

第四章 研究結果與討論

本章分為三節，第一節民國九十年第一次國中基本學力測驗數學科校準後所呈現量尺分數之研究結果，第二節民國九十年第二次國中基本學力測驗數學科校準後所呈現量尺分數之研究結果，第三節綜合討論。

第一節 民國九十年第一次國中基本學力測驗數學科校準後所呈現量尺分數之研究結果

為了解所有考生與本研究所獲得資料間的差異，茲將敘述性的統計列於表一，民國九十年第一次國中基本學力測驗數學科所參加的總人數為 299368 人，若以每題答對得一分來計算，平均分數為 18.45，分數的標準差為 7.26，此次考試的信度 α 為 0.89，而此研究所獲得的樣本數為 4999 人，平均分數為 18.62，分數的標準差為 6.74，信度 α 為 0.88。因此樣本平均數略高於母體 0.17，而標準差的差異則較母體少了 0.52，因此樣本分數的分布將較母體的分數集中。

表一 第一次國中學測母體與樣本資料統計表

	母體	樣本
總人數	299368	4999
原始平均分數	18.45	18.62
原始分數的 SD	7.26	6.74
信度 α	0.89	0.88

第一次國中學測所取得的樣本資料經過 BILOG-MG 以 1PL 及 3PL 模式估計後，和學測所題供之 b 值呈現於表二中，本研究經 1PL 模式校準後估計所得 b 值與學測所提供之 Rasch 模式 b 值頗為相近，平均難度相差 0.01，其中以第 26 題差異最大為 0.14，第 8、10、25 和 29 幾乎相等，標準差相差為 0.02。

表二 第一次國中學測 Rasch、1PL 及 3PL 模式之試題參數估計值

題號	學測估計		樣本估計		
	Rasch_b	1PL_b	3PL_a	3PL_b	3PL_c
1	-1.52	-1.45	0.85	-0.55	0.39
2	-2.04	-1.99	1.46	-1.15	0.20
3	-1.51	-1.50	1.09	-0.82	0.23
4	-2.06	-2.02	0.83	-1.49	0.20
5	-1.62	-1.57	0.81	-0.71	0.38
6	-1.70	-1.75	0.86	-1.26	0.17
7	-1.00	-0.97	1.63	-0.18	0.30
8	-1.55	-1.55	1.85	-0.81	0.17
9	-1.46	-1.38	1.00	-0.79	0.21
10	-0.35	-0.35	1.18	0.32	0.28
11	-0.49	-0.45	1.09	0.04	0.20
12	-0.39	-0.43	1.73	-0.10	0.11
13	-0.80	-0.85	1.45	-0.23	0.23
14	-0.08	-0.07	1.06	0.63	0.27
15	-1.11	-1.14	0.83	-0.59	0.23
16	0.03	0.06	0.77	1.23	0.33
17	-0.51	-0.55	1.59	0.09	0.26
18	0.14	0.09	1.56	0.29	0.11
19	-0.43	-0.42	0.49	-0.10	0.14
20	0.32	0.29	1.22	0.65	0.18
21	0.19	0.14	1.12	0.69	0.23
22	0.13	0.09	1.23	0.51	0.19
23	-0.19	-0.12	1.26	0.46	0.25

24	0.43	0.40	0.85	1.47	0.29
25	-0.41	-0.41	1.85	0.10	0.21
26	0.06	0.20	1.18	0.96	0.28
27	0.29	0.28	0.90	0.56	0.13
28	0.17	0.25	1.17	1.14	0.31
29	0.70	0.70	0.94	1.07	0.15
30	0.74	0.70	1.01	1.20	0.19
31	1.12	1.09	1.50	1.73	0.22
32	0.62	0.57	1.06	1.52	0.28
平均	-0.45	-0.44	1.17	0.18	0.23
SD	0.88	0.86	0.34	0.87	0.07

樣本資料經過 1PL 模式校準後，所得答對題數之能力估計如表三，1PL 所估得之能力範圍為-3.05 至 2.41 之間，其所採用的量尺平均為 0、標準差為 1，轉換為平均 30、滿分為 60 後的量尺分數呈現於表三的第 4 欄，結果發現，答對題數在 19 題以下，除了答對 13 題高於學測所公布的量尺外，餘皆等於或小於學測所公布的量尺，答對題數高於 20 題的情形者剛好相反，因此量尺分數若是以此種情形呈現，將有利於前半段能力較好的考生，若以變化率而言，此種估計與學測只有 36.36%(12/33)相同。

若以 1PL 模式校準後，再將估得之能力配合題目難度估出真分數，所呈現之真分數分佈如表三之第 5 欄，真分數範圍為 0.98~31.46。此與原始答對題數範圍(0~32)不同，這是因為除非考生的能力值為 $-\infty$ 或 ∞ 才有可能真分數達到 0 或 32。

