

電腦科學領域高被引文章與一般 文章之書目計量比較分析

A Bibliometric Analysis of Highly Cited and General
Papers in the Computer Science Field

黃慕萱

Mu-Hsuan Huang

國立臺灣大學圖書資訊學系 教授

Professor

Department of Library and Information Science

National Taiwan University

黃玫溱

Mei-Jhen Huang

國家圖書館研究組 組員

Research Division

National Central Library

【摘要 Abstract】

本研究利用書目計量法分析比較電腦科學領域高被引文章與一般文章之特性，包括基本特性、合著特性，及老化特性，並對臺灣地區機構所屬作者發表文章同樣進行特性探討。高被引文章蒐集自ESI資料庫，包括1996至2006年共2,112篇文章，其中臺灣高被引文章有20篇；一般文章檢索自WOS資料庫，包括1996至2006年共2,200篇文章，其中臺灣一般文章有89篇。研究結果顯示，高被引文章平均首次被引年齡較一般文章低，主要由歐美國家所發表；在合著率方面，高被引文章與一般文章皆呈現上升趨勢，以高被引文章合著率較高，跨國合著率同樣是高被引文章較高；分析高被引文章與一般文章之老化曲線可知，高被引文章被引狀況

投稿日期：2008.8.18；接受日期：2008.10.24

email: 黃慕萱mhhuang@ntu.edu.tw；黃玫溱mjhuang@ncl.edu.tw

較持久，一般文章則是較早達到被引高峰且較早老化。此外，臺灣地區機構所屬作者發表文章特性與所有文章大致相近，但臺灣的高被引文章數在各國之排名遠不如一般文章，且臺灣的一般文章平均被引次數較所有一般文章低，顯示臺灣文章數量雖不少但品質仍有待提升。

This study explores the characteristics of the highly cited papers and general papers in the computer science field; papers by Taiwan authors are specifically analyzed and compared to the entire group of papers. Paper characteristics examined include their general features (e.g., citation conditions, journal impact factors, authorship conditions, etc.) as well as the collaboration and obsolescence conditions of those papers. Data about the highly cited and general papers from 1996 to 2006 are obtained from the Essential Science Indicators (ESI) and Web of Science (WOS) databases. Study results suggest that highly cited papers, mainly published by American and European authors, are cited sooner than general papers. Highly cited papers have higher rates of collaboration and international collaboration than general papers. Moreover, highly cited papers are cited for a longer period of time, while citations to general papers decrease rapidly. Papers by Taiwan authors have similar characteristics. But Taiwan authors published less highly cited papers; the average number of citations of general papers is also lower than that of the overall papers. This suggests the paper quality of Taiwan authors requires improvement.

[關鍵字 Keywords]

書目計量；高被引文章；合著；文獻老化

Bibliometrics; Highly cited papers; Collaboration; Obsolescence

壹、前言

期刊是科學研究傳播的重要管道，而研究者在發表期刊文章時，往往建構於該領域過去累積的知識上，藉由引用他人著作支撐個人論點或作為後續研究基礎，因此引用常用來代表對前人研究的認同或評價，也大致顯示兩文獻之間具有相當程度關聯性。一般而言，文獻被引用顯示其所含的知識或資訊在科學交流中曾被人使用，研究者會引用的文獻通常是關於特定研究主題之重要作品，所以期刊文章被引用次數多寡，可

作為學術水平和價值之測度，反映文章在學術上的重要性，被引用的次數愈多在學術上的價值也愈高（黃慕萱，1994；黃裕惠，2004）。相較之下，一般文章的被引次數偏低，甚至不曾被引用，其文章品質可能不如高被引文章，對研究者與後續研究的影響力也較低，藉由兩者的對照比較更能突顯高被引文章異於一般文章之處，故本研究以一般文章作為高被引文章比較對象。

當前國內外關於高被引文章的研究，主要探討高被引文章在評估之應用，由於社會大眾對科學傑出表現之重視，高被引文章愈來愈常被作為學術評估指標，如用於比較中國大陸各大學研究表現（Zhu, Wu, Zheng, & Ma, 2004）、衡量歐洲各國研究強度（European Commission, 2001，轉引自黃裕惠，2004）等。只有部分研究針對高被引文章之發表時間、各年被引次數等特徵進行討論（Plomp, 1990; Aksnes, 2003），其中合著分析能瞭解近來學術群體研究趨勢，文獻老化分析可觀察文章使用隨時間改變之過程，可能為高被引文章之重要特性。

合著文章融合多位研究者之專業知識，往往被期待具有較高品質（Avkiran, 1997, p. 175），不少研究即指出合著文章與文章被引次數有相當關聯性，合著文章被引次數通常高於單一作者文章，易形成高被引文章（Smart & Bayer, 1986; Narin & Hamilton, 1996; Glänzel, 2002）；然而目前合著與被引次數之研究，多由特定領域中隨機抽樣文章進行分析，只有Aksnes（2003）等極為少數的研究是針對高被引文章分析合著狀況，顯示高被引文章合著特性之討論仍甚缺乏，因此本研究將合著特性列為重要分析項目。

文獻老化分析能呈現文章被引狀況在時間推移下之變化，反映文章價值之更替，通常被引次數較高的文章較晚達到被引高峰，被引次數下降速度也較慢（Aversa, 1985），先前的研究顯示，高被引文章老化曲線包含「快速上升，快速下降」和「緩慢上升，緩慢下降」兩種曲線（Aversa, 1985; Cano & Lind, 1991）；但目前仍少有研究同時分析高被引文章與一般文章之老化現象，而且以歷時法觀察可合理呈現文章後續被引情況（Stinson, 1981，轉引自王淑儀，1998，頁18），因此本研究透過歷時法觀察高被引文章老化特性，並與一般文章進行比較。

近年來政府積極推動科技產業建設，尤其對電腦科學領域發展相當重視，列為重點科研項目，導致愈來愈多研究者投入其中，展現的研究成果更讓電腦科學領域迅速成長。根據美國科學資訊所2006年的統計，臺灣各領域文章數佔世界各領域文章數之比率以電腦科學領域最高，比率為4.13%（張偉嶠，2006），顯示在政府推動與相關因素影響下，電腦

科學領域逐漸受到關注而成為重要的研究範疇，故本研究選擇電腦科學領域進行文章特性分析。

由於目前高被引文章特性研究不多，且甚少研究同時探討與之對應的一般文章，在文章特性掌握上較不全面，因此本研究擬透過書目計量法，分析比較電腦科學領域高被引文章與一般文章之基本特性、合著特性，和老化特性，並探討臺灣地區機構所屬作者發表文章（簡稱臺灣文章）之特性。期望經由本研究之分析，協助研究者瞭解電腦科學領域高被引文章與一般文章特性之差異，進而提供學術界作為領域發展與研究評估之參考。

