

第二章 相關文獻探討

第一節 智力相關理論

智力的本質為何？如何測量智力？關於智力的研究可以說是眾說紛紜。在十九世紀末，英國心理學家 F.Galtan 以各種感覺辨識及動作反應的能力來推估個人的智力，此時期的智力測驗大多以心物題目（psychophysical items）為主，這類型題目被認為無法確實估計出智力，與學習適應等表現的關聯也不大。而在二十世紀初，Binet 和 Simon 認為智力是一種普通智慧能力（general intellectual capacity），智力的表現是個體對於生活或社會情境的推理、判斷、實際經驗及環境適應的整體能力之反映，一個人若有很強的理解力，則其歸納等能力應該也會很強。而在西元 1904 年，Spearman 提出了二因論（two-factor theory），認為智力包含普通因素（general factor）及一種或多種特殊因素（specific factor），普通因素的認知活動主要牽涉到關係推理（education of relations）及關聯推理（education of correlates）兩種能力，特殊因素則代表個人的特殊能力，為學習專門知識技能所需要的能力。而在 1938 年，Thurstone 提出七項基本心理能力（primary mental abilities），即語文理解（verbal comprehension）、語詞流暢（word fluency）、歸納與演繹推理（inductive and deductive reasoning）、數字運算（number）、機械記憶（rote memory）、知覺速度（perceptual speed）、空間或視覺化（space or visualization）等能力，相對於 Spearman 的二因論，Thurstone 的理論被歸類為多因論（multi-factor

theory) (蔡崇建，1991)。

第二節 性向測驗

性向測驗是為了測量特殊能力或能力群而設計的工具，可以用來預測一個人
在某些學科或職業上成功的可能性，這類測驗不像智力測驗一樣同時涵蓋多種心
理官能並求出智商作為單一代表值，性向測驗將各種性向加以區分，並使用不同
的方法來測量（葛樹人，1988）。

由於智力測驗涵蓋的範圍受到相當程度的限制，大部分的智力測驗主要在測
量語文能力和一些處理數目及其他抽象符號關係的能力，這類測驗往往只能測量
到某些方面的智力（路君約，1992）。智力是一種綜合性的普通能力，本質偏重
在抽象思考和理解；性向則是方面性及特殊性的能力，與某些特定的學習或技能
活動有關，性向上的個體間差異相當普遍，即使是兩個智商相同的人，在各種性
向上的發展也可能截然不同（葛樹人，1988）。

性向測驗受因素分析研究的影響，發展出各種多元性向組合(multiple aptitude
batteries)，可以對個人提供多項特質的單獨分數，而不是一個總分或是智商，這
類測驗組合對個人內在分析或是區域性的診斷提供了一個適當的工具（路君約，
1992）。

性向測驗的分類主要可分為下列三項：

1. 普通性向測驗：普通性向測驗即是所謂的智力測驗，亦被稱為普通學業能力

(general scholastic ability)、普通教育能力 (general educational ability)、學校性向 (academic aptitude) 或學業性向 (scholastic aptitude) 等，基本上是在測量學校式工作 (school type tasks) 上成功的能力。

2. 特殊性向測驗：本類測驗的發展是為了獲得個人特殊技能或能力的數據，以利進行職業輔導和人員揀選與分類。
3. 多因素性向測驗：早期的心理測驗乃是廣泛運用智力測驗再加上某些特殊性向測驗，但在 1940-1950 年代即迅速發展出多因素性向測驗，其中以美國於 1947 年發展的職業導向普通性向測驗組合 (USES General Aptitude Test Battery, GATB，臺灣譯為通用性向測驗)，與學業導向的區分性向測驗 (Differential Aptitude Tests, DAT) 為代表，多因素性向測驗可以加強對個人內在差異的認識，並補足普通智力測驗未涉及的部份。此類測驗被應用於職業諮商、教育輔導與分類。

