

終端使用者詞彙選用與概念一致性之研究—以台灣大學學生使用
PsycLIT 光碟資料庫為例

A study of Consistency in the Selection of Search Terms and Search
Concepts: A Case Study in National Taiwan University

黃慕萱

Mu-hsuan Huang

國立臺灣大學圖書資訊學系教授兼系主任

Professor and Chairman, Department of Library and Information Science,
National Taiwan University

摘要 Abstract

本文主要探討光碟資料庫終端使用者選用詞彙與詞彙概念的一致性，及其與檢索結果之關係。本研究以國立臺灣大學學生為研究對象，以 PsycLIT 光碟資料庫為檢索系統，總計有 10 個檢索問題，分別由 31 位檢索者進行 59 個題次的光碟檢索，總計產生 609 個檢索詞彙。本研究詞彙選擇一致性在第一層次上是 22.14%，第二層次上是 35%，與前人研究結果差異不大；在檢索概念上，本研究所得之檢索詞彙概念一致性偏低，且越到較後的層次差距越大。至於檢索結果的部分，本研究無論就檢出書目或判斷相關書目之重複性均較前人研究低，且重複性減少。

This article analyzes the consistency in the selection of search terms and search contents of college and graduate students in National Taiwan University when they are using PsycLIT CD-ROM database. 31 students conducted pre-assigned searches, doing 59 searches generating 609 search terms. The study finds the consistency in selection of search terms of first level is 22.14% and second level is 35%. These results are similar with others' researches. About the consistency in search concepts, no matter the overlaps of searched articles or judge relevant articles are lower than other researches.

關鍵詞 Keywords

資訊檢索；詞彙選擇；檢索詞彙；檢索概念；一致性

Information retrieval; Term selections; Search terms; Search concepts;
Consistency

壹、前言

人們為了解決所面臨的資訊爆炸(information explosion)困境,而有資訊系統的發明;資訊系統的功能不僅在於儲存處理資訊,更在於讓人們以最有效率的方式獲取所需資訊,故有資訊科技的不斷研發。然而,效率的提升僅憑系統單方面加強其儲存容量與處理速度是不足的,尚須使系統儘量配合使用者的想法與行動來提供資訊,所以今日系統設計的指導原則逐漸由系統導向轉為使用者導向,唯有從使用者角度所設計的資訊系統,才能充分發揮溝通的效益;因此由使用者的角度進行有關資訊檢索的研究,已成為不可避免的趨勢。

檢索問題是使用者面臨的難題、資訊尋求的動機和目的,以及內在知識狀態的呈現,對於其要檢索「什麼」、以及「如何」進行檢索的決定,影響相當大。(註1)檢索問題可反映出資訊需求者已發展的有關其資訊需求的心智模型,資訊需求者的認知型態則影響到檢索問題的型態,同時決定檢出書目是否符合其資訊需求。(註2)至於檢索詞彙則是使用者與資訊系統溝通的主要方式,亦可說是檢索者與系統互動過程中的主要媒介。在資訊系統中,詞彙可以指出文獻和檢索問題中具有相同概念的事物,同時透過詞彙本身的相關詞、狹義詞及廣義詞,說明這些概念間關係的親疏遠近(註3),因此檢索詞彙更是表達檢索問題的重要工具。

有關檢索詞彙一致性之研究,主要有 Saracevic & Baxter(註4)、Bellardo & Saracevic (註5)、Fidel(註6)、Saracevic & Kantor(註7)、Harter (註8)與 Iivonen(註9)等人之研究。其中 Iivonen 將「一致性」定義為「在不同時機(occasion)執行同一工作的一致程度»,在圖書資訊學裡,一致性研究其實包括索引者與檢索者內部一致性(intra-actor consistency)與外部一致性(inter-actor consistency)的探討(註10)由於本研究的關注焦點是集中在檢索者的外部一致性,所探討的文獻亦以外部一致性的研究為主。茲將上述各學者之研究內容整理列於表 1。

表 1 檢索詞彙選擇一致性研究

研究者 (發表時間)	研究重點	研究對象	樣本數/ 分析事例	檢 索 問 題 來 源
Saracevic & Baxter (1984)	以一先導研究測試不同檢索者檢索同一問題的檢索敘述一致性之三種測量值。	修習線上檢索進階課程之學生	16 人/256 題次 (每人檢索 16 題)	研究者預先設定的 16 題,工商、科技、社科人文、時事各四題
Bellardo & Saracevic (1987)	不同檢索者檢索同一問題的檢索詞彙選用重複性。	專業檢索者	36 人/200 題次 (每題 5 人次)	一般使用者 40 人提供的 40 題
Fidel (1987)	探討不同檢索者及同一檢索者不同時機檢索同一問題之檢索詞彙選擇一致性,以及檢索問題之難易對其之影響。	專業檢索者	10 人/40 題次 (每人每題各檢 索 二 次)	研究者預先設定的二題,一難一易
Saracevic & Kantor (1988)	資訊檢索行為的重覆性研究	研究計畫之工作人員	3 人/160 題次 (每題 4 人次)	一般使用者 40 人提供的 40 題
Harter (1990)	探討檢索詞彙選用在每一基礎篇數組(elementary posting set)中之重複性	研究者本身	一人/一題次 (個案研究)	研究者本身檢索之問題
Iivonen (1995)	探討不同檢索者及同一檢索者不同時機檢索同一問題之檢索詞彙選擇一致性,以及檢索問題種類與檢索環境對其之影響	來自四種檢索環境的專業檢索者	32 人/640 題次 (其中的八題檢 索 二 次)	研究者蒐自其他圖書館的 12 個真實檢索問題

在上述諸研究中,計算出不同檢索者對同一問題的檢索詞彙選用一致性約在 27-43%之間,同一檢索者於不同時間中對同一問題的檢索詞彙選用一致性則較高,約在 50-55%之間。此外,在一致性的計算上,有兩點頗值注意:其一是計算公式的問題,研究者呼籲計算一致性的公式不宜再增加,採用同一種公式所得之數據,研究結果的比較方有意義。(註11)其二是測量基準(basis)的問題,檢索詞彙非唯一的檢索一致性測量基準,Saracevic & Baxter 與 Iivonen 均另提出以檢索概念作為比對一致性的測量

基準(註12);另外, Saracevic & Baxter 的邏輯測量值(logic measure)與 Harter 的基礎篇數組(Elementary Posting Set)均納入以組合檢索詞彙的邏輯運算元作為比對一致性的測量基準。基本上, 針對同一檢索問題, 同一檢索者的詞彙一致性高於不同檢索者; 而無論同一檢索者或不同檢索者, 檢索概念選擇的一致性均高於檢索詞彙。此一現象反映出不同的檢索者之間對檢索問題的詮釋較為一致, 但在查詢過程中對檢索問題的的表達(以詞彙形式)則較不一致。(註13)

Iivonen 之所以將檢索詞彙一致性的比對分為詞彙與概念層次, 係著眼於檢索詞彙一致性測量基準的考量, Iivonen 認為, 檢索詞彙一致性的比對應逐字母(character by character)進行, 因為即使只是詞彙拼法的些許差異亦會影響到檢索結果。但如此一來似過於嚴苛, 故界定了詞彙一致性的四個層次, 也就是從詞彙到概念由嚴至寬四個層次。其中第一層次為拼法完全一致之詞彙; 第二層次納入拼法稍有異之詞彙, 包括單複數、切截變化及同義詞; 第三層次再納入具廣狹義階層關係之詞彙; 第四層次再納入有組合(combination)關係, 以及 13 種被認可的相關關係之詞彙。(註14)因此不同檢索者間檢索一致性的計算單位, 從檢索詞彙改為檢索概念後, 計算出的一致性將隨各層次的外擴而逐步提升。

其他學者在實證研究的設計中亦納入近似檢索概念的觀念, 如 Saracevic & Baxter 提出範圍測量值(scope measure), 即為計算檢索概念一致性之測量單位(註15); Harter 提出用以計算檢索詞彙選用之重複性及檢索結果精確率之基礎篇數組, 需先有查詢層面(search facet)的觀念才能執行, 也就是將同類的詞彙聚在一起, 與檢索概念的意思類似(註16); 國內黃慕萱教授的檢索問題、檢索詞彙及心智模型對資訊尋求和資訊檢索之影響研究中, 檢索者在列舉檢索詞彙時, 即被要求先將其檢索問題分解成幾個概念, 再分別列出相關的詞彙。(註17)

在檢索結果重覆性的相關研究方面, Bellardo & Saracevic 之研究結果顯示檢索結果重覆率更較檢索詞彙一致性為低。在 5 位檢索者均有檢出的少數書目裡, 100%均被判斷為相關; 有 3 位檢索者檢出的重覆書目裡, 精確率降至 73%; 而單一檢索者檢出的書目, 精確率僅 47%, 也就是說, 被越多檢索者檢出的書目, 被判斷為相關的機會越大。(註18)另一研究為 Saracevic & Kantor 等人所進行, 研究結果亦顯示檢索結果重覆率更較檢索詞彙一致性為低, 平均為 17%, 59%集中於 0%-5%的重覆率; 而判斷相關及部分相關書目的重覆率有 18%, 亦有 59%集中於 0%-5%的重覆率(註