貳、文獻分析

被引次數一般代表文章的影響力與品質（Wade, 1975, p. 429），所以被引次數高的文章，其影響力與品質相對較佳。在引用文獻網絡中，這些高被引文章構成了網絡中的主結點，它們是對學科發展具關鍵作用之核心文章，抑或學科的奠基文獻（孟連生，1996，頁17）。然而大部分文章很少或從未被後續文獻引用，只有少數文章能得到相當大量的引用。

Aksnes（2003）利用挪威全國科學指標（National Science Indicators）資料庫和全國引用報告（National Citation Report）資料庫取得1981-1996年挪威高被引文章的書目資料，將發表五年後文章被引次數為同領域文章平均被引次數17倍以上者視為高被引文章，統計高被引文章及所有文章的各年平均被引次數，描繪成圖形後發現兩者相似，以此說明高被引文章數量雖少，但對全部文章的平均被引次數影響大；並由引用文獻分析發現，挪威的高被引文章引用來源多為美國及歐洲國家，平均每篇文章被65種不同期刊引用；而且高被引文章所在期刊之影響係數相當高，56%的高被引文章發表於影響係數1.5到4的期刊，20%的高被引文章發表於影響係數大於4的期刊，相較之下其他文章所在期刊的影響係數多不到1。

為瞭解文章被引次數與合著之關聯性，Smart和Bayer（1986）抽取臨床心理領域、管理科學領域、教育評估領域文章，計算文章發表後十年內的被引次數，結果三領域合著文章之平均被引次數皆比單一作者文章高，尤其是管理科學領域。Glänzel（2002）選擇生物醫學、化學、數學領域文章，分析文章出版當年及後續兩年之被引次數，證實這些領域合著文章之被引次數比單一作者文章高，其中數學合著文章之平均被

引次數更超過一般文章兩倍。Aksnes（2003）的研究亦指出文章合著與文章品質之相關性，多數高被引文章為合著文章，高被引文章跨國合著率達63%，高於一般文章國家合著率29%，高被引文章平均作者數為8.9人，高於一般文章平均作者數3.7人。

此外，Rousseau（2001）探討合著與國際合著文章之被引次數，以數學統計領域、理論物理領域、化學領域的文章為研究對象，分析文章被引次數、作者數、及國家數。結果顯示：數學統計領域合著文章之被引次數高於單一作者文章，國際合著文章被引次數亦高於單一國家文章；化學領域的合著文章與國際合著文章之被引次數同樣較高。Katz和Hicks（1997）探討生命科學、自然科學、工程材料等領域之文章，亦證實合著文章之平均被引次數較單一作者文章高，尤其是國際合著文章，被引次數更甚於單一國家文章。

然而，有些研究者對合著文章與被引次數之關係抱持不同意見。Hart（2007）驗證學術圖書館學領域合著與高品質文章之關聯，統計此領域期刊文章發表後十年內之被引用次數，結果指出：未必每種期刊合著文章被引次數均高於單一作者文章被引次數。Bridgstock（1990）以1984-1986年出版之鳥類、海洋學、天文學，和物理期刊文章為對象，分析文章發表後兩年內的被引狀況，發現合著與文章被引次數無顯著相關。Avkiran（1997）探討財務金融領域文章的合著與被引次數關係，統計十四種財務金融期刊之文章，在發表後四年內的被引次數，研究結果同樣指出合著文章之被引次數未必較高，合著文章與單一作者文章之品質沒有明顯差異。

一般而言，文章在出版一段時間後被引機率會大幅下降，老化曲線可呈現文章被引的整個過程。Arunachalam和Singh（1984）以23篇主題為超導電性的高被引文章為研究對象，分析各年被引次數之分布，結果顯示被引次數在發表後第二年快速增加，維持四年的穩定狀態後才下降。Aversa（1985）針對發表於1972年的自然科學領域文章，定義其於1972-1977年間被引用達30次以上者為高被引文章，而後分析此400篇高被引文章之老化曲線。研究顯示：被引次數較多的文章，得到引用的速度較慢，老化速度也較慢；被引次數較少的文章，得到引用的速度較快，老化速度也較快。以集群分析得到兩種老化曲線：

1. 緩慢上升，緩慢下降：文章數佔58%，在出版六年後才達到被引次數比率高峰，之後逐步下降，此曲線平均被引次數為146次。
2. 快速上升，快速下降：文章數佔42%，在出版三年後達到被引次數比率高峰，之後三年急速下降，此曲線平均被引次數為108次。

McCain和Turner (1989) 隨後對分子生物領域期刊文章進行老化分析，挑選1978到1980間出版，且出版後五年內被引次數超過30次的11篇高被引文章，觀察這些文章在發表後八年間的引文分布，經過集群分析後，發現這11篇文章有兩種老化曲線，特徵與Aversa在1985年提出的老化曲線類似，分別為「緩慢上升，緩慢下降」以及「快速上升，快速下降」，前者在文章發表六年後才達到被引高峰，且被引次數比率衰退緩慢；後者在文章發表2-3年內即達被引高峰，但之後快速下降。

Cano和Lind (1991) 探討醫學和生物化學領域10篇高被引文章，及10篇一般文章之每年被引次數與成長率，結果顯示高被引文章以不甚平滑的直線上升，而一般文章上升曲線更加不平滑。研究再針對高被引文章分析老化曲線，結果顯示不同學科文章的老化曲線不盡相同，醫學領域文章的被引次數一開始急速上升，六年內累積被引次數即達總數的四分之三，但之後被引次數增加趨緩；生物化學領域文章六年內累積被引次數不到總數的三分之一，但被引次數維持原先的速度增加，持續力較佳。

由上述文獻可知，不同學科領域本質不同，合著與文章被引次數之關係未必在任何領域皆成立，但大致而言，合著文章的被引次數常高於單一作者文章，國際合著文章之被引次數亦常高於單一國家文章，合著或國際合著文章易形成高被引文章。此外，在自然科學領域及分子生物領域中，高被引文章皆呈現「緩慢上升，緩慢下降」與「快速上升，快速下降」兩種老化曲線，但其他領域之老化曲線還需進一步探究。

參、研究方法

本研究係以書目計量法分析比較電腦科學領域高被引文章與一般文章之特性，包含基本特性、合著特性、及老化特性，並對臺灣電腦科學領域文章同樣進行分析。本研究蒐集的資料可分為三個集合，包括高被引文章集合、一般文章集合，與臺灣文章集合。

