

國中基測量尺分數計算方式之探討

題庫發展組

一、緣起

國民中學學生基本學力測驗（以下簡稱國中基測）自二〇〇一年實施以來，一直是每年全國參加考生人數最多的重要測驗。依據高中暨高職多元入學方案（教育部，1998），國中基測分數為各入學管道的重要依據。為避免社會大眾長久來對聯考「一試定終身」的詬病，國中基測每年實施兩次。基於國中基測一年兩試及不同次測驗分數可以擇優的設計，國中基測採用量尺分數的理由，在於使兩次測驗分數可以直接和公平地比較，傳統聯考的計分方式則無法達到這個目的。

國中基測量尺分數是根據考生的答對題數加以轉換得到的，依據Kolen、Hanson 和 Brennan（1992）所描述的程序，透過利用正弦反函數轉換（arcsin transformation；Freeman & Tukey, 1950）使得各個分數點上的條件分配測量標準誤都大致相等。此做法是期望在量尺化過程中，加入分數精準度的訊息，讓大眾認識分數差距非常小的考生，其能力估計並無顯著差異，進而提升後續分數使用的適當性與正確性。另外，為了使各科目的重要性一樣，我們把各科分數都設定平均數為30且範圍在1-60分之間的量尺分數。

然而國中基測分數經過多年的入學使用，產生幾個實際分數運用與原本規劃不同的地方。第一，多元入學方案中雖無明確規定國中基測為申請入學管道唯一依據，但許多高中及高職在設定申請標準時，只提出國中基測相關入學申請標準，並未落實原本多元評量的理想，同時也讓國中基測背負更重的選材責任。第二，登記分發入學中，各考生依據其測驗總分高低來分發，實際運作的結果可能導致總分差1分的考生，被分發至不同學校。這種分發機制，讓考生對於分數差異相當計較，與原本分數計算方式期望加入的測量誤差概念完全抵觸。第三，雖然各科滿分皆設定為60分，但各科題數不同，測驗難度也不盡相同，登記分發入學是以各科總分作為依據，導致部分考生或家長提出「同樣是五科總題數錯五題，為什麼各科都錯一題的量尺總分比五題集中錯在某一科的量尺總分低？」的質疑。

國中基測屬於高風險測驗，任何與測驗分數公平與正確性相關的議題，一直受到高度的關注。綜合這些年社會大眾對於國中基測量尺分數的評論，主要有下列二者：其一為對於目前考生全錯或只答對數題時，皆對應到量尺分數1分的情形；其二為單科答錯一題時所得到的分數，與全對的考生所得的60分差異過大。其中第二個議題也與各科測驗題數與試題難度分配有關，如果測驗題數較多如社會科、自然科，或試題難度分配難易適中，全對與只錯一題考生的分數就不會差異過大。



二、尋求新的國中基測量尺分數產生方法

本文目的在於探討國中基測各科量尺分數現行計算方式與其他計算方式的差異，並比較其優缺點。在維持量尺分數為 1-60 分以及固定平均數的架構下，我們考慮用另外兩種量尺轉換的方式，第一種方式為將原始分數線性轉換以得到量尺分數，也就是將原本量尺計算方式中的正弦反函數轉換步驟去除。原本正弦反函數轉換目的是使各分數的測量標準誤一致，但測量誤差的概念並未落實在分數解釋與實際使用上，而原本的非線性轉換造成高分群範圍極大的分數落差，這個現象加上國中基測各科中等偏易的原則，導致高能力考生較大的能力估計誤差，透過線性轉換方式來建立量尺分數的公式如下：

$$s^* \equiv s^*[x] = A\{x - \mu(X)\} + \mu_s^* \quad (1)$$

$$A = \frac{60 - \mu_s^*}{K - \mu(X)} \quad (2)$$

其中 μ_s^* 是我們想要有的平均數； K 為測驗的題數；而 $\mu(X)$ 為利用考生考完之後實得分數的資料，所得到原始分數的分配機率模式算出 X 的平均數。

