

錯誤行為分析 -- 以國立台灣大學之終端使用者為例 Error Analysis of End users in Online Searching : A Case Study of National Taiwan University

黃慕萱 Mu-hsuan Huang

國立台灣大學圖書館學系副教授

Associate Professor Department of Library Science, National Taiwan University

[摘要]本文以Dialog系統之書目資料庫為檢索系統，檢索者的來源則為本身具有資訊需求的國立臺灣大學教職員生。本文擬對線上資訊尋求行為進行系統化的分析，其目的在探討檢索者線上檢索時之錯誤行為及其發生位置，分析零筆資料的原因，了解時間經驗對錯誤行為的影響，並比較中美終端使用者線上資訊尋求行為之異同。

[Abstract]The purpose of this article is to analyze the searching behavior of end users. The subjects are real end users from National Taiwan University. This article • uses the DIALOG command language as its searching tool. It attempts to investigate reasons for errors and location of errors. In addition it tries to identify reasons for zero posting and compare the behavior difference between Chinese and American end users.

關鍵詞：線上檢索；錯誤；資訊尋求行為

Keywords : Online searching; Errors; Information seeking behavior

一、問題陳述

• • 從錯誤行為來分析線上資訊尋求行為，是研究線上資訊尋求行為一個非常好的方向。一般而言，錯誤分析的前提是錯誤可以反映出檢索者心智模型（mental model）和實際行為之間的差異，是透析檢索行為本質的最佳工具之一。然而，心智模型非常不容易研究，因為其無法透過直接觀察，必須經由間接的證據來推論。事實上，推測檢索者心智模型的最好來源是檢索者在使用系統時所犯的錯誤，因為錯誤可以反映出使用者當時的心智模型，同時也指出了介於使用者模式與真實系統模式間的矛盾與不一致。（註1）例如Lewis和Mack發現，文字編輯的錯誤反映了使用者過去操作打字機的情形。（註2）根據目前的研究看來，大部分有關終端使用者和資訊中介者的檢索行為研究，大多只調查了移動• move •的狀態，很少有關於錯誤方面的分析，這

正是本文系統化地分析錯誤行為的重要原因。

- 本研究是以Dialog系統的書目資料庫作為檢索系統，檢索者的來源則是本身具有資訊需求的國立台灣大學學生和教職員。在本研究中，資料分析的來源是檢索策略及結果（search transcripts），研究者仔細閱讀檢索策略及結果，根據上下文顯示內容及檢索者對資料庫檢索之專業知識，推得錯誤原因；換言之，本研究探討終端使用者在線上檢索時最常犯的錯誤，其目的在深入了解國人之線上資訊尋求行為，所得研究結果可視為設計線上互動系統及線上救援系統最好的參考來源之一。

二、研究設計和研究對象

- 本研究利用線上電子佈告欄（Bulletin Board System, BBS）及人工張貼佈告的途徑，於國立台灣大學校園內公開徵求資訊需求者。在一學年內，一共招募了47位願意參加本研究之資訊需求者，總共進行了60次檢索，也就是說，進行2次或多次檢索之實驗對象並不多。一般而言，資訊需求者在第一次檢索前，除其曾在其他地方受過Dialog使用訓練外，都須接受研究者半小時的使用指導。為了讓所有的說明一致，研究者必須根據其所撰寫之Dialog檢索使用說明，逐步介紹Dialog系統之檢索方式，以控制因訓練方式不同所產生的變因。而後，這些實驗對象可以在整個學年內無限次數地使用Dialog資料庫，檢索與他們的論文或是學期報告相關的資料。同時，研究者將會適時地提供實驗對象一些資訊，例如建議檢索者使用哪些資料庫及解釋指令用法等，但研究者只能在檢索進行前後提供這些建議，絕不會在檢索進行中提及這些資訊。根據實際觀察及閱讀檢索策略及結果發現，一般而言，檢索者多能自行設計檢索策略。

- 基本上，檢索指導的內容包括：布林邏輯與相近運算元的概念、主題、作者與題名檢索的操作、六個基本指令的使用（包含begin、select steps、display sets、display、type及logoff）、檢索策略的設計方式[以分區組合檢索（building • block）為主]、切截的介紹、以及用一個檢索實例說明Dialog系統之檢索方式。在半小時的訓練中，要包含上述所有內容實在不太容易，但因冗長的說明會降低檢索者參與研究計畫的意願，所以無法增加檢索說明的長度。也就是說，由於受到時間因素的限制，本研究中之檢索指導，均只提供基本的程序訓練（procedural training），而不是概念的訓練（conceptual training）。