19)，檢索結果的重覆情形中僅 2.5%可歸因於相同的檢索詞彙。然而以上二者均非以一般終端使用者作為研究對象。

終端使用者個人差異一直在檢索中扮演關鍵的角色，但個人差異的範圍有多大，其對檢索結果所造成的影響又是如何，種種問題在國內外較乏人探討。本研究即系統化地探討不同檢索者在同一檢索問題中，其詞彙選擇的變化和此變化對檢索結果的影響（可透過選用詞彙之一致性及檢索結果之一致性進行比對），對於瞭解終端使用者的資訊檢索行為，將具有相當的意義。

換言之，完全相同的檢索題目，由於檢索者個人感情認知與智慧認知上的不同，必存在相當的個別差異，不可能產生完全相同的檢索策略和檢索結果。本研究以國立台灣大學心理系所學生為研究對象，透過多人檢索同一問題，探討檢索者選用詞彙概念的一致性與檢索結果的重複性，以及二者間的關係及與其他檢索特性的關係，並進一步探討選用詞彙的一致性及各種檢索結果之關係。以下將分別從選用詞彙與檢索概念之一致性兩方面加以探討。

貳、研究對象和研究設計

為控制研究對象學科背景的差異性對研究結果的影響，本研究以研究工具 PsycLIT 的最大使用群--台大心理學系所學生為研究對象。為控制研究對象的教育程度對研究結果的影響，根據選定問題需求者之教育程度來決定計畫檢索者應具備的教育程度。也就是說，由於選出之 10 個檢索問題中，有 8 個係由研究所碩士班學生提出，其餘 2 個由大學部學生提出，因此每一題均需招募 5 位以上教育程度與之相配合的檢索者。由於目標對象人數並不多，故未公開徵求受試者，而是兼採電話逐一徵詢意願、拜訪研究室與受試者互薦等方式進行，研究期間共招募到 33 位願意參與的計畫檢索者(連無效樣本併計)。

在 33 位計畫檢索者當中，有 2 位檢索者的 3 個題次因故未完成檢索，或未完整抓取檢索過程，故對「檢索詞彙」與「檢索結果」而言，均為無效樣本，因此，本研究所分析的「檢索詞彙」有效樣本共有 31 位計畫檢索者的 59 個題次。此外，尚有 6 位檢索者的 9 個題次在選擇檢出書目的過程有瑕疵，其中 3 位兩題均有瑕疵；另 3 位僅一題有瑕疵，故僅對「檢索結果」而言為無效樣本，因此，本研究所分析的「檢索結果」有效樣本共有 28 位計畫檢索者(包括 22 位研究所碩士班學生與 6 位大學部學生)的

50 個題次，每題均有 5 次檢索。以下有關研究對象的各種敘述統計，均是以「檢索詞彙」有效樣本的 31 位計畫檢索者來進行。

就性別比例而言，31 位有效的受試者之中，男性有 10 人(占 32.3%)，女性有 21 人(占 67.7%)，若以碩士班一年級研究生總數男女各半的比例而言，女性的參與意願為男性的 2 倍。就教育程度而言，研究所碩士班學生有 25 人(占 80.6%)；大學部學生有 6 人(占 19.4%)次之。這是為控制研究對象的教育程度對研究結果的影響，根據所選定問題提出者之教育程度(碩士班：大學部=1:4)來決定計畫檢索者之教育程度。

此外，本研究亦透過文書處理軟體、程式設計、試算表軟體三方面瞭解受試者對電腦使用的熟悉程度。統計結果以文書處理軟體的熟悉程度最高，以五等第量表計算，熟悉度平均達 3.87；試算表軟體的熟悉程度次之，熟悉度平均也有 2.97；受試者對程式設計最不熟悉，平均僅有 1.84 的熟悉度。而在對於研究工具--PsycLIT 資料庫的使用經驗上，以前一學期使用總次數計算，平均為 7.55 次，但次數分布則相當懸殊，最少為零次，有 3 人(占 9.7%)，其實這四位受試者不僅前一學期未曾使用，更是從未使用過 PsycLIT 資料庫者；使用次數最多者為 30 次，標準差高達 7.28。對於圖書館線上公用目錄(OPAC)的使用經驗，亦以前一學期使用總次數計算，平均為 15.29 次，高於研究工具 PsycLIT 的使用經驗；次數分布亦非常懸殊，從最少的 1 次到最多的 100 次，標準差高達 18.64。可見本研究中受試者對 OPAC 的使用經驗較 PsycLIT 豐富，但使用頻率略為兩極化。在對其他書目資料庫的檢索經驗方面，以書目資料庫種類而言，比例最高的前三種分別是中文博碩士論文索引(14 人，占 45.2%)、中華民國期刊論文索引(13 人，占 41.9%)與 Medline(4 人，占 12.9%)，其餘的書目資料庫使用過的比例均未超過 10%；分析其原因就語言別與學科性質而言，居首二位的「中文博碩士論文索引」及「中華民國期刊論文索引」均屬中文、綜合類的書目資料庫，主要係因中文書目資料庫因較無精細的分科，尤其在期刊論文資料庫方面更僅有中華民國期刊論文索引資料庫一個選擇，較之分科較細的西文書目資料庫，使用上自然較為集中；此外，心理學研究所研究生檢索西文文獻時，一般均以收錄最完整心理學書目資料的 PsycLIT 為主，而本項所問的是 PsycLIT 以外的其他書目資料庫，因此西文、專科書目資料庫的使用自然較為分散，故比例低於中文、綜合書目資料庫。

在研究設計上，本研究採用實證研究法，過程中使用了文獻分析法、問卷法、觀察法、查詢紀錄分析法(transaction logs)、訪談法等，其中問卷

法係利用「受試者基本資料問卷」以瞭解受試者的個人背景資料，包括性別、系級組別、電腦熟悉程度(包括文書處理、程式設計與試算表軟體三類)以及資訊檢索系統的檢索經驗(包括圖書館線上公用目錄與其他書目性資料庫)等；此外研究對象還需填寫「查詢詞彙問卷」，目的是引導受試者根據問題資料上所呈現之檢索問題的口頭陳述之謄錄，以及問題提出者學科系級背景、查詢目的等資料，將檢索問題劃分為檢索概念與每一概念所屬之檢索詞彙(包括同義詞、相關詞與欲排除之詞彙)，並指出概念間的邏輯關係及對檢索主題的熟悉程度，除協助受試者利用分區組合檢索建立其檢索策略外，亦可藉此瞭解受試者對檢索主題的熟悉程度與主題間關係的想法。觀察法則記錄檢索過程中無法顯示在查詢記錄中之檢索者輸入的詞彙和詞彙選用的過程，以及相關的問題。

在研究對象檢索完成之後，利用訪談法確認觀察中發現的問題，如新增詞彙的概念歸屬；並配合查詢記錄分析法，透過 PsycLIT 光碟資料庫本身的功能抓取檢索結果，以掌握檢索者所採用之檢索詞彙、詞彙選用的過程和檢索結果。由於查詢過程記錄中並未完整記載所有詞彙選擇的互動過程，故另設計「互動記錄單」，觀察記錄受試者鍵入詞彙以對應至主題詞以及瀏覽索引典的過程。

在研究實施地點方面，由研究者主導，盡量安排安靜不受打擾的空間。由研究者提供具中英對譯功能的線上雙語翻譯軟體之個人電腦，以排除因語言問題造成的檢索互動障礙。此外，受試者於檢索進行中可參考其事前準備好的備忘紀錄，但研究環境不提供除線上雙語翻譯之外的任何參考工具，受試者亦不被允許於檢索進行中利用電話詢問他人任何關於該次檢索之問題，以避免因此對檢索行為造成的影響。

開始進行檢索前，受試者均須接受檢索系統使用訓練，以控制因對系統熟悉程度不同產生的差異。檢索開始後，受試者可自由翻閱書面使用說明，亦可隨時詢問研究者系統操作方面的問題。研究者採用此介入行為的理由在協助受試者排除因不諳操作造成的障礙，以順利貫徹其檢索策略，希望能獲得 Spink & Saracevic 所謂「促成使用者與系統間表層互動(註20)」的效果。

參、選用詞彙一致性之分析

在探討計畫檢索者選用詞彙之一致性前，必須先界定選用詞彙、一致性、一致性之層次、一致性之計算公式與各層次一致性之計算方法。

在檢索詞彙方面，本研究採用 Iivonen 對檢索詞彙的界定--「以布林運算元作為分隔的字串，或是開始/結束某一檢索敘述的字串」(註21)，前者係指檢索敘述中以布林運算元將檢索詞彙加以組合的情形；後者則指檢索敘述中未曾運用布林運算元組合檢索詞彙的情形。就 PsycLIT 系統而言，在本研究所蒐集到之檢索過程中曾出現過的「使用」詞彙，包括下列七種使用方式：

- 1.KW:採用關鍵詞模式進行檢索之詞彙。關鍵詞模式的作用為在 Title (題名,簡稱 ti), Abstract(摘要,簡稱 ab), heading word(主題索引用語,簡稱 hw), ty(圖書章節名稱中的用語),Key Phrase Identifiers(關鍵片語,簡稱 id)五欄位中蒐尋,系統將回應以其中任一欄位符合輸入條件的書目紀錄。
- 2.ti:指定於單一的 Title 欄位檢索。
- 3.jn:指定於單一的 Journal Name(期刊全名)欄位檢索。
- 4.SH:採用主題詞模式進行檢索之詞彙。主題詞模式的作用為將輸入條件對應至內建的主題索引典,由系統提供相近的主題詞供使用者挑選,並回應以主題詞欄位符合的書目紀錄。
- 5.sh:指定於單一的 Subject Heading(主題詞)欄位檢索。
- 6.MAP:採用主題詞模式後改用系統所提供主題詞進行檢索前,原輸入作為對應(map)主題索引典的詞彙。該詞彙在當次並未實際用以檢索。
7. THE:僅用以瀏覽其索引典但在當次並未實際用以檢索之主題詞。

