

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，

量尺分數的標準差為 12.3，

表五 第二次國中學測母體與樣本資料統計表

	母體	樣本
總人數	167440	4737
原始平均分數	18.67	18.76
原始分數的 SD	7.4	7.35
信度 alpha	無	0.9

第二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 Rasch 模式 b 值呈現於表六中，本研究經 1PL 模式校準後估計所得 b 值與學測所提供之 b 值差異大於第一次學測，平均難度相差達 0.16，其中以第 16 題差異最大為 1.00，第 23 題最小差異為 0.01，標準差相差也較第一次學測大，相差達 0.37。

表六 第二次國中學測 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	-1.61	-1.24	1.62	-0.67	0.20
2	-0.35	-0.50	1.18	0.23	0.31
3	0.13	-0.29	0.91	0.29	0.24
4	-1.80	-1.03	1.40	-0.51	0.20
5	-1.01	-0.92	1.39	-0.47	0.17
6	-1.27	-0.98	2.05	-0.33	0.26
7	-0.15	-0.55	1.48	0.12	0.29
8	0.04	-0.24	0.91	0.02	0.10

9	-2.15	-1.61	1.25	-0.73	0.42
10	-0.65	-0.54	0.92	0.15	0.29
11	-0.76	-1.16	1.87	-0.13	0.45
12	-0.87	-0.82	0.84	-0.46	0.18
13	0.66	0.15	1.60	0.52	0.18
14	0.07	-0.50	1.71	-0.14	0.13
15	-1.57	-1.33	1.45	-0.46	0.39
16	-2.82	-1.82	1.17	-1.22	0.26
17	-0.34	-0.53	1.15	-0.10	0.18
18	-0.80	-0.70	1.44	-0.23	0.19
19	0.14	-0.17	1.56	0.06	0.08
20	-1.24	-1.47	1.10	-0.85	0.29
21	0.03	-0.30	1.59	0.29	0.26
22	0.71	0.12	2.15	0.68	0.25
23	0.35	0.36	2.10	0.60	0.15
24	-0.39	-0.44	1.58	0.05	0.21
25	0.41	-0.14	1.89	0.86	0.38
26	0.42	0.04	1.43	0.63	0.25
27	1.02	0.38	2.06	1.02	0.27
28	0.76	0.19	1.24	1.46	0.37
29	1.04	0.37	1.64	0.68	0.16
30	0.67	0.16	1.18	1.20	0.33
31	0.64	0.10	0.94	0.78	0.24
平均	-0.34	-0.50	1.44	0.11	0.25
SD	0.99	0.62	0.38	0.64	0.09

樣本資料經過 IPL 模式校準後，所得答對題數之能力估計如表七，IPL 所估得之能力範圍為-2.82 至 2.10 之間，為了配合第一次學測的能力範圍，因此將其拓寬使其相符後，再轉為滿分為 60 的量尺分數，結果呈現於表七的第 4 欄，結果發現除了答對 6 與 30 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，因此量尺分數若是以此種情形呈現，將使得大部份的學生的量尺

分數低於學測的量尺分數，若以變化率而言，此種估計與學測只有 28.13%(9/32) 相同。答對 2 題空白，乃所得資料並無恰好答對 2 題者。另以 3PL 模式估計，結果答對 1 者，其能力竟高於 3 題與 4 題者，此乃樣本資料答對 1 題的只有 1 人，而答對 3 與 4 題者其平均反而小的緣故。

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，所呈現之真分數分布如表七之第 5 欄，真分數範圍為 0.97~29.47。

同樣地，本研究亦採用 3PL 模式來校準估計題目及學生能力參數，所得之結果如表六第 4, 5, 6 欄與表七第 6 欄，由表六之估計結果可以看出，事實上鑑別度及猜測度亦是存在的。鑑別度從 0.84~2.15，難度範圍為-1.22~1.46，猜測度為 0.08~0.45，其平均分別為 1.44, 0.11, 0.25。能力估計值的範圍則是 -2.01~2.00。若先轉換成與第一次學測所估的能力範圍，再轉成 0~60 的量尺，結果呈現於表七的第 7 欄，結果發現除了答對 1、3、19、20 及 21 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，因此量尺分數若是以此種情形呈現，將使得大部份的學生的量尺分數低於學測的量尺分數，若以變化率而言，此種估計與學測只有 15.63%(5/32) 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 真分數分布，如表七第 8 欄，真分數範圍為 8.22~30.3。

表七 第二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	-2.82	1	0.97	-2.01	1	8.22
1	1	-2.50	1	1.58	-1.74	5	8.55
2	1						
3	4	-2.13	2	2.7	-1.85	5	8.4
4	6	-1.92	4	3.6	-1.76	5	8.52
5	8	-1.66	8	5	-1.67	6	8.67
6	10	-1.47	11	6.25	-1.60	7	8.8
7	12	-1.38	11	6.91	-1.52	8	8.96
8	14	-1.34	12	7.21	-1.46	9	9.1
9	16	-1.29	13	7.6	-1.32	11	9.47
10	18	-1.19	15	8.42	-1.16	14	10.01
11	20	-0.99	17	10.19	-0.98	16	10.77
12	21	-0.74	21	12.61	-0.85	18	11.45
13	23	-0.57	23	14.33	-0.62	22	12.91
14	25	-0.49	24	15.15	-0.52	23	13.65
15	26	-0.46	25	15.46	-0.47	24	14.05
16	28	-0.44	25	15.67	-0.43	25	14.37
17	29	-0.41	25	15.97	-0.30	26	15.48
18	31	-0.33	26	16.8	-0.13	29	17.03
19	32	-0.15	29	18.61	0.12	33	19.39
20	34	0.10	32	21	0.31	35	21.19
21	36	0.30	35	22.72	0.40	37	22.03
22	37	0.40	37	23.5	0.44	37	22.4
23	39	0.44	37	23.8	0.45	37	22.49
24	41	0.49	38	24.16	0.46	38	22.58
25	43	0.59	39	24.84	0.52	38	23.12

26	44	0.80	42	26.08	0.72	41	24.86
27	47	1.06	46	27.28	1.06	46	27.31
28	49	1.27	49	28.01	1.28	49	28.46
29	51	1.46	51	28.51	1.37	51	28.83
30	54	1.72	55	29.02	1.54	53	29.39
31	60	2.10	60	29.47	2.00	60	30.3

表八 第二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	32.54	13.12
樣本_1PL	31.03	13.42
樣本_3PL	30.65	14.01

第三節 綜合討論

本研究的量尺分數計算，是基於所有題目皆來自題庫，且題庫中的題目參數都具有相同的量尺的假設，即所有試題的參數皆已知並具相同的量尺，在此假設下，其實兩個測驗間即不需做等化，因將此兩測驗施測於學生，獲得作答資料後，即可利用軟體求得學生的能力估計值，一般的測驗軟體，都會將考生的能力估計設定於平均數為 0、標準差為 1 的分佈。所以，所估得之能力量尺再