- 在本研究中，每一位實驗對象必須填寫二種問卷，第一種問卷的目的是調查檢索者的個人背景（如年級和學科背景等），了解檢索者的檢索經驗及電腦經驗等。上述資料（年級、學科背景、電腦及線上檢索經驗等）均被視為影響線上檢索行為的潛在變因。由於Kuhlthau所做的研究證實了檢索問題的文字性陳述，可以有效地反映出檢索者在某一特定時間上的思考重點與思路變化。[\(註3\)](#)因此在每一次檢索前，檢索者還要填寫一份簡要的問卷，他們

必須以自然語言陳述檢索問題、說明已掌握到那些相關資訊、預測檢索到相關資料的可能性、以及描述目前已進行到那一個研究階段等，希望藉此掌握檢索問題的特性。

- 本研究一共有47名檢索者，其中男性31名（佔66%），女性16名（佔34%），這可能是台大男性教職員生多於女性教職員生的一種自然反應。在這47名檢索者中，其中有43位是研究生（91.5%），包括碩士班學生29人（61.7%）和博士班學生14人（29.8%），其他族群則包括大學部學生2人（4.2%），教師及研究助理各1人（各佔2.1%）。許多研究顯示，線上終端使用者的最大族群是研究生，本研究中研究生所佔的比例高達91%，和國外的研究結果不謀而合。若以實驗對象的學科背景來看，檢索者分別來自四個學院，其中以工學院之檢索者最多（23人，48.9%），文學院10人（21.3%），農學院7人（14.9%），理學院6人（12.7%），管理學院則僅有1人（2.1%）。若以系所為單位進行更深入的分析，可以發現檢索者以來自圖書館學系的10人（21.3%）和材工所的10人（21.3%）較為集中，其餘系所都十分分散，除化工系3人（6.4%）外，其餘系所之檢索人數皆僅有一至二人。造成圖書館學系及材工所檢索人數較多的原因，可能和這二個系所之授課教師鼓勵學生充分利用Dialog檢索系統有關，因此，研究中若能掌握數位願意配合的老師，研究對象的參與程度可能會大為提高。

- 由於掌握檢索者對電腦工具及各種檢索系統的經驗，有助於了解檢索者對Dialog系統的應用程度，因此本研究以問卷法蒐集此方面之背景資料。一般而言，47名檢索者中，半數以上均表示熟悉文書處理軟體（26名，55.3%），但對於試算表、應用軟體、及程式設計則較不熟悉。同時，本研究中絕大多數之檢索者，都曾使用過光碟檢索系統和線上公用目錄系統（Online Public Access Catalog, OPAC），其中使用過光碟資料庫之檢索者（40位，85.1%）比用過線上公用目錄者（34位，72.3%）為多，但在使用次數方面，檢索者使用線上公用目錄的平均次數（11.11次）則較其使用光碟資料庫的平均次數（5.61次）高出甚多。至於Dialog系統之檢索經驗，本研究中大部分之檢索者未曾使用過Dialog系統，僅有13人過去使用過此系統（27.7%），而其中除一名化工系教授較常使用國際百科系統外，其餘曾使用Dialog系統之檢索者對該系統均不太熟悉，這很可能是資訊需求者不願意進行多次檢索之主要原因。

三、錯誤行為的原因分析

- 在本研究所蒐集之60次檢索中，總計產生417次錯誤行為，實驗對象在每次檢索中平均犯下7.07個錯誤，錯誤次數則介於1到21個之間。表1顯示大部分檢索者所犯的錯誤多介於1到9個之間（70%），其中產生0至4次錯誤者計有20次檢索（33.3%），5至9次檢索者計有22次檢索（36.7%），至於錯誤數目介於10次至14次之檢索有12件（20.0%），介於15次至19次之間的則僅有4件（6.7%），超過20次錯誤的檢索也僅有1件（1.7%）。一般而言，影響錯誤數目的原因相當多，例如檢索時間的長短（檢索時間越長，所犯的錯誤可能越多）、檢索者對指令語言的熟悉程度、檢索問題的主題與特性、以及學習效

果等。雖然上述原因都是可能影響檢索者犯錯數目的因素，但由於時間精力上的限制，這些因素並未在本研究中加以探討。

表 1：錯誤行為發生之次數分析表

錯誤次數	發生次數	百分比
0-4	20	33.3
1	6	10.0
2	2	3.3
3	5	8.3
4	7	11.7
5-9	22	36.7
5	8	13.3
6	7	11.7
7	2	3.3
8	3	5.0
9	2	3.3
10-14	12	20.0
15-19	4	6.7
20-24	1	1.7
無資料	1	1.7
總計	60	100.0

• • 由於錯誤次數與檢索長度有很大的關係，所以如能將一個檢索區分成數個循環（cycles），可以更加了解時間和經驗對錯誤次數的影響。一般而言，循環的定義是在讀者列印（print）或顯示（display）結果（可能是連續列印或顯示）之前所下的所有連續指令。舉例來說，一個循環可能由選擇資料庫（在Dialog中的指令是begin），或選擇檢索詞彙（在Dialog中的指令為select或select steps）的指令開始，結束於顯示（包含全螢幕顯示及連續顯示）的指令中。圖1是一個典型的循環：

