

破除 SCI 至上:終點又回到了起點?

朱 劍

2020 年 2 月 18 日,教育部、科技部聯合印發《關於規範高等學校 SCI 論文相關指標使用 樹立正確評價導向的若干意見》(以下簡稱“《意見》”),與以往頂層管理部門很少就某個具體評價指標發表意見不同,僅從標題即可看到已上升為“評價導向”高度的鮮明態度。如此特別的《意見》公開發佈,不出意外地引起了學界和學術期刊界關於學術評價問題新一輪的熱議。反應迅捷的網絡上最先出現的當然是叫好聲,但擔憂的聲音很快也就出現了。緊接著,教育部科技司負責人以答記者問的形式對文件精神作出了進一步解釋,但網上的議論並沒有就此平息,相反,憂慮甚至質疑的聲音仍不斷響起。無論叫好抑或擔憂,SCI 指標淡出學術評價必將成為事實,那麼,我們該如何看待評價指標的這一變化?學術評價又將迎來怎樣的未來?筆者以為,只有追根尋源,從 SCI 是什麼,緣何被應用於學術評價之中、應用給評價帶來了哪些變化一路回溯,弄清楚它的來龍去脈,才能對剔除 SCI 後的替代品能否帶來更合理的評價、等待我們的評價未來可能是什麼有個大致判斷。

一、SCI 指標是如何越界成為學術評價標準的?

《意見》的確是從頭說起的,關於 SCI 是什麼,《意見》說得很清楚:“是國內外廣泛使用的科技文獻索引系統”,但一個“文獻索引系統”的指標為何以及如何能夠進入今天的學術評價並產生了決定性作用,《意見》並沒有分析。其實,這是一個不能忽略的問題,因為一定是有特別緣由的。

這就要追溯到 SCI 指標的產生。1950 年代,美國文獻學家尤金·加菲爾德發現了科學文獻引文分佈的規律,即多數引文集中在少數刊物上,由此開創了文獻檢索的新途徑——引文檢索。他創立了“科學情報研究所”(ISI),設計並開發了 SCI 這個以集中了主要引文的少數科學期刊作為數據源的引文數據庫(最初是紙質出版物)。這在學術期刊大量湧現的時代的確是個創舉,有了它,研究者蒐集相關信息就能既省時又方便。為了及時準確地追蹤學科進展,必須為來源期刊的調整(遴選和剔除)設立指標,這就是 SCI 指標的由來。這些指標,只與引文相關,並不涉及內容,更不對引用做價值評判,稱“體系”似不無誇張,不過從其獨特性和封閉性來說也未嘗不可,因為正是在“體系”的意義上,這套指標的含義和用途都是特定的,即專門用於引文數據庫來源期刊的調整。

那麼,SCI 指標是如何越界而被引入學術評價中來的呢?雖說這套指標設立之初目的就非常明確,即根據引文找出影響力相對大的期刊,但從宏觀角度看,影響力與學術質量往往是具有正相關關係的(單一的引文數據能否準確描述影響力另當別論)。所以,在 SCI 指標設立之初,就因其所“天然”具有的評價屬性而埋下了未來越界的伏筆。一方面,只要依據這套指標製作的數據庫價值得到公認乃至被視為權威,那麼,這套指標體系關於期刊“影響力”的描述也一定會得到廣泛認同;另一方面,這套指標體系還有一個“天然優勢”,就是既“客觀”,又“量化”,一目了然,童叟無欺,頗有說服力。因此,只要出現合適的機會,它的越界就是件大概率的事。

近半個世紀以來，國際間國力的比拼很大程度上就是科研實力的較量。科研是需要資源保障的，一個大學或科研機構，想獲取資源就得證明自己有實力。何以證明？唯有比較，有比較才能有鑑別。我們今天把這樣的比較和鑑別稱為“學術評價”，這裡的“評價”與學術史上的“批評”已遠不是一回事，而與分配資源的權力直接相關。按說這樣的評價理應基於學術質量，但從技術角度來看，若以質量評價為目的，以大學和學科為對象，需要處理的樣本是天量的，更不用說評價者的德行和學識了，故很難實操。即使對學者個人學術水平的評價，也會因類似的原因而困難重重。相比之下，如果根據權威的定量數據來排名，則既省事又有說服力，這就為 SCI 指標越界提供了機會。

