

回歸學術初心 優化學術生態

——關於當前我國人文社會科學學術評價的幾點思考

喻 陽

近日,教育部、科技部聯合下發了《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》),明確提出了應扭轉當前科研評價中存在的SCI論文相關指標片面、過度、扭曲使用等現象,推動高等學校回歸學術初心、淨化學術空氣、優化學術生態。這份文件是針對高等學校而言的,集中指涉SCI論文相關指標,而未直接涉及非高校系統和人文社會科學學術評價,但其所聚焦的SCI論文指標乃係近年來學術評價中最具代表性和影響力的指標,而且是以教育部、科技部這樣的高層級權威機構以文件形式下發的指導意見,雖然內容還比較宏觀,指涉的範圍、學科還不够廣泛,執行效果還有待觀察,其制定頒發仍舊是改進我國現行學術評價體系的重要舉措,至少有方向性引領作用。對自然科學、工程技術如此,對人文社會科學亦復如此。下面,筆者參照這份文件的有關精神和內容,結合當前我國人文社會科學研究和學術評價現狀,就幾個相關問題談點看法。

一、理想與現實:學術的本質與學者的初心

學術是對特定問題相對較為專門、系統、深入的學習與探索。其與基礎和專業知識的學習和積累、文獻的收集與整理、實驗操作、社會調查等行為密切相關,細緻、深入的觀察,材料的精心收集與整理,嚴謹的邏輯分析,創造性想像與直覺,都是其活動不可或缺的手段。學者則為具有較豐富的基礎和專業知識、以不同方式從事學術活動、具有較高探究能力的各類人員。依靠文獻梳理、邏輯分析、實驗檢驗、社會調查等手段探究真相、拓展和深化認知、解決問題,應為學術活動最為根本的動機與目的。學術本為志業,學術活動多為清苦、持久、崇高的勞作與努力。追求真知、解決問題、積累和傳承文化、服務社會、改善人類,當為學者的初心與使命。

近年來,我國學術評價體系注重成果數量和發佈載體(機構、平台)的所謂等級,加之職稱評定、績效考核、名譽評比相關規定和要求的不當引導,使快出、多出學術成果,在某種等級的學術媒介(如SCI刊物、中文核心期刊等)發表學術成果成為了許多教學研究人員的主要追求。快出成果使其對學術產物缺乏必要的檢驗和完善,降低了學術成果的質量;為了多出成果,便必然需要縮短單一成果的產出時間,也勢必不同程度地影響成果質量;一些人則更是將一項研究成果以分拆的方式變為多項研究成果,影響成果的完整性,浪費載體資源和讀者閱讀時間資源。更有甚者,一些人為了快出、多出學術成果,不惜抄襲、剽竊、材料數據造假以事之,從而鬧出大大小小的學術醜聞。對發表成果載體某種等級的強調,造成媒介資源的某種緊張,權力、人情甚至金錢等因素因而滲透其間。買賣、“贈送”論文蔚為奇觀。這些都嚴重敗壞了學術風氣,惡化了學術生態。許多學術從業者將學術當作了謀生和獲取名利的手段,急功近利心態蔓延,嚴重背離了其本應秉持的初心和本

當承擔的使命。

二、我們爲什麼要反對或淡化量化評價

在有關學術評價的討論中，對過分重視量化評價的批評可謂持久熱烈，卻又收效甚微。各種形式的量化評價依然故我，大行其道。究其原因，主要在於量化評價的方式簡便易行，可操作性強，因而爲教學科研單位的有關領導和組織管理人員喜好和奉行。其中一些人也認同、深知量化評價中存在的種種問題，但從便於操作的角度考慮，仍信其爲不二之選。在花樣翻新的各種評比、排名中，量化評價仍爲其首選方式。如果不以數量比較爲主要依據，各種評比必定更加耗時費力，甚至難以開展。計算機技術的迅猛發展，使得量化評比更加簡便易行。

量化評價之所以大行其道，其背後的根源除上面指出的簡便易行外，顯然還與學術錦標賽、學術 GDP 主義關係密切，在於量化學術成果與資源配置、人才吸引和社會聲譽等緊相勾連。解決之道自然便是改革有關體制、弱化甚至去除這種關聯。《意見》第八條明確建議和要求：“學校在績效和聘期考核中，不宜對院系和個人下達 SCI 論文相關指標的數量要求，在資源配置時不得與 SCI 相關指標直接掛鉤。要取消直接依據 SCI 論文相關指標對個人和院系的獎勵，避免功利導向。”這樣的意見值得在對人文社會科學領域教學科研的管理中借鑒，參考制定相應的指導辦法。唯有在職稱評定、職務任命、績效考核、榮譽評比、資源配置中大力弱化甚至明確斷絕其與某種類型期刊的發文數量的關係，才能從根本上消解量化評價的動力，抑制學術錦標賽、學術 GDP 主義盛行，從而通過適當減少學術成果數量，特別是大力減少劣質論文數量，增加高水平學術成果數量，來從總體上提高學術成果供給質量，優化學術生態。