圖一 • 循環實例

• • •

```

b 35

ss deed? and proper?

type s3/5/all

```

• •

• • 表2顯示錯誤所在的檢索序數，由表2可以看出超過95%（403次，96.7%）的錯誤是發生在檢索者的第一次檢索和第二次檢索中，其中發生於第一次檢索的錯誤高達351次（84.2%），發生於第二次檢索的錯誤則減為52次（12.5%），而發生於第三次檢索及第四次檢索的錯誤次數則微不足道（只有14次，3.4%）。由此可知，在前幾次檢索中，尤其是第一次檢索，終端使用者對指令的不熟悉和線上經驗不足，因此比較容易有錯誤行為發生。

表2：錯誤行為所在之檢索序數表

檢索序數	錯誤的次數	百分比(%)
1	351	84.2
2	52	12.5
3	13	3.1
4	1	0.2
總計	60	100.0

• • 表3顯示檢索者錯誤所在的循環序數，由表中可以看出，95的錯誤都發生在前5個循環中，其中又以第一次循環所發生的錯誤次數最多，高達204次（48.9%），第二次循環則有72次錯誤（17.3%），第三次循環有67次錯誤（16.1%），第四和第五次循環的錯誤次數分別為31次（7.4%）和22次（5.3%），第六次及其以後循環發生錯誤的次數均在10次以下。由此可知，錯誤很明顯地集中在前幾個循環中，顯示經過幾個循環的學習後，檢索者所犯的錯誤次數會逐漸下降。換言之，在前幾次循環中，由於終端使用者對指令的不熟悉和線上經驗不足，因此比較容易產生錯誤行為。

表3：錯誤行為所在之循環序數表

第幾次循環	錯誤的次數	所佔百分比(%)
1	204	48.9
2	72	17.3
3	67	16.1
4	31	7.4
5	22	5.3
6	8	1.9
7	6	1.4
8	2	0.5
9	2	0.5
10	2	0.5
11	1	0.2

總計

60

100.0

• • 在討論任何錯誤行為之前，必須先了解錯誤發生的原因。從檢索者的檢索策略及結果中，研究者根據上下文及其對資料庫檢索的專業知識，總共找出了80種錯誤原因，其中28種錯誤原因只發生過1次（35%），17種錯誤原因只發生過2次（21.3%），12種錯誤原因曾發生過3次（15%）。換言之，超過7成的錯誤原因其出現次數均為3次或3次以下，由於錯誤的原因相當分散，表4僅列舉9種犯錯次數超過10次錯誤原因。由表4可以得知，檢索者最常犯的錯誤是「應用自然語言檢索時，卻使用控制語言來檢索」，此過錯發生的頻率高達47次，約佔所有錯誤的十分之一（11.3%）。排名第二位的錯誤行為是「一次開啟多個資料庫時，沒有利用 "rd" 除去重複的資料」，此錯誤發生46次，約佔所有錯誤的十分之一（11.0%）。排名第三的錯誤行為是「沒有善用前面檢索過程中產生的檢索組號碼，再次輸入相同的檢索詞彙」，這樣的錯誤發生31次（7.4%）。而排名第四者則為「將筆數很少的檢索結果和另一個檢索限制交集，因而導致零筆資料」（30次，7.2%）。此外，「檢索詞彙拼錯或輸入有誤」亦出現26次（6.2%）之多，「誤用鍵盤右邊的修正鍵修改」則發生22次（5.3%），雖說檢索訓練中曾提及僅能使用回格鍵（Del）鍵修改資料，而「重複輸入相同的檢索策略或檢索敘述」亦造成14次錯誤（3.4%）。至於其他二種發生頻率較高的錯誤行為皆與資料庫的限制有關，其中「資料庫不提供選定的後置欄位檢索」有17次（4.1%），而「該檢索組合在資料庫中原本即不存在」亦出現10次（2.4%）。

表4：最常發生之錯誤原因表

錯誤次數	百分比	錯誤原因
47	11.3	應用自然語言檢索時，卻使用控制語彙檢索。
46	11.0	一次開啟多個資料庫時，沒有利用 " rd " 除去重複資料。
31	7.4	沒有善用前面檢索過程中產生的檢索組號碼，再次輸入相同的檢索詞彙。
30	7.2	將筆數很少的檢索結果和另一個檢索限制交集，因而導致零筆資料。
26	6.2	檢索詞彙拼錯或輸入有誤。
22	5.3	誤用鍵盤右邊的修正鍵修改。即系統不接受使用鍵盤最右邊的修正鍵，雖然在螢幕上會出現修改過的情況，但實際上卻是修改無效。
17	4.1	資料庫不提供選定的後置欄位檢索。例如：/maj . /k . /ti . /id.
14	3.4	重複輸入相同的檢索策略或檢索敘述。
10	2.4	該檢索組合在資料庫中原本即存在，並非檢索者指令錯誤或策略不佳所致。