隨著 SCI 的價值得到越來越多學者的認可，其權威性逐漸確立，它畢竟搜羅了各學科最優秀的期刊。正因為有評價的客觀需要，也因為 ISI 亦需營利，在數據庫問世後，其評價作用也被開發了出來。這些所謂的評價並沒有重新設立指標體系，而是直接利用遴選來源期刊的某些指標，再加上基於某些算法的一些衍生指標，對學者、期刊、機構、作品等進行排名，製作成各類排行榜，成為既與數據庫捆綁又可獨立於數據庫而存在的所謂“評價產品”。因為指標體系不變，這些產品仍然只關引文，不涉內容，原本複雜的學術評價由此也成了十分簡單的數字遊戲。

作為 SCI 的副產品，這些排行榜在社會上引起了很大反響，潛移默化地影響乃至塑造了從普通民衆到政府部門對頂尖到一般大學、學科、期刊和學者及其作品的基本看法，同時也為 SCI 拓展了廣闊的市場。由此，SCI 終於從一個“文獻索引系統”步入了排行榜領域，而 SCI 在數據庫和評價兩個領域的成功，也使得一些仿製的數據庫（如“社會科學引文索引”〔SSCI〕、“藝術和歷史引文索引”〔A&HCI〕等）和排行榜相繼問世。

在這種情況下，影響力排名逐漸成了最重要的評價。客觀地說，由加氏定律生發的 SCI 指標應用於長時段大樣本、結果簡單粗疏的影響力排序時，只要數據沒有受過人為污染，加之對來源期刊分區和附加的糾偏算法，的確不失為一種低成本且容易產生說服力的辦法。因為不管怎樣，我們都得承認 SCI 網羅的各學科優秀期刊的總體質量明顯優於非 SCI 期刊。某個機構、團隊或學者能在 SCI 期刊，尤其是其中排名靠前的期刊持續發表較多數量的論文，且被同行大量引用，的確可視為科研實力的一種間接證明。但是，這樣的排行榜一旦作為影響力評價的結果在社會廣泛流傳並形成對公衆和政府的影響後，影響力與學術質量之間的界限就變得模糊不清了。

SCI 指標的越界並沒有到此為止，隨著排行榜對高校和科研機構的影響逐漸顯現，不用說一般高校，就連原本並不怎麼在意這些排行榜的著名高校和科研機構也不得不重視起來，特別是這樣的排名足以影響到學校的聲譽、師生來源及資金來源之後。當為獲取更好排名而不得不迎合這些指標成為一些學校的選擇時，排行榜的“指揮棒”效應就開始顯現了。SCI 指標很快進入了高校的管理領域，向傳統的同業評議發起了挑戰。至此，以影響因子為代表的 SCI 指標終於完成了從選擇數據庫來源期刊的封閉性標準到學術評價中普遍採行的普適性標準的跨越。

二、SCI 指標何以引入中國並替代同行評議的？

SCI 為中國學界廣泛知曉大概是 1980 年代後期的事。從 1987 年起，國家科委即委託中國科學技術情報研究所利用 SCI 等檢索工具，對我國學者在國際學術期刊發表論文的情況進行統計。接著，各高校和科研機構又紛紛以該所的這份統計資料作為依據進行比較排行。可見，與 SCI 的檢索功能相比，其指標體系似乎影響更大。自然科學界以外特別是管理部門對 SCI 的最初接觸和理解，不是數據庫的檢索服務功能，而是學術評價功能。這與當時特定的時代背景是有直接關係的。