而為使情形單純化,將本研究中比對詞彙的範圍限定在前五種,也就是只比對實際用以檢索,留存於查詢過程記錄的部分,排除未實際用以檢索的 MAP 與 THE 二種。此外,同一次檢索中若有拼字完全相同但使用方式不同之詞彙,若其可能導致不同的檢索結果時,均以不同檢索詞彙計之,如 KW 與 SH;即使同為 SH,若以延展方式(exp.)使用時,由於會導致相當不同的檢索結果,故亦視為不同的檢索詞彙。但檢索詞彙不計重複的使用次數,也就是說,在一次檢索過程中,同一詞彙以同樣的使用方式出現多次者,均只計為一次。

至於一致性的層次,本文沿用 Iivonen 的說法,將「一致性」界定為「在不同時機執行同一工作的一致程度」(註22),可依此將計畫檢索者詞彙選用的一致性界定為「不同的計畫檢索者執行同一檢索問題時選用詞彙的一致程度」。一般均認為檢索詞彙一致性的比對應逐字母(character by character)進行,因為即使只是詞彙拼法與用法的些許差異亦會影響到檢索結果。但如此一來似過於嚴苛,故 Iivonen 提出由詞彙逐漸擴充到概念的

四個自嚴格至寬鬆的一致性比對層次標準，其中第一層次為拼法完全一致之詞彙；第二層次納入拼法稍有異之詞彙，包括單複數、切截變化及同義詞；第三層次再納入具廣狹義階層關係之詞彙；第四層次再納入有組合(combination)關係，以及 13 種被認可的相關關係之詞彙。(註23)本研究參考 Iivonen 對從檢索詞彙到檢索概念一致性層次的分法，根據研究中所蒐集到的情況以及執行比對的可行性，界定了詞彙與概念一致性比對的層次，其中屬於詞彙部分的有兩個層次，比對標準分別為：

第一層次：逐字母比對，拼字及檢索欄位必須完全一致，若可能造成檢索結果不同，即視為不一致。

第二層次：第一層次外再加上

1. 拼字稍異者：包括單複數、切截變化及拼錯字者。
2. 二詞彙間有受試者認定的同義關係，或系統的索引典設定為有同義關係(USE, UF)。
3. 延展敘述語與非延展敘述語之間。
4. 有縮寫與全稱關係者。
5. 符合以上情況但檢索欄位有別者。

一致性的計算公式上，則採用近年來在檢索者一致性的研究中常用的非對稱公式，以便於研究結果的比較，令研究成果更具意義。非對稱公式如下：

$$\text{A 的一致性} = \frac{\text{A,B 共用之詞彙數}}{\text{A 使用之詞彙數}}$$

$$\text{B 的一致性} = \frac{\text{A,B 共用之詞彙數}}{\text{B 使用之詞彙數}}$$

其中 X、Y 為兩個不同的檢索者。非對稱公式的最大特點在於二檢索者之間的一致性有兩個數據，因此可個別顯示每一個檢索者的一致性程度。X、Y 二者之一致性即為該二數據的平均值，其公式如下：

$$tc = \left(\frac{100 \times A}{A + M} + \frac{100 \times A}{A + N} \right) / 2$$

tc：詞彙使用的一致性

A：X、Y 一致使用的詞彙個數

M：被 X 使用，不被 Y 使用的詞彙總數

N：被 Y 使用，不被 X 使用的詞彙總數

在每一檢索問題的檢索詞彙一致性平均數方面，由於每一題有 5 到 6 位檢索者，因此採取兩兩成對計算後再計算平均數。其中有 5 位檢索者的檢索問題由於檢索者兩兩之間有 10 種可能的不同組合，故該檢索問題的檢索詞彙一致性平均數算式為：

$$(S_{1,2} + S_{1,3} + S_{1,4} + S_{1,5} + S_{2,3} + S_{2,4} + S_{2,5} + S_{3,4} + S_{3,5} + S_{4,5}) / 10$$

(其中 $S_{x,y}$ 意謂第 x,y 位檢索者兩兩之間的詞彙使用一致性)

有 6 位檢索者的檢索問題由於檢索者兩兩之間有 15 種可能的不同組合，故該檢索問題的檢索詞彙一致性平均數算式為：

$$(S_{1,2} + S_{1,3} + S_{1,4} + S_{1,5} + S_{1,6} + S_{2,3} + S_{2,4} + S_{2,5} + S_{2,6} + S_{3,4} + S_{3,5} + S_{3,6} + S_{4,5} + S_{4,6} + S_{5,6}) / 15$$

(其中 $S_{x,y}$ 意謂第 x,y 位檢索者兩兩之間詞彙使用的一致性)

將每一檢索問題的檢索詞彙一致性平均數相加後，再除以檢索問題總數，即為本研究所得的檢索詞彙一致性總平均。

至於各層次一致性的計算方法，則當計算共同使用的詞彙數時，除第一層次必然為單純的一對一情形外，第二層次可能面臨到跨層次的一對多情況，使計算變得較為複雜，茲以檢索問題 A 的部分兩兩比對組合為例，說明如表 2 所示：

表 2 詞彙比對第一層次與第二層次的計算列示
(以檢索問題 A 的 S1,2、S1,3、S1,4 為例)

檢索者兩兩比對	層次	A	M	N
S1,2	L 1	0	1-0=11	8-0=8
	L 2	0+1=1	1-0-1=10	8-0-1=7
S1,3	L 1	1	11-1=10	9-1=8
	L 2	1+0=1	11-1-1=9	9-1-0=8
S1,4	L 1	1	11-1=10	12-1=11
	L 2	1+2=3	11-1-3=7	12-1-2=9

如表 2 所示，其中第二層次的共同使用詞彙個數 A 的算法是第一層次加上第二層次新增，且完全未與第一層次重覆的部分；而第二層次的個別使用詞彙個數 M、N 算法則是第一層次減掉所有歸屬於共同詞彙，且未與第一層次重覆的部分。

此外，本研究除了前述兩兩比對的情況外，另針對詞彙一致性的第一層次計算了二人以上使用同一詞彙的詞彙數，檢視其遞檢情形。其中有 5 位檢索者的共同檢索詞彙平均數算法為：

$$3 \text{ 人的一致性平均} = (S_{1,2,3} + S_{1,2,4} + S_{1,2,5} + S_{1,3,4} + S_{1,3,5} + S_{1,4,5} + S_{2,3,4} + S_{2,3,5} + S_{2,4,5} + S_{3,4,5})/10$$

$$4 \text{ 人的一致性平均} = (S_{1,2,3,4} + S_{2,3,4,5} + S_{1,3,4,5} + S_{1,2,3,5} + S_{1,2,4,5})/5$$

$$5 \text{ 人的一致性平均} = S_{1,2,3,4,5}$$

(其中 $S_{x,y,\dots}$ 意謂第 x,y,\dots 位檢索者兩兩之間詞彙使用的一致性)

有 6 位檢索者的共同檢索詞彙平均數算法為：

$$3 \text{ 人的一致性平均} = (S_{1,2,3} + S_{1,2,4} + S_{1,2,5} + S_{1,2,6} + S_{1,3,4} + S_{1,3,5} + S_{1,3,6} + S_{1,4,5} + S_{1,4,6} + S_{2,3,4} + S_{2,3,5} + S_{2,3,6} + S_{2,4,5} + S_{2,4,6} + S_{3,4,5} + S_{3,4,6} + S_{1,5,6} + S_{2,5,6} + S_{3,5,6} + S_{4,5,6})/20$$

$$4 \text{ 人的一致性平均} = (S_{1,2,3,4} + S_{1,2,3,5} + S_{1,2,3,6} + S_{2,3,4,5} + S_{2,3,4,6} + S_{2,3,5,6} + S_{1,3,4,5})$$

$$+ S_{1,3,4,6} + S_{1,3,5,6} + S_{1,2,4,5} + S_{1,2,4,6} + S_{1,2,5,6} + S_{1,4,5,6} + S_{2,4,5,6} + S_{3,4,5,6} / 15$$

$$5 \text{ 人的一致性平均} = (S_{1,2,3,4,5} + S_{2,3,4,5,6} + S_{1,3,4,5,6} + S_{1,2,4,5,6} + S_{1,2,3,5,6} + S_{1,2,3,4,6}) / 6$$

$$6 \text{ 人的一致性平均} = S_{1,2,3,4,5,6}$$

(其中 $S_{x,y}$... 意謂第 x,y ... 位檢索者兩兩之間詞彙使用的一致性)

本研究中檢索者在 59 次檢索裡，共使用了 609 個檢索詞彙，每次檢索所使用的詞彙數，最多者為 54 個，最少者為 2 個，從分組分配的角度觀之，如表 3 所示，可知其中以使用 6-10 個檢索詞彙者最多(25 次，占 42.4%)，其次為使用 1-5 個者(15 次，占 25.4%)；使用 16 個詞彙以上者僅有 7 次(占 11.9%)。平均每次檢索使用了 10.32 個檢索詞彙，標準差為 8.02。這一數據與 Saracevic 及 Kantor 等人所提出，且經國內研究者如黃慕萱教授、陳佳君等驗證過的「典型檢索過程」(typical search)中，應包含 10 個檢索詞彙的數據(註24)極為相近。