• • 由上述常犯的錯誤可以看出，使用者所犯的錯誤大多與認知負擔超載（cognitive overload）或檢索知識不足有關。由於實驗對象大多是初學者或線上檢索經驗不多的檢索者，其資訊處理容量通常不足，為了減低學習與認知上的負擔，傾向以花費最少的精力和負擔來完成檢索，因此沒有善用前面過程中所產生的檢索組號碼、拼錯字或按錯鍵、重複輸入、或是誤用鍵盤右端的修正鍵，都是很自然的外顯現象。至於檢索知識不足所導致的錯誤，例如自然語言和控制語言的區別、除去重複資料、將筆數很少的檢索結果和另一個檢索限制交集、和一些較為特殊的後置欄位檢索，都是屬於較具深度的檢索知識，初學者在使用上發生困難，也是一種相當普遍的現象。

四、錯誤行為發生的位置

• • 本研究企圖了解錯誤所在的功能（functions）、指令（commands或moves）及要素（elements），也就是說，分析每一個錯誤行為所發生的位置。首先探討錯誤行為發生在哪些功能上。本研究所使用的功能分析表，是沿用筆者在博士論文中所發展出的功能分析表。功能可以定義為一個或數個（通常為一個）指令所欲完成的目的，筆者在博士論文中所使用的功能共有七種，其中包含無效功能（null function）(註4)。無效功能是指無法辨識其所屬功能的指令，就像檢索者鍵入 a 或 ss type 這種奇怪的指令。表5顯示超過四分之三的錯誤行為發生在建立與發展檢索組（create and develop sets）上（317次，佔76.0%），而發生在顯示資料上的錯誤行為較預期的少很多（28次，佔6.7%）。值得一提的是，中立指令所產生的錯誤高達52次（12.5%），其中絕大多數都發生在除去重複資料的指令上。由於研究者適時說明除去重複資料的重要性，因而造成此指令的高使用率，因之也產生較多的錯誤行為。其他功能所發生的錯誤行為均不多，分別為選擇資料庫8次（1.9%）、無效功能6次（1.4%）、尋求協助5次（1.2%）及離線1次（0.2%）。這可能和上述功能在檢索中出現的比率較低及檢索者大多經驗不足或是初學者有關，他們不知道有這些功能或指令可以使用，就算知道這些功能，他們也沒有足夠的資訊處理能力去使用這些功能，因此錯誤發生在這些功能的次數也相當低。

表5：錯誤行為所在之功能表

錯誤次數	百分比	錯誤行為所在之功能及副功能	
8	1.9	功能1	選擇資料庫
7	1.7	1.1	選擇資料庫
1	0.2	1.2	更改資料庫

317	76.0	功能2	建立與發展檢索組
48	11.5	2.1	輸入檢索組號碼，非直接輸入檢索辭彙
148	35.5	2.2	直接輸入檢索辭彙
7	1.7	2.3	非主題欄位的檢索（如：la、py、an）
81	19.4	2.4	輸入檢索組號碼及辭彙
3	0.7	2.5	檢索辭彙及非主題欄位
28	6.7	2.6	檢索組號碼及非主題欄位
2	0.5	2.7	檢索組號碼、檢索辭彙及非主題欄位
28	6.7	功能3	顯示或列印
1	0.2	3.1	顯示或列印完整紀錄的所有資料筆數
14	3.4	3.2	顯示或列印部分紀錄的所有資料筆數
4	1.0	3.3	顯示或列印完整紀錄的部分資料筆數
9	2.2	3.4	顯示或列印部分紀錄的部分資料筆數
5	<1.2	功能4	尋求協助
2	0.5	4.1	回顧檢索過程
3	0.7	4.2	找尋相關的檢索辭彙
52	12.5	功能5	中立指令
1	0.2	5.2	儲存檢索步驟
51	12.2	5.3	除去重複資料
1	0.2	功能6	離線
1	0.2	6.1	離線
6	1.4	功能7	無效功能
417	100.0		總計

• • • • 如果將每一個主要功能區分成副功能，表5顯示在「建立與發展檢索組」上，以「直接輸入檢索詞彙」所導致的錯誤最多（148次，佔35.5%），其次是「同時使用詞彙與檢索組號碼」（81次，佔19.4%）。在顯示或列印資料中，一半的錯誤行為（14次，50%）是發生在「顯示或列印部分紀錄的所有資料」之副功能中，此可由檢索指令分析中看出端倪（格式7是檢索者的最熱門選擇，其屬於部分紀錄，而顯示全部範圍者在所有顯示列印指令中佔52.0%）。此外，由功能分析和副功能分析可以看出，中立指令所產生的錯誤行為次數高於顯示資料所產生的錯誤行為，這是非常奇怪的現象。一般而言，Dialog顯示與列印指令之設計過於複雜，初學者很難將其操作自如，但實驗對象卻出人意外的將其操作的非常好，這是非常值得深入探討的現象。值得一提的是，除去重複資料之指令在指令分析中只出現28次，但其產生的錯誤次數卻高達51次，這是因為檢索者在應該使用「remove • duplicates」指令時卻未使用該指令（典型錯誤原因如「一次開 • 多個資料庫時，沒有利用 • rd • 除去重複的資料」，或是「顯示完全或部分的資料後，才進行除去重複資料（rd）」等），因此rd指令發生的次數比錯誤次數少23次之多。