首先,國家對科研的投入在 1990 年代步入了快速增長期。這些投入大多是以項目和工程的形式分配的,無論項目還是工程,都是集中資源辦大事的舉國體制的體現,向誰集中,就是讓誰得到機遇,得到機遇就意味著可能實現跨越式發展,甚至遠超同儕,一騎絕塵。這就需要評價。掌控資源分配和管理的行政權力部門對評價的要求就是權威性和簡明性,只有這樣才具有可操作性。但要同時滿足這兩個要求是極其困難的,這時,目的就決定了手段,能否最有效地促進優秀成果產出,就成為是否採信某種評價的決定性標準。評價機構的圖情專家及時送上了貌似符合這兩個要求的 SCI 評價:“SCI 對某項研究成果的評價……直接的方法就是通過系統地檢索 SCI,從中查出某項成果的實際引用次數;間接的方式即是看發表成果的雜誌處在什麼水平上。顯然兩種方式都具有很強的可靠性。相對而言直接評價方式比較繁瑣……而間接評價方式就要簡單快捷得多,一經發表便能確認”。“以刊評文”就是這麼來的,從中不難看到建議者對管理者的瞭解和體貼。

其次,時至 1990 年代,學術界長期以來積極提倡的“與國際接軌”逐漸被“國際化”所取代。較之“接軌”,“國際化”的意涵要豐富得多,最大的不同在於前者是單向的,後者是雙向的,不難看出,後者對前者的取代表達了一種強烈的輸出願望,要使中國的知識產品走向世界,讓世界聽到中國的聲音。這與國家對科研的巨額投入又是交織在一起的,巨量投入是為了迅速提高中國的科研能力和水平,而參照系就是當今世界科技強國,要能比肩世界科技強國,國際化就是必由之路。而在評價機構的圖情專家看來,最快捷的學術國際化途徑就是在國際學術界公認的 SCI 期刊上發表論文。

再次,高校教師和科研人員對公正評價的渴望。今天的許多學者在批評 SCI 指標或類似計量評價時,往往都是以“受害者”的身份發言的,但有件事大家可能都遺忘或忽略了,就是 1990 年代學界對量化評價的呼喚。在 1990 年代之前,評價往往是各學科專家的同行評議,但隨著評價與利益的關係日漸緊密,同行評議的公正性遭遇了前所未有的挑戰,特別是職稱評審中的論資排輩和人情氾濫現象遭到了有才華的青年學者的痛恨,加之當時住房和福利制度尚未改革,教師不僅工資收入,更重要的是住房等福利分配都跟職稱掛鉤。因此,學術界對公正和客觀評價的呼聲開始高漲,認為解決評價問題的最好辦法就是引入客觀的量化指標,把量化指標視為公平的象徵。

最後,SCI 指標的越界“潛力”和評價機構、圖情專家的大力推薦。SCI 指標之所以容易獲得信任,除了前述原因,還有兩個優勢:其一,數據單一,算法簡單,哪個專業的人都能看得明白;其二,所有的數據都可以查證,相互之間也可以比較。有了這些,公正似乎就有了保證。最先把 SCI 評價功能介紹給國內的是一些圖情研究人員,他們不僅將 SCI 描述為通往國際化的快捷甚至唯一路徑,而且成功地用 SCI 指標高度的單一性消解了學科和文化傳統的差異性,從而將 SCI 指標打扮成放之四海而皆準的學術評判標準,終於打動了行政權力部門和學者們。

可見,管理部門對簡單評價的需求、強烈的知識輸出願望、學術界對公正評價的呼喚、評價機構對 SCI 的推崇,共同促成了 SCI 指標運用到了學術管理中。這是一次管理部門、學術界和評價機構的共謀,儘管出發點不盡相同。共謀的結果就是引入 SCI 指標。之所以引入,是因為三方都認為現行評價有重大缺陷。至少在當時看來,SCI 指標有修補甚至替代現行評價的優勢。

正是如此機遇和需求使正在國際上走紅的 SCI 之進入中國水到渠成。1993 年的一篇文章曾描述道:“近年來,國內科研管理界也開始採用引文分析法來開展定量評價……目前我國在評定國家自然科學獎時,把作者發表的論文是否被‘SCI’等國際權威性檢索工具收錄列為重要的必備條件之一。機電部排重點院校時也將‘SCI’收錄各校論文篇數列為條件之一。”這樣的指標一旦登堂入室,成效果然立竿見影。一方面,無論是路徑還是標準,在國際化訴求空前強烈和同行評議顯失