多因素性向測驗主要的編製技術，為因素分析法 (Factor-analytic technique)，這也是該測驗容易被人接受的原因之一 (路君約，1995)，一般的多因素性向測驗，多半由 4 至 12 個分測驗所組成，其內容與類別因測驗目的而異。西元 1994 年路君約、盧欽銘與歐滄和等編制了一個適用於國中及高中學生的多因素性向測驗，內含八個分測驗，各分測驗條列說明如下：

1. 語文推理：目的在評估學生抽象或概括的潛力，而非簡單的語文能力或字彙認識。

2. 數學推理：目的在測量受試者了解數目關係以及處理數目概念的能力。
3. 機械推理：主要用來評估需要了解普通物理力學原則的課程或職業。
4. 空間關係：測量空間關係知覺的能力，要求受試者將二度平面在心理上操弄成三度實體，以心理想像補充平面的不足。
5. 抽象推理：乃是學生推理能力的非語文量數，要求受試者由非文字圖案中類化和歸納各種原則。
6. 錯別字和文法修辭：這兩個分測驗較接近於成就測驗，所用的字彙和成語，多選自現行國高中各類教科書。
7. 知覺速度與確度：為本多因素性向測驗中唯一的速度測驗，要求受試者將成對的注音符號、英文字母或數目加以核對，辨別其異同。

第三節 歸納推理

歸納推理 (inductive reasoning) 是人們對所得到的訊息加以整合，找出其共同的型態 (pattern) 或是共通的法則 (rule) 的過程，透過歸納推理，人們可以從所接觸的有限案例中，解釋整體現象發生的原因，並對未來做一定程度的預測，如此可以降低問題的不確定性 (Simon & Lea, 1974; Greeno, 1987)。就邏輯上而言，除非能夠窮舉所有的案例，否則我們無法證明之前歸納出來的法則之正確性。1956年 Bruner、Goodnow 和 Austin (1956) 的研究發現，一般大學生在進行歸納推理時，會將第一個案例所具有的特徵作為目前的假設，然後根據此假設預

測接下來的案例，若新的案例皆符合其假設，則接受此假設為真，若有案例不符合其假設，則修正其假設。此類歸納方式有一定的危險性，只要有一個案例不符合其原先假設的法則，則整個法則將被推翻，故較好的歸納法則方式為證偽 (falsification)，唯有在該法則沒有被任何案例推翻的情況下，我們才能暫時接受該假設法則。有許多方法可以用來測量推理能力，但是較常被使用的有類推測驗 (analogy)、矩陣問題 (matrix problems)、系列完成測驗以及分類測驗 (classification task) (Lohman, 2002)。

第四節 虛擬題庫

「虛擬題庫」一詞由廖文偉 (2002) 於「應用影像處理技術之虛擬題庫設計」所提出。相對於傳統題庫有限的題目數量，虛擬題庫可以利用試題生成法則 (item generation rules) 以產生極大量的試題。蕭惠云 (2004) 亦在「互動式提示虛擬題庫中介練習系統之設計與評估」中運用此概念建立中介練習系統，利用虛擬題庫題目可以自動大量產生的特性，提供給受測者充分的練習機會，然後觀察受測者經過練習之後的進步狀況。

在虛擬題庫此名詞出現之前，劉子鍵 (2001) 亦提出過「圖形推理能力測驗之自動命題及施測系統」的設計、實作與評估，其試題由試題自動產生引擎 (item-generation engine) 依試題產生算則產生，和虛擬題庫同樣可以避免傳統題庫不易更新、增加試題及試題曝光將會影響試題效度的困境。

第五節 系列完成測驗

許多智力測驗都包含有文字系列完成測驗的部份，Thurstone 的基本心理能力測驗即為一例，進行文字系列完成測驗可以測試受試者解決問題和歸納推理的能力（Wriessnegger，2002）。

系列完成測驗最早由 Thurstone 於 1941 年運用在測量智力上，而 Simon 與 Kotovsky 在 1963 即運用 Thurstone Letter Series Completion task 發展出一套系列完成測驗的規則用以描述各種系列問題，並根據各系列的描述（description）判斷系列的難度。表 2-5-1 是 Simon 與 Kotovsky 採用的文字系列完成測驗之測驗題目及其描述格式。