量化評價之所以廣受批評，在於其評價結果有失公允，在某種程度上、以不同方式降低了知識生產品質、學術成果質量，助長了投機取巧、急功近利、浮誇浮躁的學術風氣，惡化了學術生態；其以刊評文的簡單化操作，更是不負責任的表現。基於某種載體類型（等級）的量化評價的關鍵問題是將不同質的東西等量齊觀，使本當基於質量（水平）的評價變成了對不同質的東西在數量上的簡單對比，而把專業性強、評價主體素質要求高、艱苦細緻的質性評價簡單化爲某種媒介等級的考量。以刊物而論，不可否認的事實是，發表在同一刊物上的文章質量並不相等，有的可能還差別較大。人們對一個期刊的評價，往往是對其總體水平的評價，並不意味著刊於其中的文章在水平上都是整齊劃一的。至於重點基於他引頻次的影響因子，更不能與質量簡單劃等號，其間存在學科類別（人文學科與社會科學、基礎學科與應用學科）、議題屬性（動態綜述與專題研究、宏大敘事與微觀探討）等方面的區別，還可能被一些人（團體、機構）蓄意操縱。而且，“影響因子”並不同於“質量因子”、“貢獻因子”。而在許多評價活動（如職稱評定、績效獎勵、榮譽評比）中，“質量”、“貢獻”是比“影響”更爲重要的東西（何況那所謂的“影響”還是簡單基於引用的“影響”）。

在我們批評量化評價，提倡質在量先、重質輕量的質性評價時，必須清楚的是，質性評價是一種比量化評價遠更複雜、困難的方法，並不符合人們芟繁從簡、避難就易的選擇取向。原因在於對學術成果的“質”的評判要比基於學術成果數量核計的量化評價困難、複雜得多，把控不好則更容易偏離客觀、公正的軌道。既然如此，我們爲何還要大力提倡基於專業素質和學術品質的質性評價，而對看似客觀公正、簡便易行的量化評價多加批評呢？我們之所以要強調更加重視質性評價而弱化量化評價，根本原因在於，通過分析不難發現，量化評價無論如何不可能是真正客觀、準確和公平的，而做得好的質性評價無疑更接近客觀、準確和公平，儘管由於種種原因，“做得好”不是一件容

易的事情。在找不到十全十美、最優的方法時，我們只能努力尋找相對更優、次優的方法。筆者認為，在人文社會科學學術評價中，從公平、準確和導向作用方面考慮，相較於量化評價，質性評價就是一種相對更優的選擇。而強調綜合採用質性評價和量化評價方法，企圖通過兩者結合來避免各自存在的問題，其中也存在種種困難需要克服。其中方向性的引領必須是質在量先、重質輕量。只有在同質的條件下，量的比較才有意義。

三、分類評價的必要性和方法

在學術評價中，另一廣受爭議的問題即在於漠視不同類型的學科性質差異，用一把簡單的尺子衡量不同類型的學科、人才，造成事實上的不準確、不公平。這在各種類型的大學排名中表現尤甚。有鑒於此，近年來，在黨和國家、有關主管部門下發的涉及人才評價、學科評估、資源配置等的各類文件中，都明確強調要實施分類評價。但在現實中，無視學科屬性的籠統評價依舊盛行，針對性強、設計科學的分類評價難見蹤影。究其原因，主要仍是貪圖簡便的惰性思維使然。

就人文社會科學而言，除越來越豐富和深化的學科和專業外，從更為宏觀的角度考慮，亦存在人文學科與社會科學、基礎學科與應用學科的典型分野。它們在研究對象、方法、結論性質與效用等方面都存在顯著區別。人文科學(學科)主要探究人的精神活動及其品性養成，精神、文化為其核心內容，價值導向為其顯著特色，教化、涵養、“成人”為其探究主旨，文本闡釋、邏輯思辨、體驗(內省和移情)、想像為其主要方法，歷史性、批判性、主體性、相對性、多樣性為其學科特徵。文學(語言文字學、文學理論、文學批評、文學史等)、歷史學、哲學為其主幹學科。社會科學是以社會現象為研究對象的科學，其任務是描述、展示、解釋各種社會現象，預測、揭示社會變化趨勢(或曰發展規律)。在研究方法上，社會科學更注重模型建構、統計分析、社會調查。相較於人文學科，其結論更具有普遍性、實用性。政治學、經濟學、社會學為其主幹學科。