第二種方式為利用試題反應理論中的 Rasch 模式所估計得到的考生能力值 θ 直線轉換以得到量尺分數。每年第一次國中基測結束後，依據古典測驗理論中 extended beta-binomial 模式來建立各科的量尺分數。兩次國中基測分數可以互相比較是透過事前參數連結 (linking) 的機制來達成。因為兩次考生群不完全相同，利用各科所有試題的難度和考生的作答表現，去估計考生的 IRT 能力值。由於各科題目的難度已經事先被連結在同一把尺上，所以兩次測驗的 IRT 能力值是可以直接比較的，我們就是透過 IRT 能力值，把兩次的國中基測量尺分數等化。現行國中基測量尺計分與分數等化的方式分別利用古典測驗理論與 IRT 模式來進行，在測量誤差的解釋上，可能產生不一致的情況。因此，我們想利用試題反應理論模式直接來建立各科量尺分數，評估其應用於國中基測實際資料的效果。我們可用下面的公式來轉換：

$$s^* \equiv s^*[\theta] = A\{\theta - \mu(\theta)\} + \mu_s^* \quad (3)$$

$$A = \frac{60 - \mu_s^*}{\max(\theta) - \mu(\theta)} \quad (4)$$

其中 μ_s^* 是我們想要有的平均數；而 $\mu(\theta)$ 為利用考生考完後實得分數的資料，所得到原始分數的分配機率模式算出 θ 的平均數； $\max(\theta)$ 為全對考生的能力值。



三、結果與討論

國中基測各科量尺分數現行計算方式與其他兩種方式各有優、缺點，很難同時改善考生全錯或答對數題時皆對應到量尺分數 1 分的問題，以及單科答錯一題時所得到的分數與全對的考生所得的 60 分差異過大的不合理現象。針對去除正弦反函數轉換算法，其單科只錯一題時所得到的分數，與全對的考生所得的 60 分只差 1 分或 2 分，最符合社會期待。但此種算法在低分端產生比原來的非線性轉換更為嚴重的同分情況，主要是來自於國中基測各科題本中等偏易的設計，使整個原始答對題數分配呈現負偏的狀態。為改善線性轉換方式在低分端形成的嚴重同分現象，我們建議在轉換後尚未裁截到 1-60 分範圍前，加上一常數，如此能改善低分端同分的情形。以國文科為例，表一中呈現各種不同計算方式轉換後，加上常數 20 之答對題數與量尺分數對照表，結果顯示低分端同分的情況已經獲得改善。

表一 九十六年第一次國中基測國文科不同算法量尺平移 20 分比較

答對題數	正弦反函數轉換算法 (做微調)				正弦反函數 轉換算法				去除正弦反函數 轉換算法				IRT 轉換算法			
	未整數化		量尺分數		未整數化		量尺分數		未整數化		量尺分數		未整數化		量尺分數	
	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20
0	-24.6	-4.6	1	1	-18.3	1.7	1	2	-25.9	-5.9	1	1	-11.9	8.1	1	8
1	-18.3	1.7	1	2	-12.8	7.2	1	7	-24.1	-4.1	1	1	-3.7	16.3	1	16
2	-15.0	5.0	1	5	-9.8	10.2	1	10	-22.4	-2.4	1	1	1.3	21.3	1	21
3	-12.4	7.7	1	8	-7.5	12.5	1	13	-20.6	-0.6	1	1	4.2	24.2	4	24
4	-10.0	10.0	1	10	-5.4	14.6	1	15	-18.8	1.2	1	1	6.4	26.4	6	26
5	-7.9	12.1	1	12	-3.5	16.5	1	16	-17.0	3.0	1	3	8.1	28.1	8	28
6	-6.0	14.0	1	14	-1.8	18.2	1	18	-15.2	4.8	1	5	9.6	29.6	10	30
7	-4.2	15.8	1	16	-0.2	19.8	1	20	-13.4	6.6	1	7	10.9	30.9	11	31
8	-2.5	17.5	1	18	1.3	21.3	1	21	-11.6	8.4	1	8	12.0	32.0	12	32
9	-0.8	19.2	1	19	2.7	22.7	3	23	-9.8	10.2	1	10	13.1	33.1	13	33
10	0.7	20.7	1	21	4.1	24.1	4	24	-8.0	12.0	1	12	14.0	34.0	14	34
11	2.2	22.2	2	22	5.5	25.5	5	25	-6.2	13.8	1	14	14.9	34.9	15	35
12	3.7	23.7	4	24	6.8	26.8	7	27	-4.5	15.5	1	16	15.7	35.7	16	36
13	5.1	25.1	5	25	8.1	28.1	8	28	-2.7	17.3	1	17	16.6	36.6	17	37
14	6.5	26.5	7	27	9.3	29.3	9	29	-0.9	19.1	1	19	17.3	37.3	17	37
15	7.9	27.9	8	28	10.5	30.5	11	31	0.9	20.9	1	21	18.0	38.0	18	38
16	9.3	29.3	9	29	11.7	31.7	12	32	2.7	22.7	3	23	18.8	38.8	19	39