表 2-5-1 文字系列完成測驗測驗題目及其描述格式 (Simon & Kotovsky, 1963)

題號	測驗題目 (test problems)	初始值 (initialization)	描述格式 (sequence iteration)
1	cdcddcd__	M1=(c,d);c	M1,N(M1)
		---	c,d
2	aaabbbcccdd__	M1=Alph;a	M1,M1,M1,N(M1)
		M1=Alph;a	M1(3),N(M1)
3	atbataatbat__	M1=(b,a);b	a,t,M1,N(M1)
4	abmcdmefmghm__	M1=Alph;a	M1,N(M1),M1,N(M1),m
5	defgefghfghi__	M1=M2=Alph;d	M1,N(M1),M1,N(M1),M1,N(M1), M1,N(M2),E(M1,M2)
6	qxapxbqxa__	M1=(q,p);q M2=(a,b);a	M1,N(M1),x,M2,N(M2),
		---	q,x,a,p,x,b
7	aduacuaeuabuafua__	M1=Alph;d M2=Balph;c	a,M1,N(M1),u,a,M2,N(M2),u
8	mabmbcmcdm__	M1=Alph;a	m,M1,N(M1),M1
9	urtustuttu__	M1=Alph;r	u,M1,N(M1),t
10	abyabxabwab__	M1=Balph;y	a,b,M1,N(M1),
11	rscdstdetuef__	M1=Alph;r M2=Alph;c	M1,N(M1),M1,M2,N(M2),M2
12	npaoqapraqsa__	M1=Alph;n M2=Alph;p	M1,N(M1),M2,N(M2),a
13	wxawybyzczadab__	M1=Alph;w M2=Alph;a	M1,N(M1),M1,M2,N(M2)
14	jkqrklrslmst__	M1=Alph;j M2=Alph;q	M1,N(M1),M1,M2,N(M2),M2
15	pononmnmmlmk__	M1=M2=Balph;p	M1,N(M1),M1,N(M1),M1,N(M2),E(M1,M2)

註：Alph = alphabet; Balph = alphabet backwards.

表 2-5-1 中初始值欄位說明了每個序列 (sequence) 的起始字母及字母表上的順序，M1,M2,... 等變數代表起始字母，描述格式欄位中的 N(M1)代表變數 M1 設

定至下一個字母， $E(M1, M2)$ 則是使 $M1$ 等於 $M2$ 。

Simon 與 Kotovsky 發現了一個簡易的判別難度規律，較簡單的題目通常其描述 (description) 也較簡單，題目的難度幾乎可以由該題目描述的長度來判斷。但第 6 題則是例外，此題有兩種描述方式，而其難度與這兩種描述方式皆不符合。Simon 與 Kotovsky 還找到另一個可能影響難度的要素，就是該題目所佔用的立即記憶 (immediate memory) 空間，在該次測驗中，所有難度較高的題目都由 2 個分別的串列 (list) 所組成，受試者在推理該類問題時需同時處理 2 個分別的串列，相對的，難度較低的題目則皆為單一串列。Simon 與 Kotovsky 也懷疑題目的週期 (period) 長度可能是影響難度的根本要素，但此假設與當時的實驗結果不甚符合，因此 Simon 與 Kotovsky 認為比起題目的週期長度，題目所佔用的立即記憶空間更與難度有關。Simon 與 Kotovsky 並試圖解釋當時實驗中第 9 題難度遠高於其「描述」的預測，他們認為其原因是題目中有假造 (spurious) 的關係，題目中有些相同的字母其彼此之間卻沒有關係，這樣容易使受試者產生混淆。

Holzman 與 Pellegrino 在 1983 年也針對系列完成測驗進行過實驗，Holzman 等人基本上沿用 Simon 的文字系列完成測驗描述格式，但將其稍作修改以適用於數字系列完成測驗，其實驗的主要內容為探討週期長度 (period length)、模型描述長度 (pattern description length) 與工作記憶單位 (working-memory place keeper) 對於系列完成測驗的難度影響，及研究中等程度智商 (average-IQ)、高智商 (high-IQ) 及成人在數字系列完成測驗的表現。Holzman 等人發現，在週期長度、

