

世界大學科研論文質量評比的運用與限制

文／黃慕萱

國立臺灣大學圖書資訊學系教授
高等教育評鑑中心兼任研究員

為了解國內大學的論文競爭力，財團法人高等教育評鑑中心基金會委託筆者費時一年，完成「世界大學科研論文質量評比」計畫，以「學術生產力」、「學術影響力」與「學術卓越性」三大評估構面，進行2007年世界大學科研論文質量整體評比及排名，提供世界前500大名單。

評鑑中心自去（2007）年11月8日發布世界大學科研論文質量評比以來，收到各界關注與詢問，為求更清楚呈現本評比的運用與限制，茲將常見問題整理如下：

1. 本次的評比特色？

此次進行論文質量評比的特色為：高達80%為品質指標（首度應用「h指數」反映學校論文品質），且長短期指標各占50%，彰顯短期努力績效。但此次評比僅是「科研論文質量評比」，不是涵蓋教學、研究及行政之「大學評比」，亦不是著重聲譽或非凡成就的「學術表現評

比」。所設計指標適合作為新興國家研究型大學的參考。

2. 本次評比的適用對象為何？

此次科研論文質量評比的適用對象為研究型大學，特別是針對新興國家的學校，透過客觀且能反映出短期努力的指標，讓這些大學能了解其在世界排行的地位及優勢，得知自己與其他大學的差距，並能追蹤學校論文質量每年的進步情形。

3. 本次評比可否代表大學排名？

不可，此次科研論文質量評比僅針對世界各大學所發表之科研論文，分別就質量與數量、長期及短期角度設計指標，僅能呈現出各大學在科學與社會科學論文上的質量，無法延伸解釋至教學、研究及行政等績效綜合評比。

4. 本次評比可否代表大學的整體學術能力？

不完全可以，因為除了論文外，尚有研

究計畫、專利、專書、產學合作等，都是大學整體學術能力的一部分。不過科研論文的質量除人文領域外，經常被用來推估學校的學術能力。

5. 本次評比可否被視為五年五百億的考核標準？

不可，此次科研論文質量評比並非作為五年五百億計畫的考核或評估，僅可視為五年五百億計畫實施之前論文質量的基礎點，因為至2006年底，大多數利用五年五百億經費之研究成果皆尚未產出。

6. 本次評比與上海交大「世界大學學術排名」有何不同？

上海交通大學的「世界大學學術排名」採用的各項指標中，包括諾貝爾獎、高被引學者與Nature、Science的文章等指標，追求的都是非常傑出的研究成果，尤其是諾貝爾獎，並非大多數學校可以達到的目標，因此無法反應大多數學校的研究表現。而「科研論文質量評比」乃設計一套包含多數學校現階段經努力即可達成的短期指標，只要學校一有進步，成果隔年即可反映在排名上。



▲約翰霍普金斯大學。

7. 本次「科研論文質量評比」與英國《泰晤士報高等教育增刊》的世界大學排名有何不同？

英國《泰晤士報高等教育增刊》(The Times Higher Education Supplement)所進行之世界大學排名，其強調同儕評量、企業雇主評量，權重達50%，易受主觀意識影響，尤其是其問卷形式的大學排名給分，很容易成為大學聲望評比，而忽略大學的實際表現，而「科研論文質量評比」則以較為客觀的數據進行大學排名。

8. 本次「科研論文質量評比」是否有考慮學校規模因素？

是，一所學校論文的多寡，自然和該校規模有一定關係，因此本評比利用平均被引次數等指標平衡學校規模的影響。學校規模在某種程度上亦可代表學校實力，因此本次評比僅能盡量尋找一個平衡點，只是在最終排名中，仍然可看出規模的影響，明年的評比將繼續就減輕規模影響作努力。

9. 為何「科研論文質量評比」不採計人文藝術學科的論文？

由於藝術人文領域之研究範疇地域性色彩較濃、研究成果常以非英語或專書發表，而A&HCI收錄文獻以英語期刊論文為主，故本評比不採計A&HCI期刊論文。

10. 為何「科研論文質量評比」的指標之一h指數僅統計兩年？

由於h指數是相當敏感的指數，二年論文數已相當大量（以哈佛大學為例，其二年論文數為28,951篇），足以產生客觀的質量綜合指數。我們試取11年論文數

最高的哈佛大學與東京大學及分散於中間及後面的學校共47間，進行長期與短期的h指數相關係數分析，發現二年h指數與11年h指數的相關係數為0.967，具有高度相關性，實無必要進行11年h指數的計算。

11. 為何有些學校感覺排名應該要比某些學校好但卻排名在後面？

由於本評比進行的是「科研論文質量評比」，僅能忠實反應各大學當時之論文質量表現，並非「學術評比」或「大學評比」，故可能與其他大學排名有所不同。不同的評比指標、學校規模大小、有無醫學院，甚至人文社會科學占該大學比例等