除人文學科與社會科學的區分外，還存在著基礎學科與應用學科的區別，其在研究對象、方法、結論的性質和效用方面，亦存在顯著不同。基礎學科更注重對研究對象的恆常性質及其演化規律的探討，結論具有某種普遍必然性；應用學科則重在運用基礎學科提供的理論知識，理解、揭示研究對象的特質及其與環境條件相互作用的規律，解決各種現實問題，滿足種種現實需求。

除學科性質不同之外，與學術評價關係密切的另一個問題在於，學術評價除了存在主觀與客觀相互限制、相互依賴，相對性與多元性不可避免之外，本身也存在不同的維度。大其要者，即存在水平(質量)、影響和效用(貢獻)的區別。一項學術成果的水平，主要看其有無開新之處以及開新的難度、深度和大小(亦即所謂原創性及其大小)，材料是否真實充分，論證是否嚴謹，文字表達是否規範順暢；其影響則主要看其被多少人關注到，被什麼人關注到，以及關注的程度如何；其效用則在於其對於人類知識積累、文化傳承或解決實際問題是否發揮了作用或作用的大小。考量一個學術成果的效用(貢獻)，首重其對解決緊迫社會問題的直接幫助或參考作用。

區分了以上評價維度後，顯然，對於人文學科與社會科學，基礎學科與應用學科，其評價重點應該有所不同。對人文學科，水平(質量)考量是更重要的維度；對社會科學，影響和效用考量是更加重要的維度；對基礎學科而言，水平和影響是首先需要考量的因素，可以適當淡化對其短期和直接效用的考量；對應用學科而言，考量其解決現實問題的效度是更重要的方面。

正是因為人文學科與社會科學、基礎學科與應用學科在研究對象、方法、成果性質等方面都有顯著的不同，而評價又存在質量、影響力和效用的不同維度，為了更加客觀和精準，所以需要分類加

以實施。明確不同學科的不同性質和學術評價的不同維度，當可為具體實施人文社會科學的分類評價提供方向性的指引。

四、同行評價應該注意的問題

同行評價是又一個廣受關注的學術評價問題。在近年來黨和國家及有關管理部門下發的關於人才評價、學科評價、資源配置等各類文件中，亦普遍強調在實施相應評價時，應秉持同行評價原則。《意見》第四條明確提出了“完善學術同行評價”的要求。在這裏我們一方面需要指出，只有學術同行才是適格的學術評價主體，必須明確倡導、推行同行評價原則；另一方面，與質性評價之於量化評價一樣，相對於當前行政管理人員參與其中、以某種類別的成果數量為主要依據的泛化（混雜）評價相比，同行評價是一種遠更複雜、操作難度更大的評價。運作得不好，與其追求更加準確、公平的初衷相反，同行評價將是比目前多種人員參與其中、量化評價為主的方式更不客觀和公平的評價。因此，我們既要堅持同行評價的方向不動搖，又須在制度設計、操作方式釐定和選擇上下大力氣，以真正實現更加公平、準確的學術評價。

同行評議之所以必要，是因為只有從事相同或相近研究的同行（大同行和小同行）才能够更加準確地判斷一個學術成果的原創性（有無、大小、難易）和材料是否真實與充分，而這些方面是判斷其質量（水平）最為關鍵的因素，對人文學科、基礎學科尤其如此。即便是對其影響力和貢獻的評價，雖然在這些方面量化評價有一定的優勢，但同行的感受和意見較之非同行的感受和意見也往往更為真切。

為了追求客觀和公正，在學術評價的主體選擇上，學術同行是不二之選。或許將來的人工智能可以大大豐富、優化學術評價的主體選擇，但就當前現實而言，只有學術同行才是適格的學術評價主體。但也必須看到，因為“學術同行”都是現實的、處於各種複雜的社會關係中的人，如果沒有好的制度設計和操作措施，同行評價就很容易變成一種比目前盛行的量化評價更不公平和客觀的方法，這是我們必須予以高度警惕的。

事實上，從同行評議興起後不長一段時間，其在現實操作中的種種弊端便顯露無遺，因而飽受批評。在人文社會科學學術評價中，同行評議主要容易出現以下四個方面的問題。一是受立場、價值觀影響，評價者易對表現出與自己相同或相近立場、價值觀的學術成果給與積極、正面的高評價，而對與其立場、價值觀相左的學術成果給與消極、負面的低評價，即表現出某種“黨同伐異”傾向；二是受人情、關係影響，評價者易對涉及與自己有親緣、學緣或地緣等關係的人的學術成果給與高於其真實水平的評價；三是受權力、利益影響，或則有權勢人物打招呼，或則評價結果直接牽扯到自身或與自身高度相關者的利益，因而不願、不能做出實事求是的評價；四是受限於自己的認知水平，表面上看似同行，實際上不具備準確評判評價對象的能力，甚至出現“小專家”評“大專家”的情況。