模型描述長度與工作記憶單位三個因素之中，影響受試者表現最明顯的是工作記憶單位，而工作記憶單位實質上與 Simon 與 Kotovsky 研究中的「立即記憶空間」類似，此兩研究的結果也大致相符。

Butterfield 等人在 1985 年提出了另一個描述文字系列完成測驗的系統，並針對各種可能影響文字系列完成測驗難度的預測變數 (predictor variable) 作實驗，Butterfield 認為文字系列由一個或一個以上的字串 (string) 組成，他並運用 N、I 及 B 三個關係來代表字串中的關係 (relation)。

表 2-5-2 為 Butterfield 用來描述並產生文字系列的範例，N、I 及 B 為 Butterfield 用來代表 next、identical 及 back 的關係，next 為字母表中的下一個字母，identical 為字母表中的同一個字母，而 back 則為字母表中的上一個字母。以 Rule 1 而言，該規則是由兩個字串所組成，分別是 N1 I1 與 B2 I2，文字系列就由這兩個字串組合起來的規則不斷循環產生，每個循環的起始字母都被預設為上一個循環的最後一個字母。此規則中兩個字串的週期長度 (period) 皆為 2，整個規則的週期長度為 $2+2=4$ 。

表 2-5-2 描述並產生文字系列的範例 (Butterfield 等, 1985)

Rule1	
	N1 I1 B2 I2
String1:	DD EE FF
String2:	YY XX WW
Problem:	DDYYEEXXFFWW
Rule2	
	N1 B1 N1 = N1 N2 I1
String1:	CBCDCDEDE String1: C CD DE E
	String2: B C D
Problem:	CBCDCDEDE = CBCDCDEDE
	N1 B1 N1 = N1 N2 I1
String1:	CBCDCDEDE String1: C CD DE E
	String2: L M N
Problem:	CBCDCDEDE ≠ CLCDMDENE
Rule3	
	N1-first N1 N1 I2
String1:	CDE DEF EFG
String2:	X X X
Problem:	CDEXDEFXEFGX
Rule4	
	N1 N2(+N) I2 I1
String1:	C CD DE E
String2:	NN OPP QRSS
Problem:	CNNCDOPPDEQRXXE

Rule 2 舉出了一個有趣的例子，它證明了兩個相異的規則有可能產生同樣的文字系列，在例子中由第一個規則所產生的所有文字系列皆可以由第二個規則來描述，但由第二個規則產生的文字系列卻不一定可由第一個規則來形成。

Rule 3 中 -first 這個關係說明了在字串的每個循環，其起始字母是可以被指定的，在之前的例子中，所有循環的起始字母皆為上一個循環的最後一個字母，在

Rule 3 中，N1-first N1 N1 的起始字母不是上一個循環的最後一個字母，而被指定為上一個循環的第一個字母，即上個循環的 N1-first 字母，若去除-first 關係，String1 應為 C D E F G H I J K，因為 F G H 這循環以 C D E 中的 E 為起始字母，在加上-first 後，第二個循環變成 D E F，因為其起始字母被改變為 C。

Rule 4 說明了週期長度並非固定，+N 這個關係代表在每次循環中，規則中都要多增加一個關係 N，N2(+N) I2 實際表現出來的形式為

N2 I2 N2 N2 I2 N2 N2 N2 I2...

每次循環其循環長度都會增加 1，再舉個例子來說，現有一規則如下：

N1(+I) B2 I2 I2 I2(-I)

將會產生如下的文字系列（文字系列間的空白是用來明顯區隔每個循環，實際上規則並不會產生空白）：

A Z Z Z Z B B Y Y Y C C C X X D D D D W（後面接下行）

E E E E E F F F F F F G G G G G G G...