問題，皆會影響評比結果。

12. 分數接近的學校是不是代表其科研論文表現差不多？

是，由於各指標第一名學校相較其他學校數值特別高，故後面學校總分差距相當些微，加以本評比所給予的名次，只是方便讀者得知大致排名，並無意圖顯示任何絕對差距，意即排名接近且分數差距細微的學校，不宜單以名次斷定表現優劣，而應將其視為科研論文表現不分軒輊的學校。

13. 為何不將二百名後學校以201-250、251-300等區間名次標示分數接近學校的排名？

若將分數接近的學校以201-250、251-300

相同區間名次顯示，有難以訂定區間範圍的困難，且極有可能此區間的最後一名與下一區間的第一名差距不大，但兩者卻被分成二級。由於同區間名次內學校數該設定多少難以界定，且細微的差距不宜被區分為二種等級，故本評比並未採區間名次標示。

14. 「科研論文質量評比」採用的指標是否太偏量化數據而不注重質化表現？

本評比以可客觀取得的數據為依據，評估各大學之科研論文質量表現，一般在探討科研論文績效表現時，量化的分析指標較能提供客觀的描述基準，尤其量化指標較適合進行排名，故本評比採量化指標為

主，但在量化指標上加入觀測論文品質之概念，共有80%的品質指標，可謂是以質化為主的量化數據評比。

15. 以人文社會科學為主要領域的學校排名是否會較為落後或無法進入排名？

是，本評比的對象是科研論文，資料來源為SCI資料庫與SSCI資料庫，並不包含人文科學的論文（A&HCI）。同時，雖包含社會科學論文，但由於社會科學的論文數較少（SCI收錄6,771種期刊，SSCI收錄2,005種），因此以人文社會科學為主的學校，如果該校沒有其他相對較強的科學領域，就很難在本評比排名結果中出現。

16. 以醫學院為主的學校，排名是否會

較為前面？

是，由於領域特性的關係，醫學領域文章發表量與被引次數相較其他領域高出許多（1997年至2007年8月的臨床醫學（Clinical Medicine）領域論文數共有1,868,486篇，被引次數20,643,979次；工程（Engineering）領域論文數為723,371篇，被引次數為2,542,146次）；故於本評比中，有醫學院或附設醫院的學校其表現會比沒有相關系所的學校更加優異，本評比未來考慮將增加以分領域排名之方法進行年度排名。

17. 規模較大的學校是否排名會比較前面？

是，本評比今年雖使用平均被引次數等指標平衡學校規模的影響，然在最終排名中，仍可看出規模的影響，例如在西雅圖的華盛頓大學的學生數大約為40,000人左右（教師數約5,700人）且有醫學院，故排名較前面。今年已試圖減輕規模效果但仍不足夠，明年進行評比時會持續將規模影響程度降低。

18. 如何看待各國進榜學校數量與名次與心中預期不同？

本評比並非「大學聲望評比」或「大學學術評比」，故可能與不同對象心中認知的大學排名有所不同。除外，在該國家進入排名榜的學校中，也許實際名次與心中預期並不相同，但同國家入榜學校間的相對位置仍符合社會預期。

19. 為什麼評比指標這麼多，是不是可以整併？

本次評比與上海交大評比總分的相關係數為.896，名次的相關係數為.853。

同時，九個指標與上海交大總分的相關檢定皆達顯著，亦通過統計迴歸的顯著性檢定。換言之，短期指標的考量造成評比指標數目的增加，但這可使學校的努力能在次年看到成果。

20. 為何第一名與第二名學校之間總分差距非常大？

第一名與第二名差距過大是學術評比普遍存在的問題，上海交大評比第一名與第二名之總分差距約30分之多。泰晤士報評比為縮小分數差距，今年分數即依Z分數進行調整，其總分差距的縮小實來自人為調整。本評比曾嘗試以T分數調整並縮小第一名與第二名學校的差距，但為忠實呈現原始差距，還是決定未予調整。

由於排名榜中第一名哈佛大學規模較大（學生數約25,000人，教師數約2,500人），有醫學院且醫學院表現極佳（1996年至2006年之11年臨床醫學論文數即有35,492篇，被引次數為904,586次），加上該校其他領域論文質與量的表現皆非常優異，九個指標中，除了平均被引次數與卓越領域數二個指標外，皆為最高分，與其他學校相較之下，表現顯得特別傑出；第二名Johns Hopkins University的規模不算小（學生數約20,000人，教師數約3,100人），亦有醫學院且擁有不少醫學領域文章（1996年至2006年之11年臨床醫學論文數有21,205篇，被引次數為487,568次），然其長短期文章數與被引次數皆明顯落後哈佛大學；外加其於高被引文章數和高影響期刊論文數表現皆與哈佛大學有相當落差，故第一名與第二名學校之總分差距甚大。

