

運用評估工具及軟件改善學與教 SP Xpress®操作實踐

主辦：香港教育局
講者：香港教育學院評估與評鑑講座教授暨評估研究中心聯席總監莫基貞教授
教育局課程發展主任
日期：2014年9月18日（星期五）
時間：下午2:30-5:30
地點：九龍塘沙福道19號教育局九龍塘教育服務中心西座 W301室

1

認知診斷評估 Cognitive Diagnostic Assessment

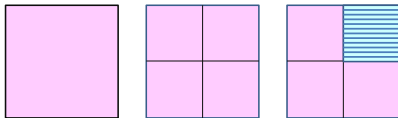
1. 診斷所涉及的…（知識、概念、技能、情意和態度）是甚麼？

- 母體
- 單位量（整體量概念: whole/fraction）
- 等分
- 分母
- 分子

2

分數的意義

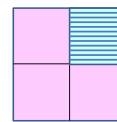
- 分數是描述一個母體被等分後的狀況
- 等分的份數稱為分母
- 被取出的份數稱為分子
- 以母體作為單位



3

認知診斷評估

2. 一位未掌握… 的考生，會有怎樣的表現？



以分數表達



A. $\frac{1}{4}$

B. 1 不明白分子和分數的分別

C. 4 不明白分母和分數的分別

D. $\frac{3}{4}$ 不明白題目要求

4

有關學生學習的診斷

- 表現
- 總分：多項選擇題、對錯題、填充題、步驟、解釋…
- 迷思
- 學生改良警告系數 (Student Modified Caution Indices, 簡稱學生MCI)
- 學生 - 題項表 (Student-Problem Chart)

5

Guttman 圖案 (Louis Guttman 1916-1987)

如果某甲的總分比某乙高，則所有某乙答對的題項，某甲也必然答對。

	Q5	Q8	Q4	Q1	Q3	Q7	Q2	Q9A	Q9B	
S11	1	1	1	1	1	1	1	0	0	7
S3	1	1	1	1	1	1	0	0	0	6
S8	1	1	1	1	1	1	0	0	0	6
S1	1	1	1	1	1	1	0	0	0	5
S5	1	1	1	1	1	0	0	0	0	5
S2	1	1	1	1	0	0	0	0	M	4
S4	1	1	1	1	0	0	0	0	0	4
S9	1	1	1	1	0	0	0	0	0	4
S10	1	1	1	0	0	0	0	0	0	4
S6	1	1	0	0	0	0	0	0	0	2
S7	0	0	0	0	0	0	0	0	M	0
	10	10	9	9	5	3	1	0	0	

M: 沒有作答

6

有關題項的診斷

- 學生 - 題項表 (Student-Problem Chart)
- 題項改良警告系數 (Item Modified Caution Index)
- 題項的容易度 (Item Facility)
- 題項的鑑別度 (Item Discrimination)
- 題項的信度 (Reliability)
- 點二系列相關系數 (Point-biserial Correlation Coefficient)
- 題項的效度
- 選項分析 (Distractor Analysis)
- 痕跡線分析 (Trace line analysis)

13

SP Xpress® 適合用於:

- 形成性評估資料的分析與診斷
- 學生人數 **30-45人** (基本上沒有上限)
- 題目**20-30題** (基本上沒有上限)
- (統計數據+視覺) 判斷
- 讓老師更深入瞭解學生的學習程度與題目困難度的關係
- 具有基本統計知識的老師使用

14

SP Xpress® 示範

Reference

莫基貞、丁彥銓、何昊璇、黃英華、謝樟南、徐坤、姚靜靜 (2011)。優化學習導向評估之SP Xpress 2.2。Hong Kong: Pace Publications Ltd.