值得注意的是，從 E E E E E 這個循環開始，string2（即 B2 I2 I2 I2(-I)）已經對

此文字系列完全不具影響力了，若將此文字系列寫成規則，將顯示如下：

N1 B2 I2 I2 I2 N1 I1 B2 I2 I2 N1 I1 I1 B2 I2 N1 I1 I1 I1 B2

N1 I1 I1 I1 I1 N1 I1 I1 I1 I1 I1 N1 I1 I1 I1 I1 I1 I1 I1...

僅包含關係 I 的字串被稱作非移動性字串（nonmoving string），例如 I1 I2 I3，但

關係 B 及 N 也有可能組合出非移動性字串，例如 B1 N1，而不屬於非移動性字串

者則為移動性字串 (moving string)。

表 2-5-3 文字系列的知識層級及範例 (Butterfield 等,1985)

1.由相同字母組成之字串。(由非移動性字串組成)	
1. OEIOEIOEI	2. KKTkKTKKT
2.由數量相等且鄰近的同樣字母組成的移動性字串。	
1. EEUFFUGGU	2. VVVWWWWWXXX
3.由數量不相等且鄰近的同樣字母組成的移動性字串。	
1. VVVWWWWWXY	2. STTUVVWXXX
4.由數量相等但非鄰近的同樣字母組成的移動性字串。	
1. MXMNXNOXO	2. RRBRQQCQPPDP
5. 由數量不相等且非鄰近的同樣字母組成的移動性字串。	
1. ABJBCJCDJ	2. AFABBGBCCCHC
6.只有在所有的字串其規則都被對應到相關的位置上時才可以完成的字串。 (字串中包含假造相同的字母)	
1. IHGHGFGFE	2. QNPOMNMLL

表 2-5-3 為 Butterfield 等所建立的文字系列知識層級，由易至難分為 6 級，由非移動性字串所組成的文字系列因最容易被受試者完成，故被分類到最簡單的第 1 級，而對於包含假造相同字母的問題，無論其工作記憶需求 (working-memory demand) 為多少，對受試者而言均特別困難，故被分類到最難的第 6 級。Butterfield 等綜合以上兩個事實，說明了在文字系列完成測驗中，根據題目中相同的字母是在同一字串裡或是分佈於不同字串中，將可以有效決定該題目的難度。

Butterfield 等進行了多個實驗研究各種預測變數來探討文字系列完成測驗的難度，有關 Butterfield 實驗中採用的預測變數說明如下：

1. KN (Knowledge): 參照表 2-5-3 文字系列的知識層級檢查文字系列，

觀察文字系列符合表中的哪個層級，該文字系列的 KN 值即為該難度層級。

2. #ML (Number of Memory Lists): 在文字系列中，移動性字串的數量。
以 RJFRKERLD 這個文字系列為例，其規則為 I1N2B3，則此字串的 #ML 值為 2，一個移動性字串為 N2，另一個為 B3，I1 不是移動性字串，故不計入 #ML 中。
3. CNT (count): Butterfield 等認為，人們必須找到適當的字母才能完成文字系列完成測驗的題目，而要找到該字母的方法則是反向推論 (working backwards)，即在文字系列完成測驗的題目中，後面的字母皆是由前面的字母衍生而來，算出要完成文字系列問題所需往回推的次數將能成為有效的指標。

文字系列 1: N1 I1 I1 BBCCDDDD count=1

文字系列 2: B1 N2 I2 KRRJSSITT count=2+2=4

文字系列 3: B1 B2 I1 MFMLELKDK count=2+3+1=6

文字系列 4: N1 B2 I3 AHPBGPCFP count=3+3+0=6

在同一字串中，若有關係 I 與前面的關係相鄰，則視為同一個單位 (unit)。觀察文字系列 1，D 欲回溯到上一個循環只需走一步即能連接到 C，故文字系列 1 的 CNT 為 1。再觀察文字系列 2，文字系列中的 SS 被視為同一個單位 (RR 與 TT 亦同理)，故 T 欲反推回 S 需走