高被引文章在本研究中為近十年各領域被引次數排名前1%的文章，高被引文章集合蒐集自Essential Science Indicators (簡稱ESI) 資料庫，出版年涵蓋1996年至2007年，由於2007年出版之文章僅有1篇故予以刪除，餘2,112篇文章則為本研究分析基本特性和合著特性之研究樣本。在探討文章老化特性時，為降低文獻還未進入老化的可能性，參考Journal Citation Reports (簡稱JCR) 資料庫，計算電腦科學領域期刊之平均半衰期，結果為6.42年，因此本研究不採用近七年之高被引文章，換言之，

本研究以實際出版年1996年至1999年共444篇高被引文章為對象，探討文章老化特性。

一般文章在本研究中意指非高被引文章，一般文章集合檢索自JCR資料庫與Web of Science（簡稱WOS）資料庫，由於WOS資料庫無法直接設定學科領域進行檢索，因此本研究先透過JCR資料庫計算電腦科學領域七個次領域之文章比例，用以分配每年200篇之抽樣文章，再將次領域所有期刊名加上文章出版年1996至2006年，在WOS資料庫中檢索並進行隨機抽樣，一般文章基本特性和合著特性分析即採用此2,200篇文章。至於文章老化特性之分析，則是由上述抽樣文章中，檢索出版年為1996至1999年的文章，共800篇。

臺灣文章是臺灣地區機構所屬作者發表之文章，臺灣文章集合是由上述高被引文章集合及一般文章集合中，檢索作者所屬機構所在國家為臺灣的文章，只要有一位作者出自臺灣，該篇文章即納入臺灣文章集合。基本特性和合著特性分析對象，為出版年介於1996至2006年的文章，符合此條件之臺灣高被引文章共20篇，臺灣一般文章共89篇；老化特性分析對象則限於1996至1999年出版之文章，臺灣高被引文章共5篇，臺灣一般文章共29篇。

肆、高被引文章與一般文章特性比較分析

本研究以電腦科學領域高被引文章與一般文章為對象，應用書目計量法分析比較文章基本特性、合著特性及老化特性，最後再將臺灣地區機構所屬作者發表之電腦科學領域文章與所有電腦科學領域文章作一比較。

一、高被引文章與一般文章基本特性比較

基本特性比較項目包括被引次數分布、首次被引時間、引用文獻數、期刊影響係數，及國家文章數分析。用以分析之電腦科學領域高被引文章共2,112篇，一般文章共2,200篇，出版年介於1996至2006年。

(一)被引次數分布

表1比較各出版年電腦科學領域高被引文章與一般文章之平均被引次數。高被引文章的總平均被引次數為65.94次，遠高於一般文章3.48次；從文章被引次數分布來看，高被引文章的被引次數多分布於100次內，一般文章的被引次數則集中在10次以內，t檢定顯示高被引文章與一般文章在文章被引次數上達顯著差異（ $t=18.139, p=.000<.05$ ）。

高被引文章在時間累積下，較早出版的文章被引次數普遍較高，1996至1999年，及2001年平均被引次數皆達百次以上，2000、2002、2003年平均被引次數亦有50次以上，且由於部分高被引文章的被引次數特別高，所以高被引文章各年平均被引次數的差距非常大，最高為160.19次，最低為5.47次，差距高達154.72次；相對地，一般文章由於被引次數普遍偏低，被引次數隨時間遞減的情況較不明顯，1996至2000年平均被引次數在4次以上，2004年後平均被引次數降至3次以下，各年平均被引次數最高與最低分別只有5.56次及0.43次，差距僅為5.13次，較高被引文章各年平均被引次數的差距小。

(二)首次被引時間

比較電腦科學領域高被引文章與一般文章首次被引年齡，一般文章被引次數為0之798篇文章未列入分析。表2顯示高被引文章平均首次被引年齡為1.69，小於一般文章之3.23，高被引文章最遲於年齡4首次被引，一般文章則有18.92%在年齡5以後才首次被引。此外，91.05%高被引文章首次被引時間集中於前兩年，一般文章則有半數的首次被引時間在年齡2或3，但年齡4以上之各年齡首次被引文章亦超過3%，加上高被引文章與一般文章之首次被引時間標準差分別為.604和1.798，顯示高被引文章首次被引時間多比一般文章早，通常集中在前幾年，但一般文章相對分散，有可能發表多年後才首度被引，甚至不曾被注意而從未被引。t檢定可知，高被引文章與一般文章在首次被引年齡上達顯著差異（ $t=-30.713, p=.000<.05$ ）。

表1 高被引文章與一般文章各年分布

年份	高被引文章		一般文章	
	文章數	平均被引次數	文章數	平均被引次數
1996	21	160.19	200	4.64
1997	121	137.46	200	4.23
1998	150	154.61	200	5.26
1999	152	110.72	200	4.58
2000	164	98.76	200	5.56
2001	177	116.53	200	3.97
2002	170	67.95	200	3.40
2003	256	53.63	200	3.01
2004	357	29.41	200	2.01
2005	372	15.33	200	1.26
2006	172	5.47	200	0.43
總計	2,112	65.94	2,200	3.48

表2 高被引文章與一般文章首次被引年齡分布

首次被引 文章年齡	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
1年	861	40.77%	148	10.56%
2年	1,062	50.28%	451	32.19%
3年	175	8.29%	329	23.48%
4年	14	0.66%	208	14.85%
5年	0	0.00%	110	7.85%
6年	0	0.00%	70	5.00%
7年	0	0.00%	42	3.00%
8年以上	0	0.00%	43	3.07%
總計	2,112	100.00%	1,401	100.00%
平均首次被引 文章年齡	1.69		3.23	

註：被引次數為0之文章未列入計算。

(三) 引用文獻數

表3比較電腦科學領域高被引文章與一般文章之引用文獻數分布。高被引文章平均引用文獻數為35.35篇，高於一般之21.17篇；一般文章引用文獻數少於20篇之比率大於高被引文章，高被引文章則是引用文獻數在20篇以上的比率大於一般文章，且近一成高被引文章的引用文獻超過70篇，加上高被引文章與一般文章之引用文獻數標準差分別為35.500和17.606，顯示高被引文章的引用文獻數較分散且部分特別高，一般文章的引用文獻數較集中偏低。由t檢定可知高被引文章與一般文章的引用文獻數達顯著差異（ $t=16.497, p=.000<.05$ ）。