表 3 計畫檢索者使用檢索詞彙個數次數分配表

檢索詞彙數	次數	百分比(%)	累積百分比(%)
1-5	15	25.4	25.4
6-10	25	42.4	67.8
11-15	12	20.3	88.1
16-20	3	5.1	93.2
21-25	2	3.4	96.6
26 以上	2	3.4	100.0
總計	59	100.0	100.0

在詞彙使用的一致性上，在第一層次方面，根據上述公式的計算結果，共有 145 組數據，平均值為 22.14，標準差為 15.49。最大值為 70，僅有一組；最小值為完全未重複(0)，有 20 組(13.79%)，其中並無任一數據達 100% 者，顯示就檢索詞彙逐字比對的情形而言，兩次檢索之間並無比對完全相同的狀況出現。在次數分布方面如表 4 所示，其中以 20 以下最多，有 75 組，已過半數(51.72%)，20-39(52 組，35.86%) 者居次；40-59(14 組，

9.66%) 再次之：一致性越高，組數越少，一致性在 60% 以上者僅有 4 組 (2.76%)。

表 4 計畫檢索者檢索詞彙選用一致性第一層次次數分配表

一致性(%)	組數	百分比(%)	累積百分比(%)
0-19	75	51.72	51.72
20-39	52	35.86	87.58
40-59	14	9.66	97.24
60-79	4	2.76	100.00
80 以上	0	0	100.00
總計	145	100.00	100.00

檢索詞彙選用的第二層次方面，根據上述公式的計算結果，共有 145 組數據，平均值為 35，標準差為 18.02，其一致性最大值為 80，僅有一組；最小值仍為完全未重複者(0)，有 7 組(4.83%)，亦無任一數據達 100% 者；在次數分布方面如表 5 所示，其中以 20-39 最多，有 55 組(37.93%)；40-59 次之，有 42 組(28.97%)；0-19 再次之，有 34 組(23.45%)。40% 以下之一致性仍佔六成以上(61.38%)。

表 5 計畫檢索者檢索詞彙選用一致性第二層次次數分配表

一致性(%)	組數	百分比(%)	累積百分比(%)
0-19	34	23.45	23.45
20-39	55	37.93	61.38
40-59	42	28.97	90.35
60-79	13	8.97	99.32
80 以上	1	0.69	100.00
總計	145	100.00	100.00

換言之，本研究詞彙選擇一致性在較嚴謹的第一層次中為 22.14%，在較為鬆散的第二層次的一致性則為 35%，和前人的研究一致性介於 27% 到 43% 之間的結果不謀而合。

除了前述兩兩比對的情況外，本研究另針對第一層次計算了二人以上

使用同一詞彙的詞彙數，檢視其遞檢情形。依前述計算方法將各檢索問題二人以上共同檢索詞彙數的平均值列於表 6：

表 6 各檢索問題共同檢索詞彙平均數
隨檢索者增加的遞減情形

檢索問題 編號	檢索者人數 (數據組數)	共同檢索詞彙個數	
		總數	平均數
A	2 人共同(15)	27	1.80
	3 人共同(20)	12	0.60
	4 人共同(15)	5	0.33
	5 人共同(6)	1	0.17
	6 人共同(1)	0	0
B	2 人共同(15)	23	1.53
	3 人共同(20)	15	0.75
	4 人共同(15)	6	0.40
	5 人共同(6)	1	0.17
	6 人共同(1)	0	0
C	2 人共同(15)	35	2.33
	3 人共同(20)	12	0.60
	4 人共同(15)	2	0.13
	5 人共同(6)	0	0
	6 人共同(1)	0	0
D	2 人共同(15)	25	1.67
	3 人共同(20)	10	0.50
	4 人共同(15)	2	0.13
	5 人共同(6)	0	0
	6 人共同(1)	0	0
E	2 人共同(15)	28	1.87
	3 人共同(20)	14	0.70
	4 人共同(15)	3	0.20
	5 人共同(6)	0	0
	6 人共同(1)	0	0
F	2 人共同(10)	32	3.20
	3 人共同(10)	15	1.5
	4 人共同(5)	5	1.00
	5 人共同(1)	1	1.00
G	2 人共同(15)	21	1.40
	3 人共同(20)	15	0.75
	4 人共同(15)	6	0.40
	5 人共同(6)	1	0.17

	6 人共同(1)	0	0
H	2 人共同(15)	35	2.33
	3 人共同(20)	9	0.45
	4 人共同(15)	0	0
	5 人共同(6)	0	0
	6 人共同(1)	0	0
I	2 人共同(15)	22	1.47
	3 人共同(20)	15	0.75
	4 人共同(15)	6	0.40
	5 人共同(6)	1	0.17
	6 人共同(1)	0	0
J	2 人共同(15)	31	2.07
	3 人共同(20)	24	1.20
	4 人共同(15)	15	1.00
	5 人共同(6)	6	1.00
	6 人共同(1)	1	1.00
總平均	2 人共同	1.97	
	3 人共同	0.73	
	4 人共同	0.40	
	5 人共同	0.27	
	6 人共同	0.11	

表 6 顯示從 2 人到 6 人共同檢索詞彙數的遞檢情形，可知除兩兩比對的平均共同詞彙個數 1.97 大於 1 之外，其餘均不到 1 個，然而亦存在 6 人均使用同一檢索詞彙的情形，因此所有的總平均數均大於零。而遞檢的程度以 2 人到 3 人之間最大，其次為 5 人到 6 人之間。若檢視個別平均數為零，也就是沒有共同檢索詞彙的比率來看，從 4 人共同檢索詞彙數才開始出現平均數為零的情形，僅占 10% 的情形，到了 5 人時，共同檢索詞彙數平均為零的情形已大幅提昇至 40%，到了 6 人時，更高達 77.8%，在九個數據中僅有兩個不為零。

進一步檢視從 2 人到 6 人共同檢索詞彙數的分布情形，如表 7 所示，表的左欄為本研究各次檢索之間曾出現過的 2 人至 6 人共同使用詞彙數，其右則依序列出 2 人至 6 人共同使用特定詞彙個數的比對組數及占有組數之百分比。由該表整體觀之，共同使用檢索詞彙的分布基本上是集中在詞彙個數較少的部份，其中個數最少的為零，也就是在比對的數個檢索之間並沒有任何一個完全相同之詞彙；從 2 人到 6 人都有此種情況，且除了 6 人（88.89%）略低於 5 人（90.91%）之外，大致上比對的人數越多，共同詞彙個數為零的比例也越高。而共同使用檢索詞彙個數最多的則是 10

個，僅在 2 人比對的情況中出現過一次，事實上比對的人數越多，共同使用檢索詞彙的個數則越少，且由 2 人到 3 人之間的變化最大，因為僅有 2 人比對時才有 4 個到 10 個共同詞彙的狀況；3 人比對時最高的共同詞彙數只有 3 個；4 人比對時的最高共同詞彙數只有 2 個；5 及 6 人比對時的最高共同詞彙數均只有 1 個。

表 7 2 至 6 人共同檢索詞彙個數之次數分配表

共同詞彙個數	2 人共同 (共 145 組)	3 人共同 (共 190 組)	4 人共同 (共 140 組)	5 人共同 (共 55 組)	6 人共同 (共 9 組)
0 個詞彙	20(13.79%)	84(44.21%)	93(66.43%)	50(90.91%)	8(88.89%)
1 個詞彙	47(32.41%)	78(41.05%)	44(31.43%)	5(9.09%)	1(11.11%)
2 個詞彙	40(27.59%)	21(11.05%)	3(2.14%)	0	0
3 個詞彙	22(15.17%)	7(3.68%)	0	0	0
4 個詞彙	7(4.83%)	0	0	0	0
5 個詞彙	2(1.38%)	0	0	0	0
6 個詞彙	5(3.45%)	0	0	0	0
8 個詞彙	1(0.69%)	0	0	0	0
10 個詞彙	1(0.69%)	0	0	0	0

以每一欄觀之，2 人比對的結果是以 1 個共同詞彙的情況最多(32.41%)；其次為 2 個與 3 個共同詞彙(各占 27.59%與 15.17%)。4 個(含)以上共同詞彙的比例僅約 1 成(11.03%)；事實上除了 2 人比對的情形之外，其餘 3 到 6 人比對的共同詞彙數均是以零個的情況最多，其後隨著共同詞彙數的增加而遞減。

肆、檢索概念一致性之分析

在報告計畫檢索者選用檢索詞彙一致性的結果之前，必須先就檢索概念、一致性的層次、一致性的計算公式，以及各層次一致性的計算方法等基本概念分別加以界定說明。

首先說明檢索概念的意義，檢索概念係指將檢索問題依主題所分析而成的要素。也就是分區組合檢索法中所指的主題層面。檢索者所使用的不同詞彙之間若有同義、近義或階層關係，則可視為同一檢索概念。一致性的層次則根據 Iivonen 從檢索詞彙到檢索概念一致性層次的分法。由於其

第三、四層次的比對標準各為納入具廣狹義階層關係之詞彙；納入有組合 (combination) 關係，以及 13 種被認可的相關關係之詞彙(註25)，實已擴充至檢索概念的範圍，因此本研究將之劃歸檢索概念比對的部分，再根據研究中所蒐集到的情況以及執行比對的可行性，界定概念一致性比對的兩個層次，比對標準如下：