15

SP Xpress® (2.2) 操作練習

A. 數據處理

1. 啟動Excel, 打開例子 EX2.xls
2. 理解例子: 11 位學生、9 道題目(沒有第6題; 共有9個給分位), 其中 Q2, Q7 (在第6個給分位), Q9A(在第8個給分位) 是多項選擇題, 其餘是對/錯題 (0, 1)。多項選擇題有四個選項 (A, B, C, D)。若學生沒有回答, 則用M代表。
3. 注意: 題號和給分位的不同; MC 顯示該題為多項選擇題; 顯示正確答案 (Key) 的方法。

16

Example : EX2.xls

	A	B	C	D	E	F	G	H	I	J	K
1			1	2	3	4	5	6	7	8	9
2			MC					MC		MC	
3			1	B	1	1	1	C	1	C	1
4			Q1	Q2	Q3	Q4	Q5	Q7	Q8	Q9A	Q9B
5	1	S1	0	C	1	0	1	B	1	C	1
6	2	S2	0	D	0	1	1	C	1	B	M
7	3	S3	1	A	0	1	1	C	1	A	1
8	4	S4	1	B	1	M	0	C	0	D	0
9	5	S5	1	B	1	1	0	B	0	C	0
10	6	S6	0	C	1	1	0	A	0	B	0
11	7	S7	0	C	0	0	0	D	0	A	M
12	8	S8	1	B	0	0	0	B	1	C	1
13	9	S9	0	A	0	0	1	M	1	C	1
14	10	S10	0	D	0	1	1	C	1	A	0
15	11	S11	1	B	1	1	1	C	1	A	0

17

B. 啟動 SP Xpress®

SP Xpress 檔

按照 閣下的電腦設定, 選取

英文版

SP Xpress 2.2 Eng.xls



中文版

SP Xpress 2.2 Chi.xls

18

B. 啟動 SP Xpress®

1. 啟用巨集(Macros)

在安全性警告
點擊「選項
(Options)」。

選取「啟用這個內
容(Enable this content)」

按「確定
(OK)」

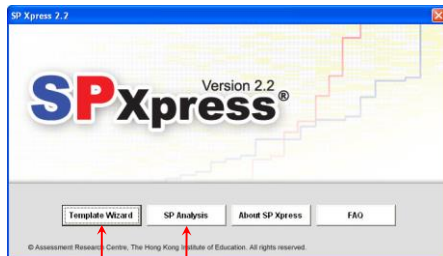
C. 執行「SP Producer」巨集

1. 同時按下 Alt 鍵及 F8 鍵

揀選
「SP_Producer」

按「執行
(Run)」

20



作用: 製作模板, 以輸入數據

21

Items with * are compulsory.

School Name: Example SP Secondary School

Level & Class: Example 1A

Subject: Example Mathematics

Test Name: Example Test on fractions

Number of Students: 11

Number of Items: 9

Generate Template Cancel

SP Xpress® Version 2.2

© Assessment Research Centre, The Hong Kong Institute of Education. All rights reserved.

D. 輸入原始數據 (SP_Raw)

MC

填上 或 貼上資料和數據

23

E. 輸入數據

別改動

24

2. 學生曲線與題目曲線

	5	7	4	1	3	6	2	8	9			
	1	1	1	1	1	MC	MC	MC	1			
	Q5	Q8	Q4	Q1	Q3	Q7	Q2	Q9A	Q9B	MC	CR	TOTAL
Student11	+	+	+	+	+	+	+	A	0	2	5	7
Student3	+	+	+	+	+	A	A	+	+	1	5	6
Student8	+	+	0	+	0	B	+	+	+	2	4	6
Student1	+	+	0	0	0	B	C	+	+	1	4	5
Student5	0	0	+	+	+	B	+	+	0	2	3	5
Student2	+	+	+	0	0	+	D	B	M	1	3	4
Student4	0	0	M	+	+	+	+	D	0	2	2	4
Student9	+	+	0	0	0	M	A	+	+	1	3	4
Student10	+	+	+	0	0	+	D	A	+	1	3	4
Student6	0	0	+	0	+	A	C	B	0	0	2	2
Student7	0	0	0	0	0	D	C	A	M	0	0	0

