究對象分成兩組，實驗組的學生透過教師個人建置的數位學習平台進行重補修課程，而控制組的學生則是在一般教室進行傳統教學的重補修課程。研究結果顯示，數位教室學生在學習成效上明顯地優於在傳統教室上課學生。由上述兩個研究發現，低成就的學生比高成就的學生更需要補救教學，且當低成就的實驗組學生使用資訊科技進行補救教學時，學習效果比低成就的控制組學生進行傳統式補救教學來得高。但若能在平日課堂中應用行動載具，立即對學生錯誤或不懂的觀念進行補救教學，或許能大幅減少需要重補修的學生人數。

在 Greasley (1999)、Mary Peat 與 Sue Franklin (2002) 比較紙筆測驗與電腦測驗(含形成性評量與總結性評量)的研究結果中，皆顯示使用電腦測驗的學生之學習成就較傳統紙筆測驗的學生來得高，雖與本研究結果相左，但本研究的實驗時間包含平板電腦訓練時間一共四堂課，而 Greasley (1999)、Mary Peat 與 Sue

Franklin (2002) 的研究皆長達兩年，因此時間是否為影響研究結果的因素之一尚待日後加以驗證。

同樣是線上評量系統，戚玉樑、楊紹明與關銘(2003)利用題目反應理論 (Item Response Theory, IRT)、電腦化適性測驗 (Computerized Adaptive Test, CAT) 及權數計分等機制，建立具自動分析能力與回饋機制之評量系統，提供學生與教師更具能力鑑別力的評量測驗平台。在試題屬性方面，題型分為單選、複選、是非；難易度分為困難、偏難、中、偏易；類型分為知識、理解、應用、分析、推理，評分方式以難易度與類型結合，建立答題得分矩陣，選題方式採取五題一循環的方式進行，這五題分別為五種不同的試題類型。本研究中的難易度依據順序理論 (OT) 建立階層，選題機制已考慮試題難易度，但題型皆為填充題，且因題型皆為計算題，因此未區分試題類型，預期未來將這兩點納入系統改良範疇，使系統之周延性更加完善。

本研究為增加教師使用該系統的彈性，由教師於課前自行出題，建立階層方面，教師可選擇以專家知識結構或順序理論為基礎，除了出題與建立階層的彈性之外，在課堂中學生以行動載具進行測驗，教師再立即觀看結果以達即時補救教學的目的。因此，本研究將形成性評量系統定位為課堂中教師決定補救教學策略的一種工具。曾彥鈞、劉育隆與郭伯臣(2006)的研究則是依照學生測驗結果將學生分類至不同的補救教學模組，再由系統中已建立好的多媒體補救教學模組進行補救教學，系統定位較屬於學生在家自我學習並由電腦扮演補救教學者的性

質。日後或許可將兩者整合，部份較適合以多媒體扮演輔助教學角色的單元，可讓學生在家自行測驗並由系統給予回饋，若是較適合在課堂中討論、同儕合作學習或需教師直接指導的單元，可在課堂中藉由行動載具與無線網路協助教師教學與學生學習，將兩種不同性質的評量機制加以結合是未來努力的目標之一。