(四) 期刊影響係數

表4比較電腦科學領域高被引文章與一般文章來源期刊之影響係數排名百分比。高被引文章來源期刊共155種，其中24種為非電腦科學領域專門期刊，9種未收錄於2006年JCR資料庫，其餘122種電腦科學領域專門期刊的高被引文章，平均期刊影響係數排名為15.07%，五成高被引文章的期刊影響係數在前10%，另有近三成文章在前20%，排名集中且偏高；一般文章由於是參考JCR電腦科學領域期刊進行隨機抽樣，所以期刊影響係數排名分布均勻，平均為49.06%。t檢定顯示高被引文章與一般文章的期刊影響係數排名達顯著差異（ $t=-41.736, p=.000<.05$ ）。

表3 高被引文章與一般文章引用文獻數分布

引用文獻數	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
0-9	165	7.81%	520	23.64%
10-19	486	23.01%	682	31.00%
20-29	520	24.62%	501	22.77%
30-39	361	17.09%	248	11.27%
40-49	216	10.23%	127	5.77%
50-59	121	5.73%	51	2.32%
60-69	70	3.31%	34	1.55%
70以上	173	8.19%	37	1.68%
總計	2,112	100.00%	2,200	100.00%
平均引用文獻數	35.35		21.17	

表4 高被引文章與一般文章來源期刊之影響係數分布

影響係數 排名百分比	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
0%-9%	653	50%	239	10.86%
10%-19%	365	28%	239	10.86%
20%-29%	82	6%	255	11.59%
30%-39%	50	4%	157	7.14%
40%-49%	40	3%	264	12.00%
50%-59%	41	3%	190	8.64%
60%-69%	25	2%	201	9.14%
70%-79%	30	2%	241	10.95%
80%-89%	5	0.39%	229	10.41%
90%以上	3	0.23%	185	8.41%
總計	1,294	100%	2,200	100.00%
平均影響係數 排名百分比	15.07%		49.06%	

註：非電腦科學領域期刊或未被2006年JCR收錄之電腦科學領域期刊未列入計算。

表5 高被引文章與一般文章國家分布

國家	高被引文章		一般文章		高被引與一般文章相對比值
	百分比 (N=2,112)	排名	百分比 (N=2,200)	排名	
美國	61.67%	1	39.27%	1	1.57
英國	8.42%	2	5.91%	4	1.42
德國	7.19%	3	6.32%	3	1.14
法國	5.92%	4	4.86%	7	1.22
加拿大	4.78%	5	5.05%	6	0.95
中國大陸	3.45%	6	5.32%	5	0.65
以色列	3.27%	7	1.50%	17	2.18
日本	3.12%	8	6.77%	2	0.46
義大利	3.03%	9	4.50%	8	0.67
瑞士	2.93%	10	1.23%	21	2.38
澳大利亞	2.41%	11	2.59%	10	0.93
荷蘭	2.32%	12	2.18%	13	1.06
瑞典	2.13%	13	1.50%	17	1.42
西班牙	1.80%	14	2.32%	11	0.78
丹麥	1.70%	15	0.59%	30	2.88
比利時	1.51%	16	0.91%	22	1.66
南韓	1.42%	17	2.32%	11	0.61
蘇格蘭	1.09%	18	1.32%	20	0.83
新加坡	0.95%	19	1.91%	14	0.50
臺灣	0.95%	19	4.05%	9	0.23

(五) 國家文章數

統計各國電腦科學領域高被引文章數，選取前二十名國家，並列出該國一般文章數，及高被引文章率和一般文章率之相對比值，表5顯示美國在高被引文章與一般文章中均排名第一，英國和德國在高被引文章中排名二和三，在一般文章為第四和第三，法國和加拿大在高被引文章中分居四、五，在一般文章中排名第七和第六。此外，在一般文章排名前二十的國家，部分並未在高被引文章中排名前二十，包括印度、俄羅斯，和希臘；相反地，瑞士、丹麥、比利時則是在一般文章中未排名前二十，但在高被引文章有排名前二十。

從高被引文章率與一般文章率之相對比值來看，有些國家相對比值較高，如丹麥、瑞士、以色列的比值分別為2.88、2.38，和2.18，表示這些國家的研究品質可能相對勝於其他國家；有些國家則是高被引文章率與一般文章率之相對比值較低，如日本、新加坡、臺灣等，其比值分

別為0.46、0.50，和0.23，表示這些國家的研究品質可能相對比其他國家差。

二、高被引文章與一般文章合著特性比較

合著特性分析內容包括合著率與平均作者數、跨國合著率與平均國家數。用以分析之電腦科學領域高被引文章共2,112篇，一般文章共2,200篇，出版年介於1996至2006年。

(一)合著率與平均作者數

表6比較電腦科學領域高被引文章與一般文章之作者數分布與百分比，高被引文章與一般文章之合著率分別為88.83%和76.91%，以高被引文章合著率較高，且卡方檢定顯示高被引文章與一般文章的合著或單一作者文章數有顯著差異 ($\chi^2=107.178, p=.000<.05$)，被引次數相當高的高被引文章，其合著率明顯高於被引次數偏低之一般文章，此結果與Smart和Bayer (1986)、Glänzel (2002)、Aksnes (2003)等研究者的發現相似。

高被引文章平均作者數為3.44人，多於一般文章之2.52人，66.28%的高被引文章與82.13%的一般文章作者數在3人以下，但高被引文章有5-9位作者及10位以上作者之文章比率高於一般文章，高被引文章由多位作者合著之比率較一般文章高。t檢定顯示高被引文章與一般文章的作者數達顯著差異 ($t=13.458, p=.000<.05$)。

表6 高被引文章與一般文章作者數分布

作者數	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
1-4	1,698	80.40%	2,035	92.50%
1	236	11.17%	508	23.09%
2	649	30.73%	791	35.95%
3	517	24.48%	508	23.09%
4	296	14.02%	228	10.36%
5-9	356	16.86%	156	7.09%
10以上	58	2.74%	9	0.41%
總計	2,112	100.00%	2,200	100.00%
平均作者數	3.44		2.52	

分析電腦科學領域高被引文章與一般文章從1996至2006年之合著率變化，由圖1可知，高被引文章與一般文章之合著率大致呈現上升趨勢，高被引文章合著率前期約為80%，後期提高到90%，一般文章則是由前期約75%，增加至後期80%，顯示高被引文章合著率之增加幅度大於一般文章合著率；此外，高被引文章合著率從1996至2006年皆高於一般文章，此觀察呼應高被引文章平均合著率高於一般文章平均合著率之結果。