題目曲線

學生曲線

S2, S4, S9, 及 S10都有4分。他們的答題表現是否一致？

31

改良警告系數 (Modified Caution Index)

- 反映答題情況異常的程度
- 數值在 0 與 1 之間
- MCI 愈大，異常情況愈嚴重
- MCI 在 0.0 與 0.3 之間：正常
- MCI 在 0.3 與 0.5 之間：宜加注意、修正
- MCI 超過 0.5 異常情況非常嚴重

32

改良警告系數 (Modified Caution Indices)

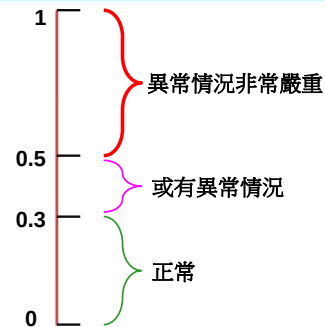
$$C_i = \frac{\sum_{j=1}^{n_{i\bullet}} (1 - u_{ij}) n_{\bullet j} - \sum_{j=n_{i\bullet}+1}^J u_{ij} n_{\bullet j}}{\sum_{j=1}^{n_{i\bullet}} n_{\bullet j} - \sum_{j=J+1-n_{i\bullet}}^J n_{\bullet j}}$$

Source: Harnisch, D. L. (1980); Sato (1975)

© Copyright Magdalena M. C. Mok

33

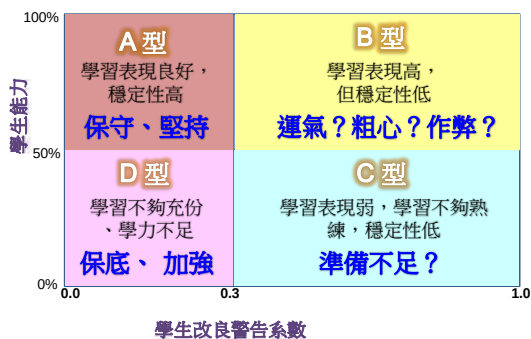
學生「改良警告系數」 (Student Modified Caution Index)