• • 為了解指令使用頻率及其錯誤次數的關係，表6將指令被使用次數及錯誤行為所在的指令（moves）顯示於同一表格中。和功能分析一樣，以發生於選擇檢索詞

彙的錯誤最多，高達317次，超過所有錯誤行為的四分之三。進一步分析指令被使用的次數和其產生錯誤的次數，選擇檢索詞彙出現的次數佔所有指令之56.7%，而其發生之錯誤次數卻高達76.0%，顯示在透過詞彙表達檢索策略的資訊系統中，選擇檢索詞彙的確是讀者線上檢索最大的困難來源。其次為與除去重複資料指令有關的51次錯誤，由於在功能分析中已作過非常詳細的描述，在此不再贅述。在顯示資料上，總計發生錯誤28次（6.7%），而此二指令（display和type）在所有指令的出現比率卻高達21.7%，顯示國人在顯示資料指令上所產生的錯誤遠較預期為少。至於其他指令所發生的錯誤次數均不多，但其在所有指令中出現次數皆佔一定比率，例如選擇或開啟資料庫70次（6.2%），但僅產生8次錯誤行為（1.9%）；線上回顧檢索策略或檢索組號碼85次（7.6%），僅產生2次錯誤行為（0.5%）；離線55次（4.9%），但僅產生1次錯誤行為（0.5%），充分透露出國人操作上述指令的能力，或許實驗對象在操作上述指令上仍有一些困難疑慮，但卻不致於產生錯誤行為。

表6:錯誤行為所在之指令表

• • • • •

錯誤行為所在之指令	發生次數	百分比	錯誤次數	百分比
選擇或開啟資料庫(begin)	70	6.2	8	1.9
選擇檢索詞彙(select steps or select)	636	56.7	317	76.0
全螢幕顯示(display)	28	2.5	8	1.9
連續顯示資料(type)	216	19.2	20	4.8
線上顯示相關詞彙(expand)	5	0.4	3	0.7
線上回顧檢索策略或檢索組號碼(display set)	85	7.6	2	0.5
將檢索結果儲存起來(save)	0	0	1	0.2
去除重複資料(remove duplicates)	28	2.5	51	12.2
離線(logoff)	55	4.9	1	0.2
無法判斷或無效的指令	0	0	6	1.4
總計	1123	100.0	417	100.0

• • 要素分析可以提供有關錯誤行為更深入的資訊，因為要素是本研究所採用之最小分析單位。要素在本文定義為指令（如選擇詞彙用的ss或開•資料庫用的begin等），或是它的變數部分（如詞彙或是作者的名字等）。舉例來說，"d • s1 / 5 / all" 包含六個要素，分別是 "d"，"s1"，"/"，"5"，"/"，和 "all"；例如 "ss • expansion?"，此指令則包含三個要素，分別是「ss」，「expansion」和「？」。表7顯示錯誤行為所在的要素。一般而言，發生於指令部分的錯誤行為較發生於檢索值部分的錯誤行為為少，而其中超過三分之一的錯誤行為（142次，34.1%）發生在檢索詞彙上，由此可知，讀者最大的問題來源還是在檢索詞彙的選擇上，到底是用哪一個詞彙最為合適，是用自然語言較好，還是使用控制語彙較為恰當，形成檢索者最大的錯誤來源。其次是發生在除去重複資料的錯誤，此種錯誤計有51次，佔12.2%，由於有關除去重複資料的錯誤在功能分析和指令分析上都有所敘述，此處不再贅述。排名第三名為發生在布林邏輯運算元"AND"上之錯誤行為，其發生次數為43次（10.3%），遠較相近運算元為多，雖說一般檢索者對布林運算元較

相近運算元熟悉，但由於布林運算元的使用次數較相近運算元多367次，其發生錯誤的機率自然相對增加。排名第四者為已建立的檢索組號碼（25次，6.0%），顯示檢索者在資訊處理容量不足時，會產生無法善用已建立之檢索組號碼或是誤用其他檢索組號碼的狀況。至於排名第五的要素則為選擇檢索詞彙（select • steps或select）指令本身（23次，5.5%），可能是由於檢索者忘了先行輸入該指令所導致。其他較常發生錯誤的要素包括「無效的後置欄位」（17次，4.1%）、「連續顯示指令」（10次，2.4%）及「切截符號」（10次，2.4%）等。