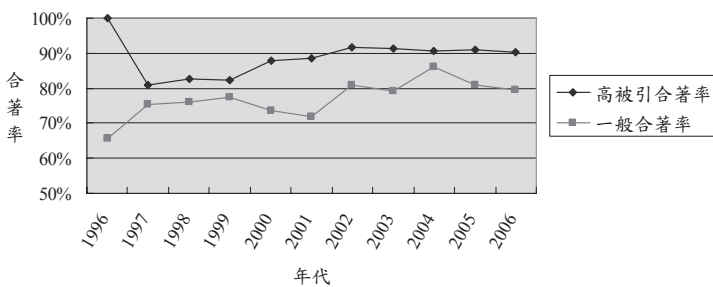


圖1 高被引文章與一般文章合著趨勢

(二) 跨國合著率與平均國家數

表7比較電腦科學領域高被引文章與一般文章之國家數分布與百分比。無論是高被引文章或一般文章，單一國家文章數皆多於跨國合著文章數，但高被引文章之跨國合著文章數，仍多於一般文章之跨國合著文章數，高被引文章跨國合著率為23.48%，一般文章為16.64%，卡方檢定顯示高被引文章與一般文章的跨國合著或單一國家文章數有顯著差異 ($\chi^2=31.153, p=.000<.05$)，被引次數相當高的高被引文章，其跨國合著率高於被引次數偏低之一般文章，此結果和Katz和Hicks (1997)、Rousseau (2001)、Aksnes (2003) 等人的研究發現相似。

在作者所屬國家數量方面，高被引文章平均有1.29個國家，一般文章平均有1.19個國家，高被引文章與一般文章皆以2個國家合著為最多，但高被引文章由3個、4個，或5個以上國家合著之比率皆高於一般文章，顯示高被引文章由多個國家合著之比率較一般文章高。由t檢定可知高被引文章與一般文章的國家數達顯著差異 ($t=6.167, p=.000<.05$)。

表7 高被引文章與一般文章國家數分布

國家數	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
1	1,616	76.52%	1,834	83.36%
2	411	19.46%	319	14.50%
3	69	3.27%	46	2.09%
4	12	0.57%	0	0.00%
5以上	4	0.20%	1	0.05%
總計	2,112	100.00%	2,200	100.00%
平均國家數	1.29		1.19	

圖2呈現電腦科學領域高被引文章與一般文章從1996至2006年之跨國合著率變化趨勢。高被引文章與一般文章之跨國合著率數年來雖有波動，但大體而言沒有明顯上升或下降趨勢，高被引文章之跨國合著率維持於兩成多，一般文章之跨國合著率則維持在一成多，沒有因為通訊科技等的發達而有顯著變化；此外，高被引文章跨國合著率長期高於一般文章，此觀察呼應上述發現：高被引文章平均跨國合著率高於一般文章平均跨國合著率。

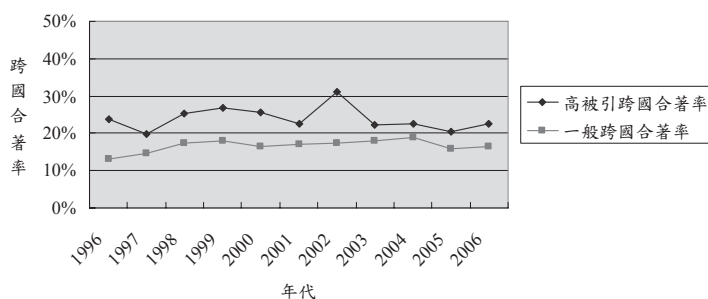


圖2 高被引文章與一般文章跨國合著趨勢

三、高被引文章與一般文章老化特性比較

老化特性分析內容包括老化曲線以及文獻半衰期。用以分析之電腦科學領域高被引文章共444篇，一般文章共800篇，出版年介於1996至1999年。

(一)老化曲線

圖3為電腦科學領域高被引文章及一般文章之老化曲線，高被引文章呈現快速上升而後持平、持續上升兩種老化曲線，一般文章則呈現前期上升而後略降之老化曲線。從被引次數比率上升幅度來看，在年齡4之前，一般文章被引次數比率上升最多，其次是高被引文章老化曲線一，高被引文章老化曲線二之被引次數比率上升最少，顯示一般文章在發表後前幾年的被引增加率大於高被引文章。

在年齡4和年齡5時，一般文章老化曲線與高被引文章老化曲線一先後進入高原期，在此期間高被引文章老化曲線一之被引次數比率略高於一般文章老化曲線，且維持一定的被引次數比率；一般文章老化曲線則是較早進入高原期，且被引次數比率在年齡7、8時開始略降；與此同時，高被引文章老化曲線二之被引次數比率卻始終穩定地上升，還未出現下降現象。顯示一般文章相較於高被引文章較早達到被引高峰，且較早開始衰退，高被引文章則是被引狀況較持久，甚至有部分文章愈來愈常被引用。

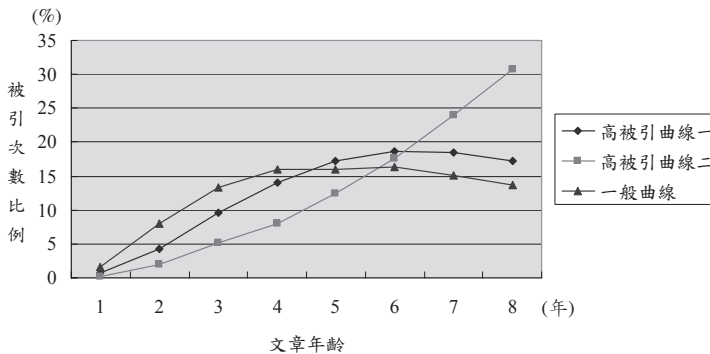


圖3 高被引文章與一般文章老化曲線

比較電腦科學領域高被引文章與一般文章老化曲線，在各年齡之被引次數比率，表8顯示年齡1至年齡4一般文章老化曲線之被引次數比率，均高於高被引文章老化曲線，且一般文章年齡1至4之增加幅度達14.36%，多於高被引文章老化曲線一之增加幅度13.41%，亦多於高被引文章老化曲線二之7.80%；此外高被引文章老化曲線一從年齡5至年齡8皆維持於17%至19%，高被引文章老化曲線二之被引次數比率則是持續增加，年齡7和年齡8被引次數比率大於20%，亦大於高被引文章老化曲線一及一般文章之被引次數比率，一般文章之被引次數比率則是維持16%左右一段時間後，在年齡8降至13.76%。