© Copyright Magdalena M. C. Mok

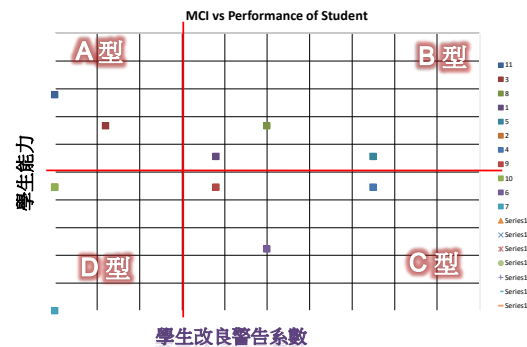
34

學生MCI X 學生能力圖

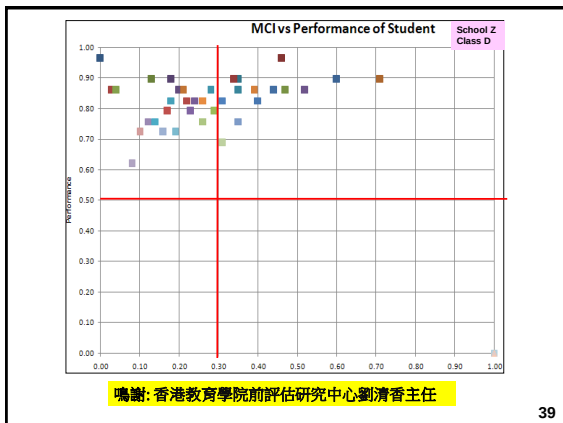
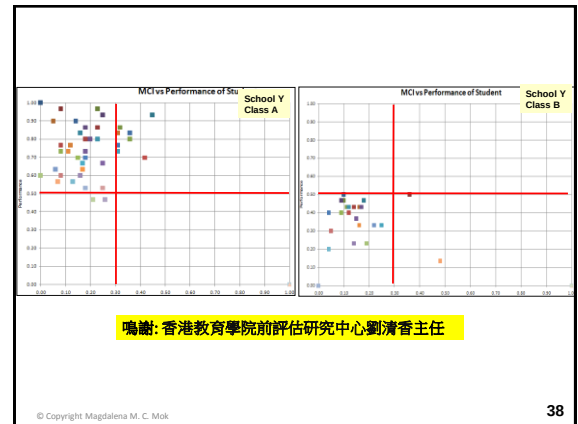
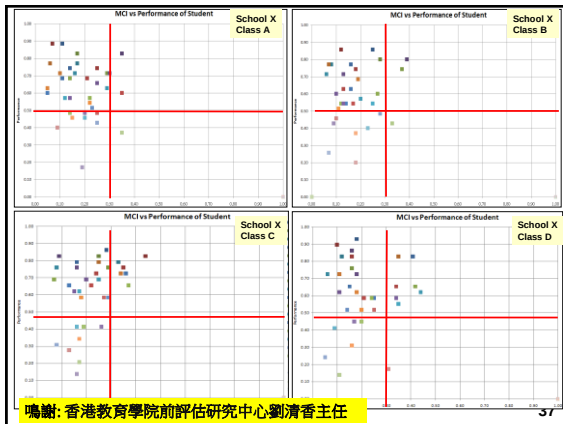


35

例：學生MCI X 學生能力



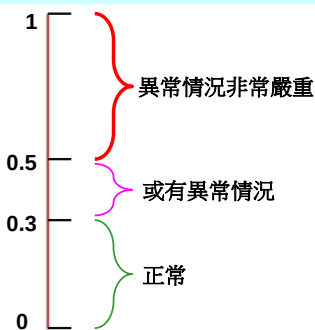
36



有關題項的診斷

- 學生 - 題項表 (Student-Problem Chart)
- 題項改良警告系數 (Item Modified Caution Index)
- 題項的容易度 (Item Facility)
- 題項的鑑別度 (Item Discrimination)
- 題項的信度 (Reliability)
- 點二系列相關系數 (Point-biserial Correlation Coefficient)
- 題項的效度
- 選項分析 (Distractor Analysis)
- 痕跡線分析 (Trace line analysis)

題項改良警告系數 (Item MCI)



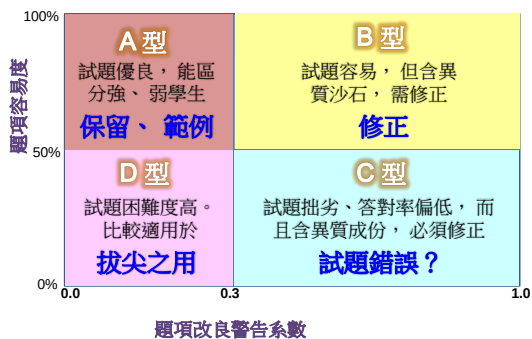
題目容易度 (Item facility)