表7：錯誤行為所在之要素表

錯誤次數	百分比	錯誤行為所在之要素
<u>100</u>	<u>23.8</u>	指令部份
1	0.2	選擇或開啟資料庫(begin)指令
5	1.2	全螢幕顯示資料(display)指令
2	0.5	線上回顧檢索策略或檢索組號碼(display set)指令
51	12.2	除去重複資料(remove duplicates)指令
1	0.2	離線(logoff)指令
1	0.2	儲存檢索(save or save temp)指令
23	5.5	選擇檢索詞彙(select steps or select)指令
10	2.4	連續顯示資料(type)指令
6	1.4	無適當指令
<u>211</u>	<u>50.7</u>	指令的變數部份
<u>142</u>	<u>34.1</u>	檢索詞彙
2	0.5	前置欄位 au
4	1.0	前置欄位 • py
2	0.5	後置欄位 • df
1	0.2	後置欄位 • ab
7	1.7	後置欄位 • de
6	1.4	後置欄位 • ti
17	4.1	無效的後置欄位
5	1.2	前置欄位(prefixes)的值(不包括詞彙)
25	6.0	已建立的檢索組號碼(僅就建立與發展檢索組而言)
<u>57</u>	<u>13.7</u>	邏輯運算元部份
43	10.3	AND
3	0.7	OR
4	1.0	NOT
7	1.7	括弧
<u>10</u>	<u>2.4</u>	切截符號 (?)

3	0.7	相近運算元部份
2	0.5	相近運算元(W)
1	0.2	相近運算元中所容許的字數
10	2.4	標點符號
3	0.7	標點符號(逗點),
2	0.5	標點符號 • 等於符號(=)
3	0.7	標點符號 • 斜撇(/)
2	0.5	標點符號 • 完成鍵enter
15	3.6	數字部份
7	1.7	資料庫號碼
1	0.2	檢索組號碼(僅就列印而言)
2	0.5	列印格式號碼
5	1.2	文件號碼
1	0.2	無效部分 (null element)

五、有關錯誤行為的進一步分析

若進一步分析上述417次錯誤行為，表8顯示這些錯誤的屬性以「未查覺的錯誤最多187次，44.8%），表示相當多的錯誤是由於檢索者所建立之概要（schema）錯誤所導致；而「查覺的錯誤」居次，計發生104次，約佔所有錯誤的四分之一（24.9%），顯示還是有相當數目的錯誤能被檢索者當場查覺；屬於「較佳策略的錯誤」較預期為少，只發生94次（22.5%），其中絕大多數（91次，96.8%）都不會影響到檢索結果，僅有3.2%（3次）的錯誤會影響到檢索結果。上述未使用更佳策略的錯誤可能來自於未能善用前面檢索過程已產生的檢索組號碼，或是將比較多資料的檢索組號碼或檢索詞彙擺在檢索敘述的前半部，因此對檢索結果毫無影響，只是會花上無意義的線上等待時間，造成時間與金錢上的浪費，無法對檢索結果產生影響。發生頻率最低的錯誤是不小心的錯誤，比例只有4.1%（17次），可能是因為大部分不小心的錯誤，檢索者都能很快發現並在線上及時修正，所以屬於這方面的錯誤行為明顯偏低。

表8：錯誤行為的屬性

錯誤行為的屬性	次數	百分比
察覺的錯誤	104	24.9
未發覺的錯誤	187	44.8
不小心犯下的錯誤	17	4.1
較佳策略	94	22.5
其它	15	3.6
總計	417	100.0

至於和邏輯有關的錯誤，總計有11個（2.6%）。表9顯示，大部分的邏輯錯誤來自於"AND"與"OR"運用錯誤（4次，36.4%），其次是沒有使用適當的括弧（3次，27.3%），再其次是和零筆資料的檢索組號碼或檢索辭彙連結（2次，18.2%），其餘和邏輯有關的錯誤發生的比例相當低。這種和零筆資料的檢索組號碼或檢索辭彙連結的錯誤，很可能和檢索者的資訊處理容量不足有關，由於需要處理的東西過多，檢索者無法有效處理這些資訊。至於和括弧或是AND與OR應用有關的錯誤，很可能是概要（schema）上的錯誤，這方面的錯誤較可能是因為檢索者缺乏正確適當的概要所引起。

表9：邏輯錯誤之複分

邏輯錯誤之複分	次數	百分比
AND與OR運用錯誤	4	36.4
沒有給予適當的括號，但系統處理有誤	3	27.3
與零筆數的檢索組號碼或檢索詞彙連結	2	18.2
OR與NOT運用錯誤	1	9.1
OR運用錯誤	1	9.1
總計	11	100.0