表8 高被引文章與一般文章老化曲線之被引次數比率

文章年齡	高被引文章		一般文章
	老化曲線一 被引次數比率	老化曲線二 被引次數比率	老化曲線 被引次數比率
1年	0.66%	0.22%	1.60%
2年	4.20%	1.88%	7.92%
3年	9.56%	5.17%	13.40%
4年	14.07%	8.02%	15.96%
5年	17.29%	12.39%	15.92%
6年	18.58%	17.60%	16.29%
7年	18.49%	24.05%	15.14%
8年	17.16%	30.68%	13.76%
總計	100.00%	100.00%	100.00%

表9 高被引文章與一般文章文獻半衰期分布

半衰期	高被引文章		一般文章	
	文章數	百分比	文章數	百分比
1年	0	0.00%	19	3.52%
2年	0	0.00%	38	7.04%
3年	3	0.68%	85	15.74%
4年	25	5.63%	81	15.00%
5年	131	29.50%	124	22.96%
6年	182	40.99%	95	17.59%
7年	84	18.92%	61	11.30%
8年以上	19	4.28%	37	6.85%
總計	444	100.00%	540	100.00%
平均半衰期	6.30年		5.17年	

註：被引次數為0之文章未列入計算。

(二)文獻半衰期

比較電腦科學領域高被引文章與一般文章之半衰期，表9顯示高被引文章平均半衰期為6.30年，一般文章平均半衰期為5.17年，在半衰期分布上，高被引文章半衰期最短為3年，一般文章半衰期最短則是1年；高被引文章和一般文章半衰期均以5和6年比率最高，有七成高被引文章之半衰期在此範圍，及四成一般文章之半衰期在此範圍，另外一般文章半衰期在4年以內之比率達四成，高被引文章則低於一成，t檢定顯示高被引

文章與一般文章的半衰期達顯著差異 ($t=12.607, p=.000<.05$)，可見高被引文章半衰期普遍較長，老化較慢。

四、臺灣地區機構所屬作者發表之高被引文章與一般文章特性分析

分析電腦科學領域臺灣地區機構所屬作者發表之高被引文章及一般文章特性，並與電腦科學領域所有高被引文章和所有一般文章進行比較，分析項目同樣包含基本特性、合著特性、及老化特性。

(一)基本特性分析

基本特性分析包括被引次數分布、首次被引時間、引用文獻數、及期刊影響係數分析。本文以1996至2006年出版之20篇臺灣高被引文章和89篇臺灣一般文章，分析文章基本特性。

在被引次數分布方面，臺灣高被引文章平均被引次數為30.85次，被引次數分布與所有高被引文章相似，均是較早出版的文章被引次數較高， t 檢定顯示臺灣高被引文章與所有高被引文章的被引次數未達顯著差異 ($t=.992, p=.390>.05$)。臺灣一般文章平均被引次數為2.25次，低於所有一般文章之3.48次，從各年平均被引次數來看，臺灣一般文章皆低於所有一般文章， t 檢定發現臺灣一般文章與所有一般文章的被引次數達顯著差異 ($t=3.270, p=.002<.05$)，顯示臺灣一般文章的被引次數低於所有一般文章。

在首次被引時間方面，臺灣高被引文章平均首次被引年齡為1.60，10%的臺灣高被引文章首次被引年齡為3，90%首次被引時間集中於年齡1或2，分布與所有高被引文章相似， t 檢定顯示臺灣高被引文章與所有高被引文章的首次被引時間未達顯著差異 ($t=.607, p=.595>.05$)。臺灣一般文章平均首次被引年齡為3.61，與所有一般文章之首次被引時間皆有七成集中於年齡2至4， t 檢定發現臺灣一般文章與所有一般文章的首次被引時間未達顯著差異 ($t=-1.493, p=.698>.05$)，首次被引時間分布相似。

在引用文獻數方面，臺灣高被引文章平均引用文獻數為19篇，主要集中於30篇以下，佔85%，所有高被引文章的引用文獻數在30篇以下者，亦佔55.44%， t 檢定顯示臺灣高被引文章與所有高被引文章的引用文獻數未達顯著差異 ($t=2.056, p=.129>.05$)。臺灣一般文章平均引用文獻數為22.35篇，約有七成集中在30篇以下，所有一般文章的引用文獻數同樣有七成集中於此範圍， t 檢定顯示臺灣一般文章與所有一般文章的引用文獻數未達顯著差異 ($t=-.626, p=.077>.05$)，引用文獻數分布相似。

在期刊影響係數方面，臺灣高被引文章之來源期刊共有7種，其中4種為非電腦科學領域期刊，其餘3種電腦科學領域期刊的臺灣高被引文章，平均期刊影響係數排名百分比為6.48%，t檢定顯示電腦科學領域中臺灣高被引文章與所有高被引文章的期刊影響係數排名百分比未達顯著差異（ $t=1.019, p=.226>.05$ ）。臺灣一般文章之平均影響係數排名百分比為59.30%，有四成文章的影響係數排名百分比大於70%，由於所有一般文章是參考JCR電腦科學領域期刊進行隨機抽樣，所以其期刊影響係數排名百分比分布較均勻，平均百分比為49.06%。相較之下電腦科學領域中臺灣高被引文章的期刊影響係數排名百分比較所有一般文章為大，t檢定亦顯示二者之期刊影響係數排名百分比達顯著差異（ $t=-3.756, p=.010<.05$ ）。」

(二)合著特性分析

合著特性分析內容包括合著率、平均作者數、跨國合著率、與平均國家數。本文以1996至2006年出版之20篇臺灣高被引文章和89篇臺灣一般文章分析文章合著特性。

在合著率方面，臺灣高被引文章合著率為75%，無論是臺灣高被引文章或所有高被引文章，合著文章數皆多於單一作者文章數，卡方檢定顯示臺灣高被引文章與所有高被引文章的合著或單一作者文章數未達顯著差異（ $\chi^2=3.777, p=.052>.05$ ）。臺灣一般文章合著率為83.15%，和所有高被引文章相同，合著文章數皆多於單一作者文章數，卡方檢定顯示臺灣一般文章與所有一般文章的合著或單一作者文章數未達顯著差異（ $\chi^2=1.888, p=.169>.05$ ）。

在平均作者數方面，臺灣高被引文章平均作者數為2.60人，臺灣高被引文章與所有高被引文章的作者數，皆有近六成為2或3人，t檢定顯示臺灣高被引文章與所有高被引文章的作者數未達顯著差異（ $t=1.342, p=.339>.05$ ），作者數分布相似。臺灣一般文章之平均作者數為2.71人，臺灣一般文章與所有一般文章皆以2或3位作者比例較高，t檢定顯示臺灣一般文章與所有一般文章的作者數未達顯著差異（ $t=-1.146, p=.150>.05$ ），作者數分布相似。