第一層次：第一、二層次外再加上二詞彙在系統的索引典中設定為有廣狹義的直接關係(BT, NT)。

第二層次：第一、二、三層次外再加上二詞彙在系統的索引典中設定為有相關的直接關係(RT)。

檢索概念一致性的計算公式，同樣依據前述的非對稱公式原理，X、Y 二檢索者檢索概念一致性的公式如下：

$$cc = \left(\frac{100 \times A}{A + M} + \frac{100 \times A}{A + N} \right) / 2$$

cc：概念使用的一致性

A：X、Y 一致使用的概念個數

M：被 X 使用，不被 Y 使用的概念總數

N：被 Y 使用，不被 X 使用的概念總數

每一檢索問題的檢索概念一致性平均數算法，其算式與檢索詞彙的部分完全相同，此處不再贅述。

至於各層次一致性的計算方法，則當計算共同使用的概念數時，兩個層次均可能面臨到跨層次的一對多情況，使計算變得較為複雜，茲以檢索問題 A 的部分兩兩比對組合為例，說明如表 8 所示：

如表 8 所示，須將詞彙一致性第二層次視為概念一致性第一層次的前一層次，概念一致性各層次的共同使用概念個數 A 的算法即為前一層次加上該層次新增，且完全未與前一層次重覆的部分；而各層次的個別使用概

念個數 M、N 算法則是前一層次減掉所有歸屬於共同詞彙，且未與前一層次重覆的部分。

表 8 概念比對第一層次與第二層次的計算列示
(以檢索問題 A 的 S1,2、S1,3、S1,4 為例)

檢索者兩兩比對	層次	A	M	N
S1,2	L 1	0+1+1=2	11-0-1-1=9	8-0-1-2=5
	L 2	0+1+1+1=3	11-0-1-1-1=8	8-0-1-2-1=4
S1,3	L 1	1+0+2=3	11-1-1-2=7	9-1-0-1=7
	L 2	1+0+2+1=4	11-1-1-2-1=6	9-1-0-1-1=6
S1,4	L 1	1+2+3=6	11-1-3-1=6	12-1-2-3=6
	L 2	1+2+3+0=6	11-1-3-1-0=6	12-1-2-3-1=5

依據前述的計數標準，在檢索概念的第一層次方面，根據上述公式的計算結果，共有 145 組數據，其一致性平均值為 39，標準差為 16.43，其中最大值為 81.25；最小值為完全未重複(0)，各均僅有 1 次，亦無任一數據達 100 者；在次數分布方面如表 9 所示，其中以 20-39 最多，有 61 組 (42.07%)；40-59 次之，有 54 組 (37.24%)；0-19 與 60-79 再次之，各有 14 組 (9.66%)。一致性 40% 以下者仍居一半以上 (51.73%)。

表 9 計畫檢索者檢索概念一致性第一層次次數分配表

一致性(%)	組數	百分比(%)	累積百分比(%)
0-19	14	9.66	9.66
20-39	61	42.07	51.73
40-59	54	37.24	88.97
60-79	14	9.66	98.63
80 以上	2	1.38	100.00
總計	145	100.00	100.00

檢索概念選用的第二層次方面，根據上述公式的計算結果，共有 145 組數據，亦無任一數據達 100 者，其一致性平均值為 46.5，最大值為 91.67；最小值為 13.33，各均僅有 1 次，標準差為 17.55；在次數分布方面如表 10

所示,其中以 40-59 最多,有 66 組(45.52%);20-39 次之,有 40 組(27.59%);60-79 再次之,有 27 組(18.62%), 最少的是一致性 20% 以下(8 組, 5.52) 與 80% 以上者(4 組, 2.76) 最少, 合計僅占 8.28%。

表 10 計畫檢索者檢索概念一致性第二層次次數分配表

一致性(%)	組數	百分比(%)	累積百分比(%)
0-19	8	5.52	5.52
20-39	40	27.59	33.11
40-59	66	45.52	78.63
60-79	27	18.62	97.25
80 以上	4	2.76	100.00
總計	145	100.00	100.00

從以上平均值與次數分配數據可大致看出,無論是檢索詞彙或檢索概念各由第一到第二層次,或是由檢索詞彙到檢索概念,由於比對標準的漸次放寬,計出的一致性均是呈現遞增的現象。

伍、檢索結果分析

本研究分析的檢索結果,包括檢索效益與重覆性兩個部分。其中檢索效益的部分由於無法得知資料庫內所有相關的書目筆數,因此本研究中的檢索效益均以精確率表達之;重覆性則指對同一檢索問題,不同的檢索者檢出書目及判斷相關書目的相同程度。以下先介紹計算基準,再陳述計算的結果:

一、檢索效益的計算基準

檢索效益的計算方式為將經由相關判斷後選出的書目筆數(R)除以檢出書目筆數(B)。其中檢出書目(B)的認定標準為在進行相關判斷之前有確實瀏覽過的書目,因此包括相關判斷筆數大於零的檢索組的瀏覽筆數,以及相關判斷筆數雖為零,但曾瀏覽完畢整個檢索組的筆數,如此一來可將沒有瀏覽動作的檢索組與未瀏覽完即判斷相關筆數為零的檢索組排除,亦可將只瀏覽部分即判斷為相關的檢索組,其尚未瀏覽的部分排除。以此標準計算,應當較直接採用相關判斷筆數大於零的檢索組筆數,更能精確表達進行相關判斷前的檢索結果;而相關判斷書目(R)的認定標準則為在逐筆

瀏覽後判斷為相關而選出的書目。

由於是將每一題次的整個檢索結果視為一體，而通常每一題次的整個檢索結果可能由好幾個檢索循環(search cycle)的檢索結果所組成，這些檢索結果彼此之間可能會有重覆的情形，因此在開始分析前，均將每題次的檢索結果去除互相重覆的部分，以獲得更為正確的數據。

二、檢索效益的分析

依上述標準的計算結果，本研究中檢索者在 10 個檢索問題共 50 次的檢索中，檢出書目筆數 B 最多為 188 筆，最少為 8 筆，平均每一次檢索的檢出書目筆數為 55.36 筆，標準差為 45；判斷相關的書目筆數 R 最多為 48 筆，最少為 1 筆，平均每一次檢索的判斷相關書目筆數為 15.24 筆，標準差為 11.17。在次數分布方面如表 11 所示，檢出書目筆數以 11-20 筆與 51-60 筆最多，但比例均不甚高(均有 8 次，占 16%)；有超過半數(56%)是在 40 筆以上，在 70 筆以上者亦有三成之多；而判斷相關書目筆數則以 1-10 筆與 11-20 筆最多，各有 19 次(38%)與 18 次(36%)，40 筆以上者僅有 3 人(6%)。可見本研究的檢索者由檢出到判斷相關，書目筆數的減少相當可觀。

表 11 計畫檢索者檢出書目筆數與判斷相關書目筆數
次數分配表

筆數分組	檢出書目筆數(B)			判斷相關書目筆數(R)		
	次數	百分比(%)	累積百分比(%)	次數	百分比(%)	累積百分比(%)
1-10	4	8%	8%	19	38%	38%
11-20	8	16%	24%	18	36%	74%
21-30	5	10%	34%	8	16%	90%
31-40	5	10%	44%	2	4%	94%
41-50	5	10%	54%	3	6%	100%
51-60	8	16%	70%	-	-	-
61-70	3	6%	76%	-	-	-
71-80	2	4%	80%	-	-	-
81-90	1	2%	82%	-	-	-
90-100	2	4%	86%	-	-	-
100 以上	7	14%	100%	-	-	-
總計	50	100%	100%	50	100%	100%

將本研究 50 次檢索的判斷相關書目筆數 R 除以檢出書目筆數 B 之後，可得出精確率的平均值僅有 34.36，還未到 50，標準差為 19.8，最大值為 80，最小值為 5.26，其中並沒有 100% 的完全精確率，以分組次數分配表觀之，如表 12 所示，以精確率在 20-29 者最多(占 28%)，有一半以上(54%)的精確率在 30 以下，有 90% 的精確率都在 60 以下，可見大部分的精確率集中在較低的部分。

表 12 計畫檢索者檢索精確率次數分配表

精確率(%)	次數	百分比(%)	累積百分比(%)
0-9	2	4.0	4.0
10-19	11	22.0	26.0
20-29	14	28.0	54.0
30-39	6	12.0	66.0
40-49	5	10.0	76.0
50-59	7	14.0	90.0
70-79	3	6.0	96.0
80-89	2	4.0	100.0
總計	31	100.0	100.0

三、檢索結果重覆性的比對基準說明

本研究在檢索結果重覆性的計算公式係沿用前述詞彙一致性的計算公式，採取非對稱性的原則，據前述的非對稱公式原理，X、Y 二檢索者檢索結果重覆性的公式如下：

$$rd = \left(\frac{100 \times B}{B + I} + \frac{100 \times B}{B + J} \right) / 2$$

rd：檢索結果的非對稱重複性

B：X、Y 都檢索到的相同文章數目

I：被 X 檢索到，但 Y 並未檢索到的文章數目

J：被 Y 檢索到，但 X 並未檢索到的文章數目

在每一檢索問題的檢索結果重覆性平均數方面，亦是採取兩兩成對計算後再計算平均數。由於檢索結果的情況較為單純，每一題的有效檢索結果均為 5 組，因此檢索者兩兩之間有 10 種可能的不同組合，故該檢索問題的檢索結果重覆性平均數算法為：