公正的當時都是有積極意義的。中國學者在 SCI 期刊發表論文的數量呈井噴式的成倍激增,世界確實聽到了中國學術界的聲音,看到了中國科研的進步;論資排輩的現象也得到了一定程度的糾正,一些青年學者憑藉在 SCI 期刊發表論文而在國內和國際學界嶄露頭角。但另一方面,卻產生了兩大消極後果:一是由於省略了自建一流學術平台而直接利用 SCI 期刊,不僅使得中國自然科學期刊因稿源外流而成為犧牲品,而且沒有自己平台的國際化必定是跛足的;二是以 SCI 指標取代此前的同行評議,使得十分複雜的學術評價變得極其簡單粗暴,今天“SCI 至上”實始於此時。

新世紀以來,SCI 指標已普遍地運用於高校內部管理,而中國期刊的衰落,更加劇了對 SCI 期刊的依賴,這樣的評價通過科層式的行政管理滲透到了基層的科研實務中。各校為了排名,紛紛在校內資源分配和管理中圍繞 SCI 指標設立了量化評價和考核制度,“以刊評文”成為一種風氣,“SCI 至上”就此形成。誠如《意見》所說:“SCI 論文相關指標已成為學術評價,以及職稱評定、績效考核、人才評價、學科評估、資源配置、學校排名等方面的核心指標,使得高等學校科研工作出現了過度追求 SCI 論文相關指標,甚至以發表 SCI 論文數量、高影響因子論文、高被引論文為根本目標的異化現象,科技創新出現了價值追求扭曲、學風浮誇浮躁和急功近利等問題。”可見,實踐結果與預期有著很大的不同。

三、回歸同行評議:學術評價又回到了原點?

近年來,隨著將 SCI 之類量化指標引入學術評價的負面影響日益彰顯,學界從一開始對量化評價的滿懷希望跌落到了深深失望,對於運用 SCI 之類指標進行的量化評價和學術管理的批評之聲從漸漸響起到不絕於耳,推波助瀾的則是層出不窮的學術不端事件,因為在各種不端行為的背後總能看到量化評價的影子,更為重要的是,捨棄了自建平台的學術國際化缺陷也日益暴露,每年數以幾十萬計的論文外流已被上升到了危及文化安全的高度。此次教育部和科技部發佈的《意見》,針對的正是 SCI 指標的濫用。但如前所述,SCI 指標的引入是有著特殊原因的,在評論和預判《意見》實施後的可能效果之前,我們不妨先看看,當年引入 SCI 指標的原因還在嗎?

首先,較之三十年前,中國的綜合國力更加強盛了。作為世界第二大經濟體,對科研的投入已不是三十年前所能比擬的,但規約科研的舉國體制並沒有變化,亦即主要的科研資源仍然來自國家財政,投入的形式仍然是項目和工程,掌握資源分配權和管理權的仍然是行政權力部門,故而對評價的依賴不僅沒有改變,而且因為所要處理的資源量級的增大,更需要科學的評價來作為分配的依據和公正的保障。因而在可以預見的將來,評價在科研體制中的地位不可撼動。

其次,學術國際化並沒有退潮。三十年前引入 SCI 指標,為的是中國學者能在國際學術舞台發聲,但能發聲不等於擁有話語權,面對 SCI 的規則,除了妥協,別無選擇。要擁有國際學術話語權,就要擁有世界一流期刊集群這樣的平台。反思 SCI 指標的引入,最為虧欠的就是自己的一流平台建設。評價如何做到既不影響學術國際化,又能盡快彌補平台這一短板,是一個亟待破解的難題。

再次,高校教師和科研人員對公正評價的呼聲依然強烈,甚至比三十年前更甚。如今,高校和科研機構對教師和科研人員普遍實行崗位聘任制,雖然不再與住房掛鉤,但不同頭銜、不同職稱、不同崗位收入懸殊,而頭銜和職稱以及日常考核仍然要靠評價來決定。舉國體制之下,用行政手段管理科研的方式不會改變,對評價的依賴性也會一如從前,只不過更需公正合理的評價而已。

最後,只要仍是行政主導的評價,就離不開指標體系,或者說離不開一個可操作的程式性方法。今天高校和科研機構實行的評價指標大多脫胎於 SCI 指標,雖然對提高科研成果產出以及在國際