此外，在上述錯誤行為中，零筆資料共出現183次（43.9%），也就是說，幾乎每五個錯誤中，就有二個屬於零筆資料的錯誤。在線上檢索或OPAC的研究中，有些研究錯誤行為的文章，尤其是有關OPAC方面的文獻，就常將零筆資料視為惟一的錯誤來源。在本計畫中，研究者共找出38種零筆資料發生的原因，其分布相當零散，其中8種原因（21.1%）只發生過2次，17種原因（44.7%）只發生過1次。因此，表10僅列出9種出現頻率超過5次的零筆資料原因，其中超過四分之一的零筆資料是因為「應用自然語言檢索時，卻使用控制語彙檢索」（47次，25.7%），即當檢索者心中用來描述檢索概念的詞彙與系統使用的敘述語不能互相吻合，檢索者若使用控制語言檢索，經常會得到零筆資料的結果，但若其使用自然語言檢索，則很有可能找出一些相關資料。發生次多的零筆資料原因是「將筆數很少的檢索結果和另一個檢索限制交集，因而導致零筆資料」，一般而言，當檢出資料筆數不多時，最為完善的策略是直接將結果顯示於螢幕上，與其他概念交集，即使沒有產生零筆資料，通常也會造成時間及金錢上的浪費。

表10：最常見零筆資料原因表

錯誤次數	百分比	錯誤原因
------	-----	------

47	25.7	應用自然語言檢索時，卻使用控制語彙檢索。
27	14.8	前面的連結檢索結果的筆數已經很少，檢索者又連結另一個檢索限制，因而造成零筆資料。
17	9.3	鍵盤的問題。即系統不接受使用鍵盤最右邊的修正鍵，雖然在螢幕上會顯現修正過的情況，但實際上卻是修正無效。
15	8.2	檢索詞拼錯或打字錯誤。
9	4.9	該檢索組合在資料庫中原本即不存在，並非檢索者指令錯誤或策略不佳所致。
7	3.8	再次輸入前面已使用過的檢索策略或檢索敘述。
5	2.7	由於資料庫本身並不儲存該檢索詞彙所造成之零筆結果，並非檢索者的錯誤。
5	2.7	進行作者檢索時(au =)，沒有利用切截來控制作者名稱的變化。
5	2.7	不會使用年代範圍檢索。

• •

至於和個人認知負擔有關的零筆資料原因則包含「誤用鍵盤右邊的修正鍵修正」(17次, 9.3%)、「檢索詞拼錯或輸入有誤」(15次, 8.2%)及「重複輸入相同的檢索策略或檢索敘述」(7次, 3.5%)，因為在資料處理容量不足的情況下，檢索者通常會比較緊張，因此比較容易按錯鍵，拼錯字或是忘記前面的檢索過程而重複輸入，上述錯誤皆可視為認知負擔不足的典型現象。

其他較常見的零筆資料原因可以其和檢索策略的關係一分為二，其中和檢索策略無關的兩種原因分別為「該檢索組合在資料庫中原本即不存在，並非檢索者指令錯誤或策略不佳所致」(9次, 4.9%)及「由於資料庫本身並不儲存該檢索詞，才造成零筆結果，並非檢索者的錯誤」(5次, 2.7%)，這二種錯誤嚴格說來並非檢索者的錯誤，尤其是當檢索者輸入不正確的敘述語時，讀者導向的系統設計均認為系統應能提供適當的轉換，幫助讀者找到適合正確的敘述語，而不是直接顯示零筆資料的結果，讓檢索者誤以為系統並沒有儲存這方面的資料。至於和檢索策略有關的零筆資料原因，則包含「進行作者檢索時，沒有利用切截來控制作者名稱的變化」及「不會使用年代範圍檢索」，這二種原因都和檢索者的檢索知識不足有關，由於其不知指令正確的使用方式，因此造成零筆資料。

前文曾根據錯誤所在的檢索序數和循環序數比較時間經驗和錯誤次數的關係，由於第一次檢索和第一次循環本身發生的次數較多，因此很難以錯誤所在的檢索序數和循環序數絕對論斷時間經驗對錯誤次數的影響，因此研究者特別找出60次檢索

及207次循環中每次檢索和每次循環所產生的錯誤次數，表11顯示第一次檢索平均發生7.47次錯誤（標準差是4.88），第二次及其以後檢索所發生的錯誤次數降為5.50次（標準差3.40），其差異雖然沒有達到統計上的顯著水準，但第一次檢索所發生的錯誤次數較第二次檢索平均高出1.97次（非常接近2次）。若以循環為單位來看，可以發現不管是第一次循環和其他循環（其中一次循環之錯誤次數為3.36次，第二次及其以後循環則降為1.48次），前二次循環及其他循環（其中前二次循環的錯誤次數2.49次，其他循環則為1.47次）及前三次循環及其他循環（其中前三次循環的錯誤次數為2.35次，其他循環為1.21次），都可以證實前面序數的循環明顯產生較多的錯誤行為，而且都到達統計上.001的顯著差異。上述數據再次顯示隨著時間經驗的增加，檢索者的檢索知識逐漸增長，檢索技巧的運用也更為靈活熟悉，因此產生的錯誤次數也會隨之降低。