檢索結果重覆性平均數算法為：

$$(S_{1,2} + S_{1,3} + S_{1,4} + S_{1,5} + S_{2,3} + S_{2,4} + S_{2,5} + S_{3,4} + S_{3,5} + S_{4,5})/10$$

(其中 $S_{x,y}$ 意謂第 x,y 位檢索者兩兩之間的檢索結果重覆性)

此外，本研究除了前述兩兩比對的情況外，另各就檢出書目與相關判斷書目計算了二人以上的檢索結果重覆性，檢視其遞檢情形。作法為將每次檢索儲存之整體檢索結果依所求之共同檢索結果人數加以交集，即可比對出重覆之書目。其重覆檢索結果平均數算法為：

$$3 \text{ 人的重覆性平均} = (S_{1,2,3} + S_{1,2,4} + S_{1,2,5} + S_{1,3,4} + S_{1,3,5} + S_{1,4,5} + S_{2,3,4} + S_{2,3,5} + S_{2,4,5} + S_{3,4,5})/10$$

$$4 \text{ 人的重覆性平均} = (S_{1,2,3,4} + S_{2,3,4,5} + S_{1,3,4,5} + S_{1,2,3,5} + S_{1,2,4,5})/5$$

$$5 \text{ 人的重覆性平均} = S_{1,2,3,4,5}$$

(其中 $S_{x,y,\dots}$ 意謂第 x,y,\dots 位檢索者兩兩之間的檢索結果重覆性)

四、檢索結果重覆性分析

在檢索結果的重覆性上，根據上述公式的計算結果，檢出書目與判斷相關書目的重覆性各有 100 個數據，其中並無任一數據達 100% 者，顯示無論就判斷相關書目或檢出書目而言，均無兩兩比對完全相同的狀況出現。檢出書目方面，重覆性平均值為 10.48，標準差為 15.79，最大值為 95.16，僅有一次；最小值為完全未重複(0)，約有 1/3(33 次，33%)；在判斷相關書目方面，其重覆性平均值為 5.72，標準差為 10.45，其中最大值為 55.56；而完全未重複(0%)的發生頻率高達近 2/3(62 次，62%)。在次數分布方面，表 13 分別列出檢出與判斷相關書目重覆性的分組次數分配情形：

表 13 計畫檢索者檢索結果重複性次數分配表

重覆性(%)	檢出書目			判斷相關書目		
	次數	百分比(%)	累積百分比(%)	次數	百分比(%)	累積百分比(%)
0-4	52	52%	52%	66	66%	66%
5-9	17	17%	69%	10	10%	76%
10-14	6	6%	75%	10	10%	86%
15-19	8	8%	83%	8	8%	94%
20-24	4	4%	87%	2	2%	96%
25-29	3	3%	90%	1	1%	97%
35-39	4	4%	94%	-	-	97%
40-44	1	1%	95%	1	1%	98%
45-49	2	2%	97%	-	-	98%
50-54	-		97%	1	1%	99%
55-59	1	1%	98%	1	1%	100%
60-64	1	1%	99%	-	-	100%
95-99	1	1%	100%	-	-	100%
總計	100	100%	100%	100	100%	100%

以分組次數分配表觀之，如表 13 所示，可知無論是檢出書目或判斷相關書目，其重複性均以 0-4 者最多，且已過半數(各占檢出書目的 52% 及判斷相關書目的 66%)，檢出書目重複性在 30 以下的比例達九成(90%)；判斷相關書目重複性在 30 以下的比例更高達 97%，可見絕大部分的檢索結果重複性均集中在 30 以下。判斷相關書目方面的重複性更低於檢出書目的數據，顯示計畫檢索者從檢出書目到瀏覽後判斷相關的書目，其決策結果的一致程度趨於降低。

與 Saracevic & Kantor 等人的研究結果相較，在重複性的部分，本研究無論就檢出書目(10.48%)或判斷相關書目(5.72%)而言均低於 Saracevic & Kantor 等人的研究結果(17%；18%)(註26)，可能是由於其以研究計畫之工作人員作為研究對象，因此受試者在檢索策略的的個別差異程度不如本研究以一般終端使用者作為研究對象的情形；其次，由於本研究無法完全避免資料蒐集期間資料庫新增之書目的影響，亦可能是造成檢索結果歧異性增加導致重複性較低之因。此外，本研究由檢出到判斷相關，其重複性明顯減少，不同於 Saracevic & Kantor 之研究的略增。

在分組次數分配方面，與 Saracevic & Kantor 等的研究中，在檢出書目與判斷相關書目的重複性 0-5% 均占 59% 的比例相較(註27)，本研究的數據

分別為 52%與 66%，是與其研究結果較為接近的部分，可見較大的差異是出現於 5%以上的部分。

若從整個檢索問題的角度來看，可將各檢索問題的檢出書目與判斷相關書目重覆性比對情形列出如表 14 所示：

表 14 各檢索問題檢索結果之重複性

題次	檢出書目	判斷相關書目	升降幅度
A	7.24	3.12	-56.9
B	2.36	2.60	+10.2
C	5.88	0.38	-93.5
D	8.39	4.99	-40.5
E	10.37	8.00	-22.9
F	9.14	5.79	-36.7
G	19.15	9.56	-50.1
H	12.96	11.77	-9.2
I	22.36	7.58	-66.1
J	6.97	3.45	-50.5
總平均	10.48	5.72	-45.4

單位：%

表 14 顯示雖整體而言，從檢出書目到判斷相關書目的重覆性總平均由 10.48%降為 5.72%，其中降幅最大的檢索問題 C 足足降低了檢出書目的 93.5%，然而仍可看出十個題目中仍有 1 題(編號 B)的平均數是提高的，表示其檢出書目的歧異性大於從中判斷相關書目的歧異性，即使瀏覽的書目不同，亦可能選出相同的書目。證實並非每一檢索問題從檢出書目到判斷相關書目的過程中，檢索者之間的重覆性必然會降低。

除了前述兩兩比對的情況外，本研究另就檢出書目與判斷相關書目分別計算了二人以上獲得同一檢索結果的書目篇數，檢視其遞檢情形。依前述計算方法將各檢索問題二人以上共同檢索結果篇數的平均值列於表 15：

表 15 各檢索問題共同檢索結果平均篇數隨檢索者增加的遞減情形

檢索問題 編號	檢索者人數 (數據組數)	共同檢出書目篇數		共同判斷相關書目篇數	
		總數	平均數	總數	平均數
A	2人重覆(10)	33	3.30	4	0.40
	3人重覆(10)	4	0.40	0	0
	4人重覆(5)	0	0	0	0
	5人重覆(1)	0	0	0	0
B	2人重覆(10)	9	0.90	2	0.20
	3人重覆(10)	0	0	0	0
	4人重覆(5)	0	0	0	0
	5人重覆(1)	0	0	0	0
C	2人重覆(10)	16	1.60	1	0.10
	3人重覆(10)	1	0.10	0	0
	4人重覆(5)	0	0	0	0
	5人重覆(1)	0	0	0	0
D	2人重覆(10)	23	2.30	4	0.40
	3人重覆(10)	6	0.60	0	0
	4人重覆(5)	1	0.20	0	0
	5人重覆(1)	0	0	0	0
E	2人重覆(10)	58	5.80	20	2.00
	3人重覆(10)	9	0.90	2	0.20
	4人重覆(5)	1	0.20	0	0
	5人重覆(1)	0	0	0	0
F	2人重覆(10)	33	3.30	7	0.70
	3人重覆(10)	3	0.30	1	0.10
	4人重覆(5)	0	0	0	0
	5人重覆(1)	0	0	0	0
G	2人重覆(10)	77	7.70	15	1.50
	3人重覆(10)	21	2.10	2	0.20
	4人重覆(5)	3	0.60	0	0
	5人重覆(1)	0	0	0	0
H	2人重覆(10)	39	3.90	18	1.80
	3人重覆(10)	11	1.10	4	0.40
	4人重覆(5)	1	0.20	0	0
	5人重覆(1)	0	0	0	0
I	2人重覆(10)	48	4.80	2	0.20
	3人重覆(10)	10	1.00	0	0
	4人重覆(5)	0	0	0	0
	5人重覆(1)	0	0	0	0
J	2人重覆(10)	20	2.00	3	0.30
	3人重覆(10)	2	0.20	0	0
	4人重覆(5)	0	0	0	0

	5人重覆	0	0	0	0
總平均	2人重覆(100)	0.76			
	3人重覆(100)	0.09			
	4人重覆(50)	0			
	5人重覆(1)	0			

表 15 顯示從 2 人到 5 人檢索結果重覆平均篇數的遞檢情形，可知除兩兩比對的平均檢出結果 3.56 篇大於 1 之外，其餘均不到 1 篇，其中 5 人產生重覆檢出書目的情形是不存在的，而在判斷相關書目的部分，更從 3 人重覆時即開始篇數為零了。若檢視個別平均數為零，也就是沒有重覆檢索結果的比率來看，從 3 人即開始出現平均數為零的情形，有 7 個數據，占 35%，到了 4 人時，重覆檢索結果平均數為零的情形提高到 16 個數據(占 80%)，到了 5 人時，更高達 100%。