$$\text{題目容易度 (Item facility)} = R/N$$

R = 本題的答對人數

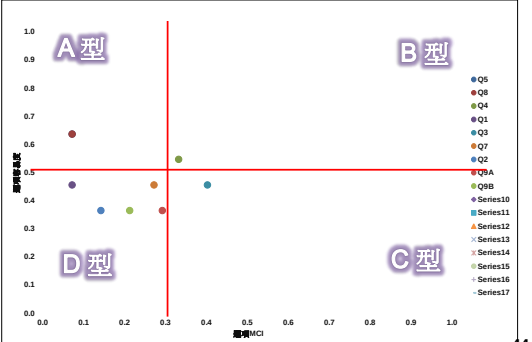
N = 考生總人數

題項MCI X 題項容易度圖



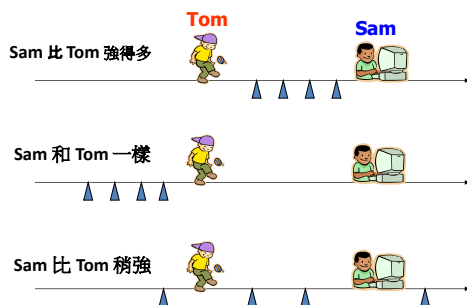
43

例:題項MCI X 題項容易度圖



44

題目與學生的互動



45

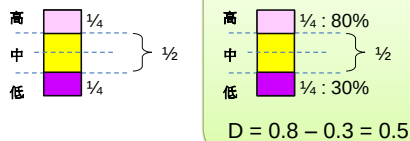
題項資料

Freq. of 1	35	35	30	30		
Freq. of 0	5	5	8	10		
Freq. of Missing	0	0	2	0		
% of 1	87.50	87.50	75.00	75.00		
% of 0	12.50	12.50	20.00	25.00		
% of Missing	0.00	0.00	5.00	0.00		
Freq. of A					28	9
Freq. of B					7	22
Freq. of C					0	6
Freq. of D					3	1
Freq. of Missing					2	2
% of A					70.00	22.50
% of B					17.50	55.00
% of C					0.00	15.00
% of D					7.50	2.50
% of Missing					5.00	5.00
Item P-value	0.88	0.88	0.75	0.75	0.70	0.55
Item MCI	0.07	0.23	0.41	0.07	0.40	0.56
Item MCI Type	A	A	B	A	B	B

46

題項的鑑別度 (Item Discrimination)

- 對於區別高能力和低能力學生之間差異的貢獻
- 高分組和低分組



- 鑑別度 $D = P_H - P_L$
= 高分組的答對率 - 低分組的答對率
- 負值的鑑別度：警告！總分低的學生答對本題，總分高的學生反而做錯！

47

鑑別度 不等於 難度

	Q1	Q5	Q6	Q2	Q4	Q3	總
S1	1	1	1	1	1	0	5
S2	1	0	1	1	0	1	4
S3	1	1	0	0	0	1	3
S4	0	1	1	0	1	0	3
S5	0	0	1	0	1	0	2
S6	1	0	0	1	0	0	2
S7	0	0	0	1	0	0	1
S8	0	1	0	0	0	0	1
總	4	4	4	4	3	2	21

© Copyright Magdalena M. C. Mok

48

	Q4	Q1	Q2	Q5	Q6	Q3	總
S1	1	1	1	0	0	0	3
S2	0	1	0	1	1	0	3
S3	1	1	0	0	1	0	3
S4	1	0	1	0	0	1	3
S5	0	0	1	1	0	1	3
S6	1	1	0	1	0	0	3
S7	1	0	1	0	1	0	3
總	5	4	4	3	3	2	21

鑑別度低
難度中等

© Copyright Magdalena M. C. Mok

49

	Q1	Q2	Q3	Q4	Q5	Q6	total
S1	1	0	0	0	0	0	1
S2	0	0	1	0	0	0	1
S3	0	0	0	0	1	0	1
S4	0	0	0	1	0	0	1
S5	0	0	0	0	0	1	1
S6	0	1	0	0	0	0	1
S7	0	0	1	0	0	0	1

鑑別度低
難度高的試卷

	Q1	Q2	Q3	Q4	Q5	Q6	total
S1	1	1	0	1	1	1	5
S2	1	1	1	1	1	0	5
S3	1	1	0	1	1	1	5
S4	0	1	1	1	1	1	5
S5	1	1	0	1	1	1	5
S6	1	1	1	0	1	1	5
S7	1	1	1	1	1	0	5

鑑別度低
難度低的試卷

	Q1	Q2	Q3	Q4	Q5	Q6	total
S1	1	1	1	1	1	1	6
S2	1	1	1	1	1	1	6
S3	1	1	1	1	1	0	5
S4	1	1	1	1	1	0	5
S5	1	1	1	1	0	0	4
S6	1	1	1	1	0	0	4
S7	1	1	1	1	0	0	4

鑑別度中等
難度低的試卷

© Copyright Magdalena M. C. Mok

50

試卷的信度 (Test Reliability)