著名學術期刊發表論文的數量都起到了直接的促進作用,但弊端卻越來越明顯,因此,管理部門和學界從來沒有放棄尋找更合理的評價指標或方法的努力,不過,科學合理的方法卻始終只是理論上的,可操作的一直闕如。沒有替代品才是各類評價機構持續走紅、越來越紅的原因。

從以上幾方面可以看出,當年指望引入 SCI 指標來解決的問題今天不僅仍然存在,而且更尖銳、更急迫、更困難了。其核心問題依然是公正、效率和質量。由於 SCI 指標引入到如今已是負面影響超過了正面意義,實踐已證明其並不能解決公正問題,效率看似很高,卻不免要以犧牲質量為代價,所以,《意見》提出的“不把 SCI 論文相關指標作為直接判斷依據”,是非常必要的。當然,評價並不會因此而退出學術管理,這些問題也不會因為停用某個指標而不復存在,必然會有另一套指標取而代之。為此,《意見》列出了十點舉措,其中關於改進評價方法的舉措,顯然就是用來替代 SCI 指標的。這些舉措是:建立健全分類評價體系、完善學術同行評議、實行代表作評價。其實,分類評價恰恰是 SCI 指標設立的前提(也就是說,現行評價如在分類方面有問題,與 SCI 指標本身無關),代表作評價當然只能由同行來完成,所以,這些舉措中,具有替代 SCI 指標作用或意義的,就是同行評議。我們知道,同行評議既是國際學術界公認的,也是三十年前中國通行的學術評價方法,因此,《意見》所倡導的實際上是向傳統評價模式的回歸。

如前所述,當年引入以 SCI 指標為代表的量化評價,就是因為同行評議不堪信任,我們當年曾經以為,引入了 SCI 指標,評價的問題就能迎刃而解,可事實證明不是這樣。那麼,今天回歸同行評議,公正問題將如何解決?若不能確保評價的公正,又怎能指望據此來合理地分配資源,並以此來促進學術國際化和國際話語權的構建?李劍鳴曾痛陳同行評議在目前中國的困境:“權力支配,人情主導,標準缺失,三者只要居其一,都會使學術評價的意義受到嚴重損害;可是,在我們當前的學術評價中,往往是三者一起發生作用,多路夾擊,來自歐美的同行評議,怎麼可能不水土不服以致徹底變質呢?”這是一個不能迴避的問題。《意見》顯然意識到了這一點,並提出了解決方案:“組織實施部門要完善規則,引導學者在參加各類評審、評價、評估工作時遵守學術操守,負責任地提供專業評議意見……並遵守利益相關方專家迴避原則。組織實施部門可開展對評審專家的實際表現、學術判斷能力、公信力的相應評價,並建立評審專家評價信譽制度。”這些措施三十年前的人大概也是知曉的,如果能做好的話,當年何需引入 SCI 指標?還是知易行難。不用說權力、人情無可避免,即如“建立評審專家評價信譽制度”,這些專家參與的評審過程、每位專家的意見都能完全公開嗎?如果不能,誰來評?怎麼評?如何讓人信服?就筆者所見,因《意見》而起的憂慮甚至質疑的聲音多出自青年教師,且無不圍繞公正問題而發出。可見,要解決同行評議的公正問題,決非易事。

回顧三十年來的討論和實踐,不難發現,評價方法始終是焦點,之所以如此,是因為總是希望找到一種方法,畢其功於一役地解決評價問題,但卻總是事與願違。筆者以為,學術評價不是單純的方法問題,根本出路應是體系的重建,這就需要規劃完整的路線圖,而不僅僅是停用或啟用某種方法。破除“SCI 至上”,實行同行評議,看似找到了解決之道,但也許只是終點回到了起點,三十年前引入 SCI 指標替代的就是同行評議,並以為找到了根本解決方案。結果,耗費了三十年時光,做的似乎只是一次試錯,讓我們再次明白學術評價應回歸以學術共同體為主體的同行評議這個常識。但僅此,懸念仍在,因為我們很難保證同行評議不會再次陷入不公正的泥潭,會不會若干年後,我們又會強烈呼喚量化評價的回歸,讓終點再次回到起點?

(作者係《南京大學學報》原執行主編、編審,全國高等學校文科學報研究會副理事長)

[責任編輯 劉澤生]