進一步檢視從 2 人到 5 人共同檢索結果的分布情形，如表 16 所示，表的左欄為本研究各次檢索之間曾出現過的 2 人至 5 人共同檢索結果的篇數，其右則依序列出 2 人至 5 人共同檢出書目及判斷相關篇數的比對組數及占有組數之百分比。由該表整體觀之，共同檢索結果的分布大部分都是集中在篇數較少的部份，其中個數最少的為零，也就是在比對的數個檢索之間並沒有任何一篇相同之檢出書目或判斷相關書目；從 2 人到 5 人都有此種情況，且比對的人數越多，共同檢索結果篇數為零的比例也越高，從 2 人比對共同檢出/判斷相關書目篇數為零各占 33%與 62%到 5 人比對時均已到 100%，也就是沒有任何一篇書目同時被 5 個人檢出與判斷相關。而共同檢索結果篇數最多的則是 28 篇，在 2 人比對的共同檢出書目中出現過一次，事實上比對的人數越多，共同檢索結果的篇數則越少，且由 2 人到 3 人之間的變化最大，因為僅有 2 人比對的檢出書目才有 11 到 28 篇共同檢索結果的狀況；3 人比對時最高的共同篇數有出現於檢出書目的 11 篇，亦是只有一次，且其次僅有 5 篇；4 人比對時的最高的共同篇數只有出現於檢出書目的 4 篇，且其次僅有 1 篇；5 人比對時則無共同檢索結果。以上各欄中最高篇數都是出現於檢出書目，若只看判斷相關書目，其共同篇數最高者則僅有 2 人比對時的 8 篇；3 人比對時最高則只有 2 篇；4 人(含)以上比對時則已無共同判斷相關書目了，共同檢出與判斷相關書目篇數的落差與前述檢索精確率偏低的現象相符。

表 16 2 至 5 人共同檢索結果篇數之次數分配表

共同檢索 結果篇數	2 人共同 (共 100 組)		3 人共同 (共 100 組)		4 人共同 (共 50 組)		5 人共同 (共 10 組)	
	檢出書目	判斷相關書目	檢出書目	判斷相關書目	檢出書目	判斷相關書目	檢出書目	判斷相關書目
0 篇	33 (33%)	62 (62%)	72 (72%)	94 (94%)	47 (94%)	50 (100%)	10 (100%)	10 (100%)
1 篇	17 (17%)	22 (22%)	14 (14%)	4 (4%)	2 (4%)	0	0	0
2 篇	11 (11%)	9 (9%)	3 (3%)	2 (2%)	0	0	0	0
3 篇	4 (4%)	1 (1%)	3 (3%)	0	0	0	0	0
4 篇	7 (7%)	2 (2%)	3 (3%)	0	1 (2%)	0	0	0
5 篇	7 (7%)	1 (1%)	4 (4%)	0	0	0	0	0
6 篇	2 (2%)	2 (2%)	0	0	0	0	0	0
7 篇	5 (5%)	0	0	0	0	0	0	0
8 篇	2 (2%)	1 (1%)	0	0	0	0	0	0
9 篇	1 (1%)	0	0	0	0	0	0	0
11 篇	2 (2%)	0	1 (1%)	0	0	0	0	0
12 篇	2 (2%)	0	0	0	0	0	0	0
13 篇	1 (1%)	0	0	0	0	0	0	0
14 篇	1 (1%)	0	0	0	0	0	0	0
19 篇	1 (1%)	0	0	0	0	0	0	0
20 篇	1 (1%)	0	0	0	0	0	0	0
25 篇	1 (1%)	0	0	0	0	0	0	0
28 篇	1 (1%)	0	0	0	0	0	0	0

以每一欄觀之，從 2 人到 5 人，其共同檢索結果篇數均是以零個的情況最多，其次為 1 篇，但百分比驟減，比對人數越多，減的幅度越大。共同檢索結果篇數在 1 篇以上者的組數略有變動，並無一定的趨勢，但比例均不高。

陸、綜合討論

選用詞彙的一致性與檢索結果二者之間是否有必然的關係？也就是說，選用詞彙一致性高的情形下，是否也會有較高的檢索結果重覆性？選用詞彙一致性高的檢索問題，是否所用的檢索詞彙數明顯較多或較少？檢索效益是否較高？本研究透過統計分析來探討這些問題。

在資料的處理上，為簡化問題，選用詞彙一致性的部分，僅以第二層次的數據為代表，理由在於詞彙一致性第一層次的標準過於嚴苛，選用詞彙的一致性偏低，而檢索概念的兩個層次已涉及語意的擴大縮小與相關，又是過於放寬，因此以選用詞彙一致性的第二層次作為分析的素材；其次，由於選用詞彙一致性的有效數據多於檢索結果重覆性的有效數據，因此僅就二者均有的部分(各為 50 次檢索的 100 個數據)加以分析。

在探討選用詞彙的一致性對檢索結果重複性的影響上，為便於統計分析，將選用詞彙的一致性依所占數據組數(此處為 100 組)的百分比平均分為低、中、高度三組，分別測試選用詞彙的一致性對檢出書目及判斷相關書目重複性的影響。其中低度一致性指選用詞彙一致性範圍在 0-10.99%，到達累積數據組數的 33% 處；中度一致性指選用詞彙一致性範圍在 11.00-20.99%，到達累積數據組數的 66% 處；高度一致性指選用詞彙一致性範圍在 20.00% 以上，到達累積數據組數的 100% 處。結果如表 17 所示：

表 17 以變異數分析表顯示檢索結果重複性與選用詞彙一致性之關係。由表中可知，就檢出書目重複性而言，重複性的平均值以選用詞彙高度一致性組最高(平均值為 16.84，標準差為 19.45)，次為中度一致性組(平均值為 10.09，標準差為 15.26)，低度一致性組最少(平均值為 4.33，標準差為 8.11)，此差異已達統計上之顯著水準($p=0.004$)，可見在檢出書目方面，選用詞彙一致性的高低確實可影響檢出書目的重複性，檢出書目的重複性會隨選用詞彙一致性的提高而增加。

表 17 選用詞彙一致性對檢出書目重複性與判斷相關書目重複性
變異數分析表

項目	平均值	標準差	F 值	P 值
檢出書目重複性				
選用詞彙低度一致性	4.33	8.11	5.783	.004
選用詞彙中度一致性	10.09	15.26		
選用詞彙高度一致性	16.84	19.45		
判斷相關書目重複性				
選用詞彙低度一致性	3.38	10.29	2.956	.057
選用詞彙中度一致性	4.56	10.65		
選用詞彙高度一致性	9.13	9.81		

另就判斷相關書目重複性而言，重複性的平均值亦以高度一致性組最高(平均值為 9.13，標準差為 9.81)，次為中度一致性組(平均值為 4.56，標準差為 10.65)，低度一致性組最少(平均值為 3.38，標準差為 10.29)，此差異雖未到達但接近統計上之顯著水準。(p=0.057)，因此雖然判斷相關書目的重複性隨選用詞彙一致性的提高而增加，增加的程度並未達統計上的顯著水準。

在選用詞彙的一致性與檢索結果重複性的相關性方面，此處亦是以 100 個數據為基礎計之，結果如表 18 所示，無論檢出書目重複性或判斷相關書目重複性，均和選用詞彙的一致性呈現正相關，且已達到統計上之顯著水準，顯示選用詞彙一致性較高者，檢索結果重複性亦較高。其中檢出書目重複性與選用詞彙一致性的相關係數，略高於判斷相關書目重複性與選用詞彙一致性的相關係數，意謂選用相同的詞彙亦檢出相同書目的可能性，高於選用相同的詞彙亦選出相同書目的可能性。

表 18 選用詞彙的一致性與檢索結果重複性之相關係數表

相關係數	檢出書目重複性	判斷相關書目重複性
選用詞彙的一致性	0.330**	0.212*

**表示達到統計上 0.01 之顯著水準

*表示達到統計上 0.05 之顯著水準

此外，可就檢索問題整體來看詞彙一致性與使用檢索詞彙個數的關係。在各檢索問題的詞彙一致性與其檢索詞彙平均數的相關性方面，統計測試的結果如表 19 所示，二者間呈負相關，但並未達到統計上之顯著水準，意謂選用詞彙一致性高的檢索問題，其檢索詞彙總平均數較低，詞彙一致性較高表示不同檢索者所選表達概念的詞彙較為集中，可能由於表達概念的詞彙較為明確，眾人所想到的都差不多，因此無須去思考其他的檢索詞彙，但是這一傾向並不顯著。

表 19 選用詞彙的一致性與檢索詞彙平均數之相關係數表

相關係數	檢索詞彙平均數
選用詞彙的一致性	-0.395

而從各檢索問題的檢索效益分別與其選用詞彙一致性及檢索結果(檢出書目與判斷相關書目)重複性的相關性統計測試的結果(見表 20)可知，雖然三項均呈正相關，但僅有一項--檢索效益與判斷相關書目重複性之間--達到統計上之顯著水準，印證了 Saracevic 與 Baxter「一致性的高低並不意味檢索品質的高低」的說法(註28)，至於其中檢索精確率與判斷相關書目重複性之間的正相關現象，是否有更多的意涵，或純是巧合，尚待進一步的研究來探討。