- 是試卷題目的內部一致性程度 (internal consistency)

- Cronbach Alpha (α) 系數：

$$\alpha = \left[\frac{m}{m-1} \right] \left[\frac{\sum S_i^2}{\sum S_i^2 + 2 \sum \sum S_{ij}} \right]$$

m 代表總題數

$\sum S_i^2$ 代表所有題目的總變異數

S_{ij} 代表 i 題與 j 題的共變數

51

試卷的信度 (Test Reliability)

	Q4	Q1	Q2	Q5	Q6	Q3	總
S2	0	1	0	1	1	1	4
S7	1	0	1	0	1	1	4
S3	1	1	0	0	1	1	4
S1	1	1	1	0	0	0	3
S5	0	0	1	1	0	0	2
S6	1	1	0	1	0	0	3
S4	1	0	1	0	0	0	2
總	5	4	4	3	3	3	22

中度相關性

強度相關性

Cronbach Alpha 系數是 試卷內所有題目的相關性的平均值

52

Alpha Without (Alpha WO)

- 假設試卷沒有這一道題目，Alpha 系數會是甚麼呢？(試卷信度會否增強？)
- Alpha Without 系數就是取消本題之後，計算所得的 Alpha 系數
- 若 Alpha = 0.7，Alpha Without = 0.8，表示本題和試卷內其他題目相關性比較低

53

點二系列相關系數 Point-biserial Correlation Coefficient

	Q4	Q1	Q2	Q5	Q6	Q3	總
S2	0	1	0	1	1	1	4
S7	1	0	1	0	1	1	4
S3	1	1	0	0	1	1	4
S1	1	1	1	0	0	0	3
S5	0	0	1	1	0	0	2
S6	1	1	0	1	0	0	3
S4	1	0	1	0	0	0	2
總	5	4	4	3	3	3	22

本題和試卷內其他題目的相關性

54

點二系列相關係數 Point-biserial Correlation Coefficient

- 對/錯與除去此題後總分的 Pearson 相關係數
- 0.7 或以上 高度相關
- 0.40 - 0.69 十分強
- 0.30 - 0.39 相當強
- 0.20 - 0.29 一般、可接受
- 0.00 - 0.19 相當弱
- 0.01 或 以下 題目和整體不配合

55

選項分析 (Option/Category Analysis)

Q3 高、中、低分組答對和答錯的比例

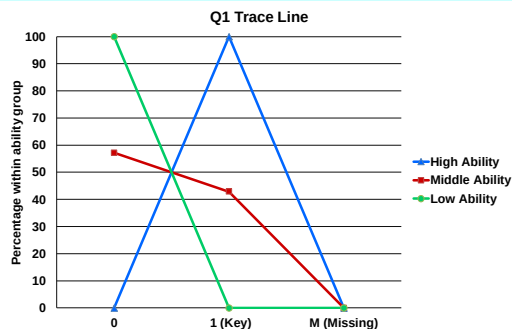
選項	高分組	中分組	低分組
0	50.00	57.14	50.00
1 (Key)	50.00	42.86	50.00
M (Missing)	0.00	0.00	0.00
Total	100	100	100

Q7 高、中、低分組所選取各選項的比例

選項	高分組	中分組	低分組
A	0.00	0.00	50.00
B	0.00	42.86	0.00
C (Key)	100.00	42.86	0.00
D	0.00	0.00	50.00
M (Missing)	0.00	14.29	0.00
Total	100	100	100

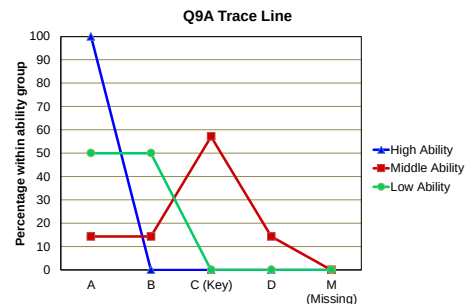
56

痕跡線分析 (Trace Line Analysis)