測

驗

萬

花

筒

答對題數	正弦反函數轉換算法 (做微調)				正弦反函數 轉換算法				去除正弦反函數 轉換算法				IRT 轉換算法			
	未整數化		量尺分數		未整數化		量尺分數		未整數化		量尺分數		未整數化		量尺分數	
	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20	原本	平移 20
17	10.6	30.6	11	31	12.9	32.9	13	33	4.5	24.5	5	25	19.4	39.4	19	39
18	11.9	31.9	12	32	14.0	34.0	14	34	6.3	26.3	6	26	20.1	40.1	20	40
19	13.2	33.2	13	33	15.2	35.2	15	35	8.1	28.1	8	28	20.8	40.8	21	41
20	14.5	34.5	14	34	16.3	36.3	16	36	9.9	29.9	10	30	21.4	41.4	21	41
21	15.8	35.8	16	36	17.5	37.5	17	37	11.7	31.7	12	32	22.1	42.1	22	42
22	17.0	37.0	17	37	18.6	38.6	19	39	13.5	33.5	13	33	22.7	42.7	23	43
23	18.3	38.3	18	38	19.7	39.7	20	40	15.2	35.2	15	35	23.3	43.3	23	43
24	19.6	39.6	20	40	20.8	40.8	21	41	17.0	37.0	17	37	24.0	44.0	24	44
25	20.8	40.8	21	41	22.0	42.0	22	42	18.8	38.8	19	39	24.6	44.6	25	45
26	22.1	42.1	22	42	23.1	43.1	23	43	20.6	40.6	21	41	25.2	45.2	25	45
27	23.4	43.4	23	43	24.2	44.2	24	44	22.4	42.4	22	42	25.9	45.9	26	46
28	24.6	44.6	25	45	25.3	45.3	25	45	24.2	44.2	24	44	26.5	46.5	27	47
29	25.9	45.9	26	46	26.5	46.5	26	46	26.0	46.0	26	46	27.1	47.1	27	47
30	27.2	47.2	27	47	27.6	47.6	28	48	27.8	47.8	28	48	27.8	47.8	28	48
31	28.5	48.5	29	49	28.8	48.8	29	49	29.6	49.6	30	50	28.6	48.6	29	49
32	29.9	49.9	30	50	30.0	50.0	30	50	31.4	51.4	31	51	29.2	49.2	29	49
33	31.2	51.2	31	51	31.2	51.2	31	51	33.1	53.1	33	53	30.0	50.0	30	50
34	32.6	52.6	33	53	32.4	52.4	32	52	34.9	54.9	35	55	30.7	50.7	31	51
35	34.0	54.0	34	54	33.6	53.6	34	54	36.7	56.7	37	57	31.4	51.4	31	51
36	35.4	55.4	35	55	34.9	54.9	35	55	38.5	58.5	39	59	32.3	52.3	32	52
37	36.9	56.9	37	57	36.2	56.2	36	56	40.3	60.3	40	60	33.1	53.1	33	53
38	38.4	58.4	38	58	37.5	57.5	38	58	42.1	62.1	42	62	34.1	54.1	34	54
39	40.0	60.0	40	60	38.9	58.9	39	59	43.9	63.9	44	64	35.0	55.0	35	55
40	41.6	61.6	42	62	40.4	60.4	40	60	45.7	65.7	46	66	36.1	56.1	36	56
41	43.3	63.3	43	63	41.9	61.9	42	62	47.5	67.5	47	67	37.2	57.2	37	57
42	45.1	65.1	45	65	43.5	63.5	43	63	49.3	69.3	49	69	38.5	58.5	38	58
43	47.0	67.0	47	67	45.2	65.2	45	65	51.0	71.0	51	71	40.0	60.0	40	60
44	49.1	69.1	49	69	47.1	67.1	47	67	52.8	72.8	53	73	41.7	61.7	42	62
45	51.5	71.5	51	71	49.1	69.1	49	69	54.6	74.6	55	75	43.9	63.9	44	64
46	54.1	74.1	54	74	51.5	71.5	51	71	56.4	76.4	56	76	46.9	66.9	47	67
47	57.4	77.4	57	77	54.4	74.4	54	74	58.2	78.2	58	78	51.8	71.8	52	72
48	63.7	83.7	60	80	60.0	80.0	60	80	60.0	80.0	60	80	60.0	80.0	60	80