同行評價雖然存在上述種種問題，但基於明確的認識和堅決的行動，非無解決之道。解決這些問題，可以從主體選擇和制度設計兩個方面著手。從主體選擇方面著手，就是要充分結合評價對象的情況，遴選具有良好道德修養、學術品質和專業水平的專家充任同行評議專家；從制度設計方面著手，則是要建立必要的迴避制度、反向公開制度（公開評審人的身份和評價內容）、誠信建檔制度、申訴與回溯制度、監督制度等，用制度的激勵與約束力量促使評審專家公正、認真履行職責。經過這樣努力完善的同行評價，雖然也不可能盡善盡美，但肯定比基於簡單限定成果發表載體品級、以發表成果數量多少為主要依據的量化評價更加客觀、公正、精準和恰當。

五、餘論：對《意見》精神的幾點補充理解

《意見》自頒發以來，在學術界、社會上迅速引起了廣泛關注。其中有些理解，筆者認為並不準確，甚至嚴重偏離了《意見》意圖引領的方向。有鑒於此，筆者再就其中幾個問題談點補充理解。

1. 《意見》的導向，不是要用其他的量化評價指標取代 SCI 論文指標，而是要“堅決擯棄以刊評文”。現在有一種意見，認為《意見》的發表，主要是要取消和淡化 SCI 論文指標的作用，避免或減少高水平論文外流，從而提升本土指標的作用。筆者認為這不是《意見》的精神主旨所在。客觀上，《意見》的執行可能會讓一部分高水平論文回流，從而有利於本土期刊文章來源的品質提升。但《意見》更為重要的精神，是要反對任何形式的“以刊評文”、量化評價盛行，糾正價值追求扭曲、浮誇浮躁、急功近利的學術風氣，營造良好學術生態。

2. 《意見》的意旨，不是要完全否定 SCI 這樣的數據統計產品的價值，而是提倡合理使用。SCI 作為科技文獻索引系統，是基於論文引用的一種數據統計結果，對於個人、機構利用有限的資源訂閱使用更優質的期刊，指導作者選擇合適的論文發表渠道和平台，無疑是具有很高參考價值的。任何精心設計、用心製作的學術分析與評價工具，對於相關從業人員都有很高的參考價值。但是，一旦將具有參考作用的東西奉為標準，更以量化評價的方式使其直接或間接與利益或名譽掛鉤，就不可避免地會發生某種異化，這才是我們要警惕和反對的。

3. 準確理解代表作制。近年來，針對以刊評文、量化評價的種種弊端，在職稱評定、人才評價等工作中，實行“代表作制”的呼聲日烈。也有一些知名高校和科研機構在相應工作中提出了推行代表作制。但必須看到，真正實施代表作制的單位還不多，一些嘗試實行代表作制的單位也沒能很好堅持。之所以出現這種情況，根本原因還是受學術錦標賽、學術 GDP 主義盛行的大環境影響和同行評議制度執行得不完善。實行代表作制的根本目的，是重質輕量，克服量化評價的不科學、不恰當、不公平。其背後的深層次原因，仍然是要扭轉浮誇浮躁、急功近利的不良學風。《意見》第五條“規範各類評價活動”中，明確提出了“實行代表作評價”，這是以高規格文件的形式，明確規定了在人才評價等評價活動中，採用代表作評價方式，從而使這一為人久久呼籲的應行制度在有關評價活動中變成了規定動作，而不是自選動作。當然，筆者認為在實行這一制度時，需要注意的是，代表作應由被評價人、機構自主提供，而不是由評價主體、機構基於刊載平台類別、品級加以規定。什麼是自己的代表作，作者本人應該有最為清楚的認知。此外，根據不同的評價對象和目的，代表作的數量以 2~5 種為宜，不可過多，以使“代表作”名副其實。

筆者始終認為，學術評價是一個知易行難的問題，其存在的弊端、改進的方向不難認知，但改進的難度超乎尋常。之所以如此，除各種內生因素外，還存在複雜的文化、社會、心理原因。如果不從根本上革除現行的科教體制弊端（重研輕教）和彌補資源配置方式缺陷（重量輕質），大力營造風清氣正的科研文化，相關改革勢必舉步維艱。但無論如何困難，明確方向，久久為功，必定進步可期。這次兩部門《意見》的頒行，就是制度建設上的一項重大舉措，意義重大而深遠。其明確的是方向和舉措，致力改變的是被扭曲、被傷害的學術風氣和生態，召喚、踐行的則是學術的初心和學者的使命。

（作者係《新華文摘》總編輯、編審）

[責任編輯 劉澤生]