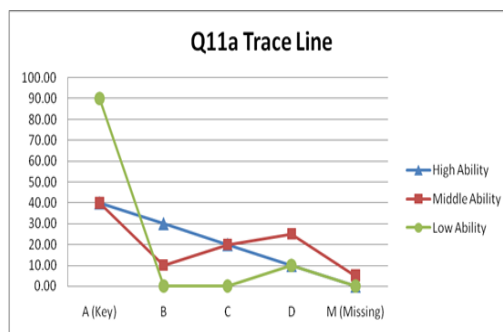
57

痕跡線分析 (Trace Line Analysis)



58

痕跡線分析 (Trace Line Analysis)



59

試卷的效度 (Test Validity)

- 試卷能達到編製目的的程度
- 試卷分數能反映學生的學習的程度

Sam 和 Tom 一樣



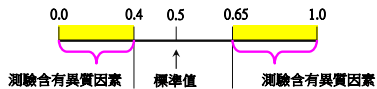
Sam 和 Tom 不一樣



60

差異係數 (Disparity Index, D^*)

- 資料來源：余民寧，2002，342頁
余民寧，《教育測驗與評量-成就測驗與教學評量》，臺北：心理出版社，2002。
- “差異係數 D^* 是指實際S和P曲線所圍面積部分，佔隨機情況下S-P曲線所圍成之面積期望值的比值”
- $0 \leq D^* \leq 1$



61

差異係數 (Disparity Index, D^*)

- 資料來源：余民寧，2002，342頁
余民寧，《教育測驗與評量-成就測驗與教學評量》，臺北：心理出版社，2002。

$$D^* = \frac{C}{4Nn\bar{P}(1-\bar{P})D_B(M)}$$

$M = G[\sqrt{Nn} + 0.5]$ 表示取高斯整數值

C是S曲線和P曲線所包含的 0 和 1 的總數

N是學生人數

n 是題目總數

計算M值之後，用下表查出 $D_B(M)$ 的值。

62

- 資料來源：余民寧，2002，342頁
余民寧，《教育測驗與評量-成就測驗與教學評量》，臺北：心理出版社，2002。

表 2-2 $D_B(M)$ 查表

M	$D_B(M)$	M	$D_B(M)$	M	$D_B(M)$	M	$D_B(M)$	M	$D_B(M)$
11	.278	29	.355	47	.384	65	.402	83	.413
12	.285	30	.358	48	.385	66	.403	84	.413
13	.291	31	.360	49	.386	67	.404	85	.414
14	.296	32	.362	50	.387	68	.404	86	.414
15	.302	33	.364	51	.388	69	.405	87	.415
16	.307	34	.366	52	.389	70	.405	88	.415
17	.312	35	.367	53	.390	71	.406	89	.416
18	.317	36	.369	54	.391	72	.407	90	.416
19	.321	37	.370	55	.392	73	.408	91	.417
20	.326	38	.372	56	.393	74	.408	92	.417
21	.330	39	.373	57	.394	75	.409	93	.418
22	.334	40	.375	58	.395	76	.409	94	.418
23	.337	41	.377	59	.396	77	.410	95	.419
24	.341	42	.378	60	.397	78	.410	96	.419
25	.344	43	.380	61	.398	79	.411	97	.419
26	.347	44	.381	62	.399	80	.411	98	.420
27	.350	45	.382	63	.400	81	.412	99	.420
28	.353	46	.383	64	.401	82	.412	100	.420

Sato, 1985

63

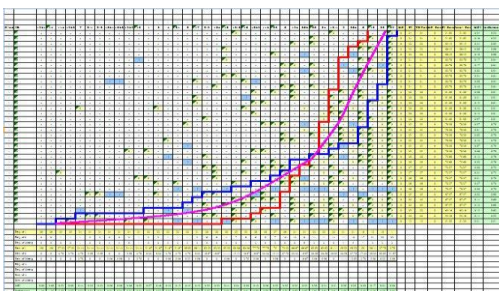
試卷的效度 (Test Validity)