表11錯誤次數在不同檢索序數及循環序數之差異表

• • • • •
• • • • •

錯誤次數	第一次檢索 7.47 (4.88)	其他檢索 5.50 (3.40)	F值 2.06	p值 0.20
錯誤次數	第一次循環 3.36 (3.33)	其他循環 1.48 (1.72)	3.76	0.001
錯誤次數	前二次循環 2.49 (2.83)	其他循環 1.47 (1.74)	2.64	0.001
錯誤次數	前三次循環 2.35 (2.72)	其他循環 1.21 (1.27)	4.61	0.001

• • • • •
• • • • •

六、結語 • • •

本文特將比較中美終端使用者錯誤行為的差異，研究發現在最常見的五種錯誤行為中，只有二種錯誤原因相同，一為「沒有善用前面檢索過程中產生的檢索組號碼」，另一為「應用自然語言檢索時，卻使用控制語言來檢索」，且其排名完全不同，由此可知國人錯誤行為的種類和歐美各國極可能有本質上的差異，充分透露出本土化研究的重要性。若比較錯誤行為所在的功能和指令，可以發現不管是國人或是歐美人士，都是以建立與發展檢索組功能（即選擇檢索詞彙指令）所產生的錯誤行為最多，但馬里蘭大學師生在此功能或指令上產生的錯誤佔43.5%[\(註5\)](#)，而國立台灣大學師生卻佔76%，顯示由於語文上的障礙，國人在選擇詞彙上所遭遇的困難最多，因此線上索引典在國內的發展空間應較國外為大。至於顯示列印指令或功能，造成馬里蘭大學師生14.4%的錯誤，但僅造成台灣大學師生6.7%的錯誤，因此可知在此複雜指令的操作上，國人的學習能力和操作能力較歐美人士為強。

- 要素分析也透露出中美終端使用者錯誤行為上的差異，對不同的實驗對象而言，同樣以檢索詞彙為最大錯誤來源，但其所佔比率不同（馬里蘭大學師生：台灣大學師生 = 24.8%：34.1%）[\(註6\)](#)，顯示台灣大學的師生對詞彙的掌握程度較馬里蘭大學為差。此外，馬大學生在相近運算元所發生的錯誤比率較高[\(註7\)](#)，而台大學生在布林運算元所發生的錯誤次數較多，除了反應使用較多的指令產生較多錯誤行為外，更顯示不管在布林運算元和相近運算元的推廣上，對國人而言，仍是有待加強訓練的基本輔助檢索技巧。

最後以Fidel所提出的操作者模式（operationalist style）和概念者模式（conceptualist style）作結論。一般而言，操作者與系統間的交互作用主要是使用系統的功能特性；而概念者則多半藉由修改檢索概念來調整他們的檢索敘述，雖說操作者有可能採取概念性的移動步驟，而概念者也有可能採取操作型的步驟。[\(註8\)](#)事實上，國人由於英語的限制，很難成為一個概念者，或許是自覺語文能力的匱乏，反倒在檢索技巧上的學習格外努力，因此複雜的顯示指令較少成為國人的困難所在，除去重複資料等進階指令的使用情況，國人的使用狀況也較歐美人士為普遍，充分透露出若依照Fidel對檢索者的分類，台大師生顯然是操作者模式，針對操作者模式所設計出線上檢索系統，自然應與概念者模式有所區隔，再度證實本土化研究的重要性。

註釋 •

- 註 1. Christine L. Borgman, "The User's Mental Model of an Information Retrieval System: Effects on Performance" (Ph. D. diss., Stanford University, 1983.)
- 註 2. C. Lewis and R. Mack, "Learning to Use a Text Processing System: Evidence from 'Thinking Aloud' Protocols," • in Proceedings of the Conference on Human • Factors in Computer Systems, (1982) : 387-92.
- 註 3. Carol Collier Kuhlthau, "A Process Approach to Library Skills Instruction," • • School Library Media Quarterly 13:1(Winter 1985) : 35-40.
- 註 4. Mu-hsuan Huang, "Pausing Behavior of End Users in Online Searching" (Ph. D.

• diss., University of Maryland, 1992). •

註 5. 此為筆者正在進行之文章，尚未發表於任何期刊或會議上。•

註 6. 同前註。•

註 7. 黃慕萱，「終端使用者之線上錯誤行為探討」，中國圖書館學會會報54期（民國84年6月），頁33-47。•

註 8. Raya Fidel, "Moves in Online Searching," • Online 9 (February 1985) : 61-74.