因為學測所有的題目皆是四選一的選擇題，以試題反應理論而言，若鑑別度 a 無法控制在一小範圍內(0.8~1.25)(王寶墉，民 85)，且猜測度 c 又無法避

免，則最好的模式應是 3PL 模式(Kolen, & Brennan, 1995)。因此，本研究亦採用 3PL 模式來校準估計題目及學生能力參數，所得之結果如表二第 4, 5, 6 欄與表三第 7 欄，由表二之估計結果可以看出，事實上鑑別度及猜測度是存在的。鑑別度從 0.49~1.85，難度範圍為-1.49~1.73，猜測度為 0.11~0.39，其平均分別為 1.17, 0.18, 0.23。能力估計值的範圍則是-2.28~2.38。以 3PL 模式校準估計，對於答對題數相同，但其答對題目不同者，會產生能力估計不同的情形發生，甚至於有可能會有答對題數較多反而能力估計值較低的情況出現，如本研究答對 3 題之能力-2.03 大於答對 4 題之能力-2.04 的情況。為了以答對題數的方式來呈現能力估計值，因此，研究者將答對相同題數的能力作加權平均後呈現出。最後再藉由此能力分布轉換成量尺分數如表三第 8 欄，在答對 17 題之前，大都低於或等於學測之量尺分數，只有答對 2 或 3 題例外，在答對 17 題之後則相反，若以變化率而言，此種估計與學測只有 30.3%(10/33)相同。若再配合估得之題目參數，可計算得配合答對題數的 3PL 真分數分佈，如表三第 9 欄。我們發現其範圍為 8.19~30.87，與 0~32 有很大的差異，這是因為我們採用 3PL 模式，猜測度存在的原因，即使學生能力再低，他都有猜對該題的機率。

表三 第一次學測答對題數與量尺分數、1PL 與 3PL 模式之能力、真分數對照表

答對題數	學測	1PL 能力估計	1PL 量尺	1PL 真分數	3PL 能力估計	3PL 量尺	3PL 真分數
0	1	-3.05	1	0.98	-2.28	1	8.19
1	1	-2.75	1	1.37	-2.28	1	8.19
2	1	-2.49	1	2.15	-2.07	4	8.47
3	3	-2.30	1	2.78	-2.03	4	8.53

4	6	-2.14	3	3.42	-2.04	4	8.51
5	8	-1.94	6	4.36	-1.91	6	8.73
6	10	-1.71	9	5.63	-1.76	8	9.03
7	11	-1.52	11	6.83	-1.63	9	9.33
8	13	-1.41	12	7.58	-1.54	11	9.57
9	15	-1.33	13	8.15	-1.42	12	9.94
10	17	-1.23	15	8.89	-1.24	14	10.57
11	18	-1.08	17	10.05	-1.07	17	11.28
12	20	-0.87	19	11.79	-0.87	19	12.25
13	21	-0.67	22	13.54	-0.69	21	13.25
14	23	-0.54	23	14.73	-0.57	23	13.99
15	24	-0.47	24	15.37	-0.49	24	14.51
16	26	-0.42	25	15.84	-0.40	25	15.13
17	27	-0.34	26	16.59	-0.30	26	15.85
18	29	-0.21	27	17.82	-0.09	29	17.47
19	30	0.00	30	19.8	0.12	32	19.19
20	32	0.20	33	21.64	0.30	34	20.66
21	33	0.34	34	22.87	0.40	35	21.47
22	35	0.43	35	23.63	0.44	36	21.78
23	36	0.50	36	24.19	0.48	36	22.09
24	38	0.61	38	25.04	0.57	37	22.78
25	40	0.79	40	26.31	0.75	39	24.07
26	42	1.02	43	27.7	1.02	43	25.8
27	44	1.22	45	28.69	1.15	45	26.55
28	46	1.38	47	29.34	1.33	47	27.49
29	48	1.55	49	29.92	1.44	48	28.01
30	51	1.79	52	30.55	1.67	51	28.96
31	54	2.09	56	31.09	2.02	55	30.09
32	60	2.41	60	31.46	2.38	60	30.87

根據學測提供之資料，若作答資料以 Rasch 模式估計，以 0~60 量尺分數呈

現，則平均量尺分數為 30，量尺分數的標準差為 12.3；以 1PL 模式估計，則平均量尺分數為 30，量尺分數的標準差為 11.81；以 3PL 模式估計，則平均量尺分數為 30，量尺分數的標準差為 11.91。統計結果如表四。

表四 第一次國中學測母體與樣本量尺分數統計表

	量尺平均分數	量尺分數的 SD
母體	30	12.3
樣本_1PL	30	11.81
樣本_3PL	30	11.91

第二節 民國九十年第二次國中基本學力測驗數學科校準後所呈現量尺分數之研究結果

為了解所有考生與本研究所獲得資料間的差異，茲將敘述性的統計列於表五，民國九十年第二次國中基本學力測驗數學科所參加的總人數為 167440 人，若以每題答對得一分來計算，平均分數為 18.67，分數的標準差為 7.40，此次考試的信度學測中心並無提供。而此研究所獲得第二次的樣本數為 4737 人，平均分數為 18.76，分數的標準差為 7.35，與母體的差異均很小，分別只有 0.09 與 0.05，信度 alpha 為 0.90。

若作答資料改以 Rasch 模式估計，以量尺分數呈現，則平均量尺分數為 30，