Test A									Test B								
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8		Q4	Q2	Q1	Q6	Q5	Q3	Q7	Q8
1	1	1	1	1	1	1	1	1	8	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	8	2	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	8	3	1	1	1	1	1	1	1
4	1	0	1	1	1	1	1	1	7	4	1	1	1	1	0	1	1
5	1	1	1	1	1	1	0	1	7	5	1	1	1	1	1	1	0
6	1	1	1	1	1	1	0	1	7	6	1	1	1	1	1	1	0
7	1	1	1	1	1	1	0	1	6	7	1	0	1	1	1	1	0
8	1	1	1	1	0	1	1	1	6	8	1	0	1	1	0	1	0
9	1	1	1	0	1	0	1	1	5	9	1	1	1	0	0	0	0
10	1	0	0	0	1	1	1	0	4	10	1	1	1	0	0	0	0
11	1	1	0	1	0	1	0	0	4	11	1	0	0	0	0	0	0
12	0	1	1	0	1	0	1	0	4	12	1	0	0	0	0	0	0
13	0	1	0	1	0	1	0	1	4	13	0	1	0	0	0	0	0
	11	11	10	10	10	10	8	8			12	9	9	8	7	6	5

B試卷的效度 比 A試卷的效度 為高

64

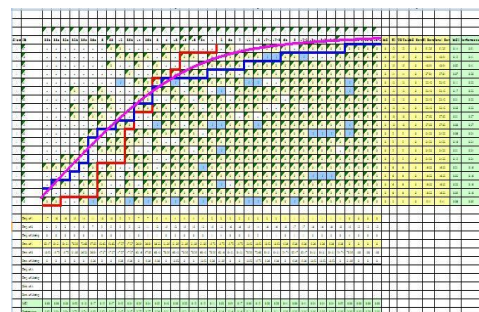
試卷的效度 (Test Validity)



太容易的試卷：低效度

65

試卷的效度 (Test Validity)



太難的試卷：低效度

66

試卷的效度 (Test Validity)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
		Q6	Q4	Q3	Q8	Q7	Q9	Q13	Q5	Q11	Q12	Q14	Q10	Q1	Q15	Q2	TOTAL
1	S2	1	1	1	1	1	1	1	1	1	0	0	1	1	0		12
2	S18	1	1	1	1	1	1	1	1	0	1	1	1	1	0		12
3	S8	1	1	1	1	1	1	0	0	1	1	1	0	1	0		11
4	S5	1	1	1	1	1	1	1	0	1	0	1	0	0	0	1	10
5	S12	1	1	0	1	1	1	1	1	1	0	1	0	0	0	0	10
6	S17	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	10
7	S4	1	1	1	1	1	1	0	0	0	1	0	0	1	0	1	9
8	S10	1	1	1	1	1	0	1	0	0	0	1	1	0	0	1	9
9	S13	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	9
10	S14	0	1	1	0	0	1	1	1	1	1	1	0	1	0	0	9
11	S3	1	1	1	0	0	1	0	1	1	0	1	1	0	0	0	8
12	S6	1	1	1	1	1	0	1	0	0	1	0	0	0	1	0	8
13	S7	1	1	0	1	1	1	1	0	1	0	1	0	1	0	0	8
14	S9	1	1	1	1	1	1	1	1	0	0	0	0	0	1	0	8
15	S15	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	8
16	S19	1	1	1	1	0	1	1	0	0	0	0	0	1	0	0	8
17	S1	1	0	0	1	1	0	0	1	0	0	1	1	0	1	0	7
18	S16	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0	7
19	S11	1	0	1	0	1	0	1	0	1	0	0	0	0	0	0	5
TOTAL		18	17	16	15	14	13	12	11	10	9	9	8	7	6	3	168

難度適中的試卷

67

謝謝聆聽

THANK YOU!

香港教育學院 莫慕貞

Magdalena Mo Ching Mok

mmcmok@ied.edu.hk

68