表 20 選用詞彙一致性與檢索精確率平均數之相關係數表

相關係數	檢索精確率平均數
選用詞彙的一致性	0.164
檢出書目的重複性	0.251
判斷相關書目的重複性	0.730*

*表示達到統計上 0.05 之顯著水準

綜合本研究之分析結果可知，選用詞彙一致性的對檢出書目的重複性確有影響，但對判斷相關書目的重複性影響不顯著；選用詞彙的一致性無論與檢出書目或判斷相關書目的重複性均呈現顯著正相關；就各檢索問題整體來看，詞彙一致性與使用檢索詞彙個數二者的關係並未達到統計上之顯著水準；各檢索問題整體的檢索效益與判斷相關書目重複性之間呈現顯著的正相關；各檢索問題整體的檢索效益分別與選用詞彙一致性及檢出書目的相關性則不顯著。

柒、結論

本研究根據統計結果，分別針對計畫檢索者選用詞彙概念的一致性、計畫檢索者檢索結果的分析、選用詞彙一致性與檢索結果(重複性、精確率等)之關係等方面作一總結。

在計畫檢索者選用詞彙與概念的一致性方面，以檢索詞彙及概念各分二個層次來呈現，其中逐字比對的詞彙第一層次，其一致性平均值與標準差為 22.14%與 15.49%；不計語尾變化的詞彙第二層次，其一致性平均值與標準差為 35.00%與 18.02%；不計廣狹義關係的概念第一層次，其一致性平均值與標準差為 39.00%與 16.43%；不計相關關係的概念第二層次，其一致性平均值與標準差為 46.50%與 17.55%；研究結果顯示四個層次均無任一數據達 100%者；在分組分配方面，一致性最低的組別所占百分比從詞彙第一層次到概念第二層次依次大幅遞減；一致性在 60%以上者所占百分比則依次遞增；所占百分比居首者，亦隨著詞彙到概念層次，由低度一致性組(0-19.99%)轉移到中度一致性(40-59.99%)，與 Fidel(註29)、Saracevic 與 Kantor(註30)及 Iivonen(註31)等研究結果相較，本研究所得之檢索詞彙概念一致性偏低，且越到較後的層次差距越大。本研究另計算出二人以上共同詞彙數的遞減情形，結果顯示從 4 人開始，共同檢索詞彙數才開始出現平均數為零的情形，到 6 人時，共同檢索詞彙數平均數為零的部份已高達 77.8%。綜而言之，本研究由詞彙到概念的一致性增長趨勢同其他研究，但各層次的數據普遍低於其他類似研究。

在計畫檢索者檢索結果分析部分，本研究各檢索結果的精確率平均值為 34.36%，標準差為 19.80%。其中 90%的精確率都在 60%以下，故大部分的精確率都集中在較低的部分；檢索結果的重覆性包括檢出書目與判斷相關書目各自的重覆性，前者重覆性的平均值為 10.48%，標準差為

15.79%；後者重覆性的平均值為 5.72%，標準差為 10.45%，其中亦無任一數據達 100% 者。無論是檢出書目或判斷相關書目，其重複性均以 0-4.99% 者最多，比例均超過半數，絕大部分的檢索結果重複性均集中在 30% 以下。判斷相關書目方面的平均重覆性更低於檢出書目的數據，顯示計畫檢索者從檢出書目到瀏覽後判斷相關的書目，其決策結果的一致程度趨於降低。與 Saracevic & Kantor 等人的研究結果相較，在重覆性的部分，本研究無論就檢出書目或判斷相關書目而言均較低，分組次數分配方面，檢出書目與判斷相關書目重覆性 0-5% 的部份與 Saracevic & Kantor 等的研究結果相近(註32)，可見較大的差異是出現於 5% 以上的部分。本研究另計算出二人以上共同檢索結果的遞減情形，結果顯示從 3 人開始，共同檢索詞彙數才開始出現平均數為零的情形，重覆檢索結果平均數為零的情形到了 5 人時，更高達 100%；綜而言之，本研究由檢出書目與判斷相關書目，重複性減少，不同於 Saracevic & Kantor 等人重複性提高的研究結果。(註33)

至於計畫檢索者選用詞彙一致性與檢索結果之關係分析，本研究的結果顯示，在檢出書目方面，選用詞彙一致性的高低確實可影響檢出書目的重複性，檢出書目的重複性會隨選用詞彙一致性的提高而增加，無論是檢出書目重覆性複性或判斷相關書目重複性，均和選用詞彙的一致性呈現正相關，且已達到統計上之顯著水準，而選用相同的詞彙亦檢出相同書目的可能性，高於選用相同的詞彙亦選出相同書目的可能性。就各檢索問題整體來看，詞彙一致性與使用檢索詞彙個數二者的關係並未達到統計上之顯著水準；各檢索問題整體的檢索效益與判斷相關書目重複性之間呈現顯著的正相關；各檢索問題整體的檢索效益分別與選用詞彙一致性及檢出書目的相關性則不顯著。

註1 Tefko Saracevic and Matthew A. Baxter, "On a Method for Studying the Structure and Nature of Requests in Information Retrieval," Proceedings of the 46th Annual Meeting of the American Society for Information Science 20 (1983) : 22-25.

註2 John G. Keyes, "Using Conceptual Categories of Questions to Measure Differences in Retrieval Performance," Proceedings of the 59th Annual Meeting of the American Society for Information Science 33 (1986) : 238.

註3 Stephen P. Harter, Online Information Retrieval : Concepts, Principles, and Techniques (Orlando, Florida : Academic Press, 1986) : 22.

-
- 註4 Tefko Saracevic and Matthew A. Baxter, "Measuring the Degree of Agreement between Searchers," in Proceedings of the 47th Annual Meeting of the American Society for Information Science, Philadelphia, Pennsylvania, October 21-25, 1984, eds, B. Flood, J. Witiak, & T.H. Hogan (White Plains, NY: Knowledge Industry Publications, 1984), 227-230.
- 註5 Trudi Bellardo and Tefko Saracevic, "Online Searching and Search Output: Relationships between Overlap, Relevance, Recall and Precision," in Proceedings of the 50th Annual Meeting of the American Society for Information Science, Boston, Massachusetts, October 4-8, 1987, ed. Ching-chih Chen (Medford, New Jersey: Learned Information, Inc., 1987), 11-13.
- 註6 Raya Fidel, "What is Missing in Research about Online Searching Behavior?" Canadian Journal of Information Science 12:3/4(1987): 54-61.
- 註7 Tefko Saracevic, P. Kantor, A.Y. Chamis, and D. Trivison. "A Study of Information Seeking and Retrieving. I. Background and Methodology," Journal of the American Society for Information Science 39:3(May 1988): 161-176.
- 註8 Stephen P. Harter, "Search Term Combinations and Retrieval Overlap: A Proposed Methodology and Case Study. Journal of the American Society for Information Science 41:2 (March 1990): 132-146.
- 註9 Mirja Iivonen, "Consistency in the Selection of Search Concepts and Search Terms," Information Processing and Management 31:2(1995): 173-190.
- 註10 Mirja Iivonen, "Completing of Quantitative and Qualitative Methods: Example from a Consistency Study." In Proceedings CoLIS 2: Second International Conference on Conceptions of Library and Information Science: Integration in Perspective, October 13-16, 1996, by The Royal School of Librarianship, 149-160.
- 註11 Saracevic, T.. "Measuring the Degree of Agreement between Searchers." In B. Flood, J. Witiak, & T.H. Hogan eds, in Proceedings of the 47th Annual Meeting of the American Society for Information Science. 227-230(White Plains, NY: Knowledge Industry Publications, 1984), 227.

-
- 註12 同註4，頁228。
同註9，頁177-179。
- 註13 同註4，頁228。
同註8，頁134-135。
- 註14 同註9，頁178-180。
- 註15 同註4，頁227。
- 註16 同註8，頁134-135。
- 註17 黃慕萱，「檢索問題、檢索詞彙及心智模型對資訊尋求和資訊檢索之影響研究I」，行政院國家科學委員會專題研究計畫成果報告(民87年7月)，頁27。
- 註18 同註5，頁11。
- 註19 Tefko Saracevic, P. Kantor, A.Y. Chamis, and D. Trivison. "A Study of Information Seeking and Retrieving. III. Searchers, Searches, and Overlap." Journal of the American Society for Information Science 39:3(May 1988): 197-216.
- 註20 Amanda Spink and Tefko Saracevic. "Interaction in Information Retrieval: Selection and Effectiveness of Search Terms." Journal of the American Society for Information Science 48:8(August 1997): 744.
- 註21 同註9，頁174。
- 註22 同註10。
- 註23 同註9，頁178-180。
- 註24 同註19，頁199。
陳佳君，「檢索詞彙來源與檢索詞彙效益之研究」，國立台灣大學圖書館學研究所，碩士論文，民國85年6月，頁52。
黃慕萱，「檢索問題、檢索詞彙及心智模型對資訊尋求和資訊檢索之影響研究I」，行政院國家科學委員會專題研究計畫成果報告(民87年7月)，頁56。
- 註25 同註9，頁178-180。
- 註26 同註19，頁199。
- 註27 同上註。
- 註28 同上註。

註29 同註6 , 頁60-61。

Raya Fidel, “Individual Variability in Online Search Behavior.” in Proceedings of the 48th Annual Meeting of the American Society for Information Science ed C.A. Parkhurst (White Plains, NY: Knowledge Industry Publications, 1985), 69-72.

註30 同註19 , 頁212。

註31 同註9。

註32 同註19 , 頁199。

註33 同上註。