在跨國合著率方面，臺灣高被引文章之跨國合著文章數低於單一國家文章數，跨國合著率為10%，所有高被引文章之跨國合著文章數同樣低於單一國家文章數，卡方檢定顯示臺灣高被引文章與所有高被引文章的跨國合著或單一國家文章數未達顯著差異（ $\chi^2=2.012, p=.156>.05$ ）。臺灣一般文章跨國合著率為17.97%，臺灣高被引文章與所有一般文章之跨國合著文章數，皆低於單一國家文章數，卡方檢定顯示臺灣一般文章

與所有一般文章的跨國合著或單一國家文章數未達顯著差異 ($\chi^2=.103, p=.748>.05$)。

在平均國家數方面，臺灣高被引文章為1.10個國家，低於所有高被引文章之1.29個國家，臺灣高被引文章最多由2個國家合著，所有高被引文章最多則由8個國家合著，t檢定顯示臺灣高被引文章與所有高被引文章的國家數達顯著差異 ($t=2.676, p=.004<.05$)，臺灣高被引文章之國家數低於所有高被引文章。臺灣一般文章之平均國家數為1.19個國家，臺灣一般文章與所有一般文章皆以2個國家合著比例較高，t檢定顯示臺灣一般文章與所有一般文章的國家數未達顯著差異 ($t=-.058, p=.969>.05$)，國家數分布相似。

(三) 老化特性分析

老化特性分析內容包括老化曲線以及文獻半衰期。本文以1996至1999年出版之5篇臺灣高被引文章和29篇臺灣一般文章，分析文章老化特性。

在老化曲線方面，本文對臺灣電腦科學領域高被引文章進行被引次數比率K-means集群分析，由於文章數僅有5篇，篇數過低，且若將被引次數比率觀察值之分群設為兩群以上，各群的群中心點差異不大，所以分群數以一群較佳。圖4為臺灣電腦科學領域高被引文章與所有高被引文章之老化曲線。臺灣高被引文章之老化曲線前期平緩，後期快速上升而後略降，曲線略接近所有高被引文章老化模式一，但波動程度較大；此外臺灣高被引文章之老化曲線，在年齡4之後被引次數比率快速上升，年齡6和年齡7之被引次數比率超越所有高被引文章，顯示臺灣高被引文章發表後的前幾年，被引狀況偏低，相較於所有高被引文章前期持續增加的被引狀況，臺灣高被引文章需要經過一段時間才會得到較多引用。

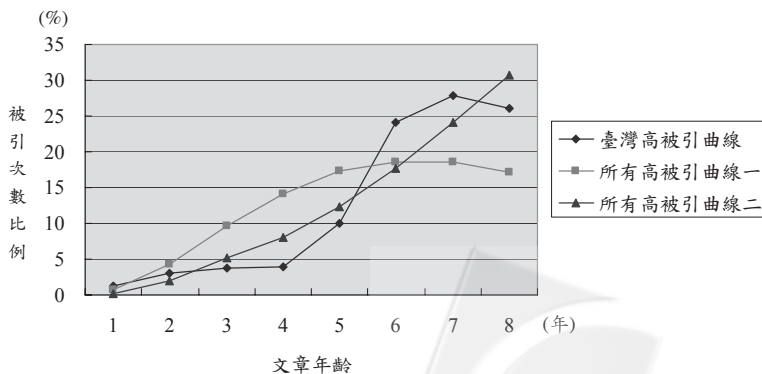


圖4 臺灣高被引文章與所有高被引文章老化曲線

透過K-means分析臺灣電腦科學領域一般文章之被引次數比率集群，被引次數為0的文章因無法計算被引次數比率，所以未列入分析，其餘文章數僅有21篇，篇數不多，因此一群為較佳的分群數。圖5顯示臺灣一般文章與所有一般文章相同，皆只有一種老化曲線。臺灣一般文章呈現前期上升之後下降之曲線，整條曲線波動大，被引次數比率從年齡1快速增加至年齡5才趨緩，所有一般文章的被引次數比率則是在年齡4之後持平；此外臺灣一般文章被引次數比率的下降幅度大於所有一般文章，臺灣一般文章從年齡6之後被引次數比率明顯減少，顯示臺灣一般文章比所有一般文章晚達到被引高峰，且臺灣一般文章被引狀況較不持續，在發表一段時間後被引狀況明顯減少，但由於臺灣一般文章被引次數很低，各年齡平均被引次數皆低於1次，所以雖然曲線看似起伏大，但實際文章被引次數其實少有劇烈升降變化。

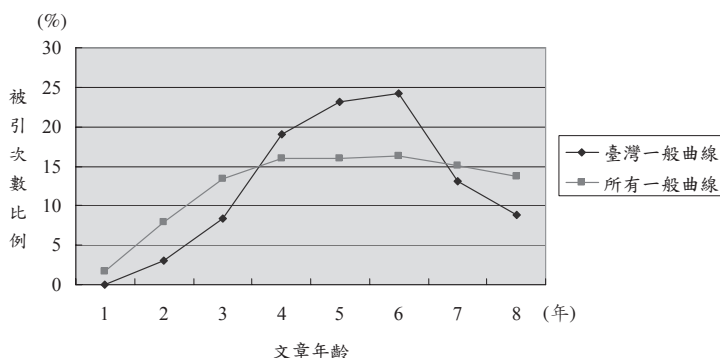


圖5 臺灣一般文章與所有一般文章老化曲線

在文獻半衰期方面，臺灣高被引文章平均半衰期為6.88年，臺灣高被引文章和所有高被引文章皆有四成比例其半衰期為6年，由t檢定可知臺灣高被引文章與所有高被引文章的文獻半衰期未達顯著差異 ($t=-1.423, p=.965>.05$)，半衰期分布相似。臺灣一般文章平均半衰期為5.17年，和所有一般文章平均半衰期相同，由t檢定結果可知，臺灣一般文章與所有一般文章的文獻半衰期未達顯著差異 ($t=.011, p=.865>.05$)，半衰期分布相似。