當使用 IRT 模式計算量尺分數時，因國中基測各科高難度的題數非常少，量尺左端能力估計誤差很大且落差也相當大。以數學科為例（參考表二），單科只錯一題時所得到的分數，與全對的考生所得的 60 分差 9 分。為改善此一問題，建議採用 eap 估計法來計算考生能力，如此全對與只錯一題之考生差距應可縮小。雖然 IRT 模式計算量尺與兩次測驗分數可比較機制所採用的模式一致，測驗理論應用一致性較高，但若國中基測維持目前難度中等偏易的組題原則，搭配大多數升學管道仍把國中基測當作唯一篩選標準，高分端分數的誤差就必須審慎處理，貿然採用此種計分方式，將可能引起更大的反對聲音。

原本國中基測的設計是評量國中畢業生三年所學習的知識與能力，透過標準化測驗研發過程，精確掌握命題原則及題目難度控制，並提供兩次測驗機會，降低考試壓力。量尺分數的應用就是為了使兩次測驗分數可以擇優使用，傳統的計分方式是無法等化兩次測驗分數的。原本正弦反函數轉換算法是希望能讓各能力點的測量標準誤一致，促進分數解釋與使用的合理與正確性。面對大眾對量尺分數計算方式的疑問，大都是因為分數使用與最初的構想有所出入，並非測驗規劃不符合學理。

為解決社會對分數計算的疑慮，我們可以重新再檢視國中基測的設計是否需要再調整。國中基測難度是否一定必須要中等偏易？透過調整試題難度為難易適中，能使高分端的能力估計誤差降低，間接改善高分端考生的問題。另外，國中基測各科分數經常加總使用，為了讓各科分數在總分的比重一致，除了設定相同分數範圍，也可考量各科相同題數的設計。這類建議都可研議規劃，但國中基測屬於高風險入學測驗，任何相關設計與規劃都必須審慎評估，避免後續對教學甚至評量產生負面影響。

表二 九十六年第一次國中基測數學科不同算法量尺比較

答對題數	正弦反函數轉換算法（做微調）		正弦反函數轉換算法		去除正弦反函數轉換算法		IRT 轉換算法	
	未整數化	量尺分數	未整數化	量尺分數	未整數化	量尺分數	未整數化	量尺分數
0	-17.93	1	-16.15	1	-20.09	1	-12.70	1
1	-11.01	1	-9.48	1	-17.66	1	-3.84	1
2	-7.37	1	-5.96	1	-15.24	1	1.51	2
3	-4.40	1	-3.10	1	-12.81	1	4.87	5
4	-1.81	1	-0.59	1	-10.38	1	7.30	7
5	0.55	1	1.68	2	-7.95	1	9.30	9
6	2.73	3	3.79	4	-5.53	1	11.01	11
7	4.79	5	5.77	6	-3.10	1	12.58	13
8	6.75	7	7.66	8	-0.67	1	13.94	14
9	8.63	9	9.48	9	1.75	2	15.22	15



測
驗
萬
花
筒

答 對 題 數	正弦反函數轉換 算法 (做微調)		正弦反函數 轉換算法		去除正弦反函數 轉換算法		IRT 轉換算法	
	未整數化	量尺分數	未整數化	量尺分數	未整數化	量尺分數	未整數化	量尺分數
10	10.46	10	11.24	11	4.18	4	16.44	16
11	12.23	12	12.95	13	6.61	7	17.58	18
12	13.97	14	14.63	15	9.03	9	18.65	19
13	15.68	16	16.28	16	11.46	11	19.72	20
14	17.37	17	17.91	18	13.89	14	20.79	21
15	19.04	19	19.52	20	16.32	16	21.79	22
16	20.70	21	21.13	21	18.74	19	22.87	23
17	22.36	22	22.73	23	21.17	21	23.86	24
18	24.03	24	24.33	24	23.60	24	24.86	25
19	25.70	26	25.95	26	26.02	26	25.94	26
20	27.39	27	27.57	28	28.45	28	27.01	27
21	29.10	29	29.22	29	30.88	31	28.08	28
22	30.84	31	30.90	31	33.30	33	29.22	29
23	32.61	33	32.61	33	35.73	36	30.36	30
24	34.43	34	34.37	34	38.16	38	31.58	32
25	36.32	36	36.19	36	40.58	41	32.93	33
26	38.28	38	38.08	38	43.01	43	34.36	34
27	40.34	40	40.07	40	45.44	45	35.93	36
28	42.52	43	42.17	42	47.87	48	37.72	38
29	44.87	45	44.45	44	50.29	50	39.79	40
30	47.47	47	46.95	47	52.72	53	42.29	42
31	50.44	50	49.81	50	55.15	55	45.65	46
<u>32</u>	54.08	54	53.33	53	57.57	58	51.07	<u>51</u>
<u>33</u>	61.00	60	60.00	60	60.00	60	60.00	<u>60</u>