2步 (I 與 SS), I 要反推回 J 也需走 2 步 (SS 與 J), 故文字系列 2 的 CNT 為 $2+2=4$ 。文字系列 3 的最後一個循環為 KDK, 第三個 K 要往回走 2 步 (D 與 K), D 要往回走 3 步 (K、L 與 E), 第一個 K 只要走 1 步 (L), 故文字系列 3 的 CNT 為 $2+3+1=6$ 。文字系列 4 的最後一個循環為 CFP, 其中 P 屬於非移動性字串, 故不計算在內, 而 C 與 F 皆必須往回走 3 步, 故文字系列 4 的 CNT 為 $3+3+0=6$ 。CNT 的值除了直接計算以外, 尚可用以下公式計算:

$$\text{Counts}=\#\text{MS}(\text{P}-\#\text{AIR})$$

#MS: 規則中的移動性字串數目

P: 週期長度

#AIR: 緊鄰於同字串其他關係的關係 I 數目

4. REL (relation): 給定關係 I 一個值為 0, 關係 N 的值為 1, 而關係 B 的值為 2, 根據文字系列的規則即可得到 REL 值, 例如 N1B2I1 的 REL 值為 $1+2+0=3$, N1N1B2B2I3 的 REL 值為 $1+1+2+2+0=6$, 但 N1 B1 B2 B3 的 REL 為 4, 因為 N1B1 為非移動字串, 不計在內, 故 REL 為 $2+2=4$ 。
5. SI (spurious identities): 文字系列中假造相同字母的數量。
6. Adjacency: 若整個文字系列中都沒有移動性字串, 則 Adjacency 為 0。若文字系列中有移動性字串, 但移動性字串的字母都彼此相鄰, 則

Adjacency 為 1。若文字系列中有任一移動性字串的字母沒有全部相鄰，則 Adjacency 為 2。

7. Equality：若整個文字系列中都沒有移動性字串，則 Equality 為 0。
若文字系列中有移動性字串，但每個移動性字串的字母數目都一樣，則 Equality 為 1。若文字系列中有移動性字串，但移動性字串的字母數目沒有全都一樣，則 Equality 為 2。

為了驗證影響文字系列完成測驗難度的因素，Butterfield 等進行了一系列的研究，針對許多預測變數（predictor variable）作實驗，實驗的結論簡述如下：

1. 當 KN、CNT 與 REL 等預測變數都被操控之後，週期長度與#ML 對難度將沒有影響力。
2. 較年長的兒童答對率明顯高於較年幼的兒童，但是當 CNT 與 REL 上升，所有兒童的答對率都會下降。
3. Simon 與 Kotovsky 所採用的預測變數#ML 在 Butterfield 等所做的實驗中對難度皆有顯著的影響力，但若加入 CNT 與 REL 這兩個預測變數一起分析，則#ML 所能解釋的難度部分將會完全被 CNT 與 REL 取代。
4. 完全由非移動性字串所組成的文字系列則格外簡單，假設有一名兒童，他無法答對所有包含移動性字串的文字系列完成測驗題目，他仍然有可能答對不包含移動性字串的文字系列完成測驗題目。
5. 包含有假造相同字母的文字系列完成測驗題目，難度將會明顯地變難。

6. 若運用 CNT、REL、SI、Adjacency 與 Equality 等 5 個預測變數應能很良

好的預測文字系列完成測驗的難度，但此結論並未經過直接實驗證明，

純粹是綜合數個實驗結果進行推測後的產物。

Butterfield 等的研究皆為分析現有的系列完成測驗題目並估計其難度，若以
虛擬題庫系統配合其研究，將有可能根據各種預測變數與規則直接產生指定難度
的題目。

欲增進測驗的預測與診斷功能，必須著重試題的鑑別度分析（簡茂發，
1987），Holzman 等雖將受試者分為 3 組，但其分析的因素只有週期長度、模型描
述長度與工作記憶單位三種，且週期長度與模型描述長度對於試題的影響並不
大，若能詳細分析各種預測變數與文字系列完成測驗題目鑑別度的關係，將能更
有效的編製文字系列完成測驗。