伍、結論

本研究從ESI資料庫蒐集近十年2,112篇電腦科學領域高被引文章，由WOS資料庫檢索近十年2,200篇電腦科學領域一般文章，以書目計量法

分析比較文章基本特性、合著特性、及老化特性，並由高被引文章與一般文章集中檢索作者所屬機構所在地區為臺灣的文章，再次進行特性分析。

研究結果顯示，高被引文章各年平均被引次數差距較一般文章大，首次被引時間亦早於一般文章，絕大多數都是在發表當年或隔年即被引，超過三年未被引用之文章幾乎不可能成為高被引文章；一般文章的首次被引時間則相對分散，可能發表多年後才首度被引，甚至從未被引，而且，因為還未被引之一般文章未來仍有機會被引，所以一般文章的實際首次被引年齡可能比本研究結果更大。此外，由於高被引文章部分為回顧性文章，這些文章的引用文獻數通常較多，所以高被引文章的引用文獻數部分特別高，一般文章的引用文獻數則較集中偏低。在期刊影響係數方面，高被引文章來源期刊之影響係數排名偏高，大多數發表於該領域高影響係數期刊，一般文章則受制於抽樣方法，排名分布均勻。分析各國發表的文章數，可發現高被引文章與一般文章之主要發表國家均為美國，表示美國發表了許多電腦科學領域文章，其中不少是極具價值之高被引文章，且歐美國家之高被引與一般文章相對比值高，其研究品質可能較佳。

此外，從合著率來看，高被引文章與一般文章近十年之合著率皆呈現上升趨勢，尤以高被引文章合著率增加幅度較大，且各年合著率皆較高，所以高被引文章平均合著率高於一般文章；高被引文章的跨國合著率同樣高於一般文章，然而高被引文章與一般文章近十年之跨國合著率沒有明顯上升或下降趨勢，顯示合著現象雖日益普遍，但跨國合著仍存在許多阻礙因素。至於文章平均作者數及平均國家數，高被引文章皆高於一般文章。

透過被引次數比例之變化分析老化曲線，高被引文章可概分出兩種老化曲線，分別是快速上升而後持平以及持續上升；一般文章則呈現前期上升而後略降之曲線。由此可知高被引文章被引狀況較一般文章持久，部分文章甚至愈來愈常被引用，被引次數比率還未下降，若延長觀察期將能更完整呈現高被引文章老化曲線變化；相對地，一般文章則較早達到被引高峰且較早衰退。從半衰期亦能印證，高被引文章半衰期較長，老化狀況較慢。

進一步探討臺灣地區機構所屬作者發表之文章，結果得知臺灣高被引文章之特性有許多與所有高被引文章相近，臺灣一般文章之特性亦有許多與所有一般文章類似。然而臺灣一般文章之平均被引次數低於所有一般文章，且臺灣的高被引文章數排名不如一般文章數排名，高被引

文章與一般文章相對比值偏低，顯示雖有許多研究者投入且發表多篇文章，但研究品質還需持續提升。另外臺灣高被引文章之老化曲線呈現前期平緩，後期快速上升而後略降，臺灣一般文章之曲線則呈現前期上升之後下降，相較於所有高被引文章與一般文章，臺灣文章被引時間較晚，且被引狀況較不持續。

綜合以上所述，高被引文章在基本特性、合著特性，及老化特性方面，都與一般文章有顯著不同，未來或可由其他角度切入，透過更多面向瞭解高被引文章與一般文章特性差異，甚至針對其他領域再作探討，與本研究結果進行比較分析，將可對高被引文章有更全面之瞭解。

參考文獻

- 王淑儀 (1998)。電腦科學文獻老化研究——兼論同時法與歷時法之特質。未出版之碩士論文，淡江大學教育資料科學學系，臺北縣。
- 孟連生 (1996)。試論引文索引法的性質與功能。資訊傳播與圖書館學，3(1)，15-19。
- 張偉嶠 (2006)。2000年—2004年臺灣學者的科研成績單。上網日期：2007年5月27日，檢自：http://www.sciscape.org/news_detail.php?news_id=2010
- 黃裕惠 (2004)。分子生物與遺傳學之文獻計量分析。未出版之碩士論文，國立臺灣大學圖書資訊學研究所，臺北市。
- 黃慕萱 (1994)。引用文獻初探。在王振鵠教授七秩榮慶論文集編輯小組編，當代圖書館事業論集——慶祝王振鵠教授七秩榮慶論文集 (頁807-816)。臺北市：正中書局。
- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170.
- Arunachalam, S., & Singh, U. N. (1984). Publication & citation patterns in the literature of a high metabolism area: The case of superconductivity in 1970. *Journal of Information Science*, 8, 93-102.
- Aversa, E. S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics*, 7(3-6), 383-389.
- Avkiran, N. K. (1997). Scientific collaboration in finance does not lead to better quality research. *Scientometrics*, 39(2), 173-184.

- Bridgstock, M. (1990). The quality of single and multiple authored papers: An unresolved problem. *Scientometrics*, 21(1), 37-48.
- Cano, V., & Lind, N. C. (1991). Citation life-cycles of 10 citation-classics. *Scientometrics*, 22(2), 297-312.
- European Commission. (2001). *Indicators for benchmarking of national research policies*. Retrieved May 30, 2004, from <http://europa.eu.int/comm/research/era/pdf/benchmarking2001.pdf>
- Glänzel, W. (2002). Coauthorship patterns and trends in the sciences (1980-1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50(3), 461-473.
- Hart, R. L. (2007). Collaboration and article quality in the literature of academic librarianship. *Journal of Academic Librarianship*, 33(2), 190-195.
- Katz, J. S., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40(3), 541-554.
- McCain, K. W., & Turner, K. (1989). Citation context analysis and aging patterns of journal articles in molecular genetics. *Scientometrics*, 17(1-2), 127-163.
- Narin, F., & Hamilton, K. S. (1996). Bibliometric performance measures. *Scientometrics*, 36(3), 293-310.
- Plomp, R. (1990). The significance of the number of highly cited papers as an indicator of scientific prolificacy. *Scientometrics*, 19(3-4), 185-197.
- Rousseau, R. (2001). *Are multi-authored articles cited more than single-authored ones? Are collaborations with authors from other countries more cited than collaborations within the country? A case study*. Retrieved July 5, 2007, from <http://doelib.uhasselt.be/dspace/handle/1942/822>
- Smart, J. C., & Bayer, A. E. (1986). Author collaboration and impact: A note on citation rates of single and multiple authored articles. *Scientometrics*, 10(5-6), 297-305.
- Stinson, E. R. (1981). *Diachronous vs. synchronous study of obsolescence*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Wade, N. (1975). Citation analysis: A new tool for science administrators. *Science*, 188(4187), 429-432.

Zhu, X., Wu, Q., Zheng, Y., & Ma, X. (2004). Highly cited research papers and the evaluation of a research university: A case study Peking University 1974-2003. *Scientometrics*, 60, 237-247.

