

理念與實踐：當前中國人文社會科學 學術評價中的幾個問題

喻 陽

[提 要] 當代中國學術評價因牽連甚廣而備受矚目與爭議，人文社會科學學術評價尤此為甚。本文在釐定學術評價關鍵概念，簡要探析學術評價主體、目的、內容、方法的基礎上，重點探討了學術成果質與量的關係、分類評價的重要性、同行評議存在的問題及應對之道。文中特別強調，評價學術成果，應秉持質在量先、以質為主、質量結合的原則。對學術成果的評價存在水平、影響力和社會貢獻等不同維度。對人文學科與社會科學、基礎學科與應用學科應堅持分類評價：對人文學科，應重點評價其水平，對社會科學，應重點評價其影響力；對基礎學科，學術水平和影響力是主要考量因素，對应用型社會科學，則應重點考量其社會貢獻。應採取有針對性的措施，切實有效解決同行評議中存在的種種問題。

[關鍵詞] 學術評價 人文學科 社會科學 分類評價 同行評議

[中圖分類號] C03 **[文獻標識碼]** A **[文章編號]** 0874 - 1824 (2018) 01 - 0106 - 14

在當今世界上絕大多數國家或地區，學術評價主要與少數人直接相關——專門的科研機構和高等院校的教學科研人員。而在當代中國，這一情形似乎有很大的不同：除專門研究機構和高等院校的教學科研人員外，人數遠更眾多的中小學教師、多數醫療機構的醫務人員、領域十分寬廣的工程技術人員，還有名目繁多的經濟師、會計師、政工師……^①甚至是絕大多數研究生，還有數量不少的本科生，都被捲入了滾滾學術洪流中，都成了不同程度的“學術工作者”。對領域如此廣泛的人員的學術成果評價，變得與其工作、生活，甚而是一生前途與命運息息相關：或則影響其單位聘用、職稱評定、業績獎懲、職務升遷，或則影響其學業完成與學位得失，多年辛苦、一生前途皆與之緊相勾連。學術評價由是而變得越發重要，由之滋生的種種問題由是而備受關注，飽受質疑。有鑑於此，本文對其中一些問題談點個人看法。

一、關鍵概念界定：學術、評價、人文學科、社會科學

任何研究，尤其是人文社會科學研究，對核心概念的辨析或約定都是其基礎工作。如果涉及

那些理解不一甚或爭議很大的概念，就更其如此。事實上，由於態度、理解力、投入時間和精力的差異，更由於語言中普遍存在的一詞多義等原因，為了讓討論避免表面熱鬧而實無交集，為了不致在“由用語的領域形成的群島”^②中迷失方向或行錯路線，對相關概念的辨析甚或約定就變得十分必要。我國著名物理學家嚴濟慈在 20 世紀 30 年代撰寫的《論公分·公分·公分》中以“公分”一詞的歧義累贅以致令人不快為例，指出了釐定名詞概念的重要性：“凡百工作，首重定名；每舉其名，即知其事，斯為上矣。”^③

依據探究對象與學科特色的域分，我們可以將“科學”、學科分為自然科學（工程技術）和人文社會科學兩大類別。兩者的關鍵區分在“自然”與“人為”。自然科學研究的對象是客觀之物的本來性質及其變化規律，不受或很少受人的思維、意志、情感等主觀因素影響，因而是“自然”的。人文社會科學研究的對象都不同程度地與人的活動有關，都會在不同程度上打上人的理智、情感或意志的烙印，因而是“人為”的。

本文後面討論的問題主要與人文社會科學有關。在人文社會科學內部，又可區分為人文科學（學科）和社會科學兩個部分，兩者在研究對象、方法和結論的性質（普遍性、可靠性）等方面，都存在顯著的區別。

另一方面，無論是自然科學還是人文社會科學，均有基礎科學（學科）與應用科學（學科）之分。基礎科學重在探究研究對象的一般特性，意在認識規律，獲得普遍知識；應用學科則重在應用相關原理，解決實際問題。基礎學科的主要功能是“求真”，應用學科的主要功能是“務實”。當然，無論是基礎學科還是應用學科，只要是學科研究的對象，都有某種程度的一般性、普遍性、共通性，不會純為個別之物，而只有最根本（一般、普遍）、較為根本（一般、普遍）和相對具體、較為具體之別，表現為範圍廣狹、程度深淺之分。

回到本文的關注重點上來，與本文議題關聯度很高而又意涵複雜，容易導致誤解、誤用的四個關鍵概念是“學術”、“評價”、“人文學科”、“社會科學”。在我們後面對學術評價問題開展集中討論之前，有必要對這四個“關鍵詞”略加辨析，或者說做一個“約定”。

1. 學術

相對而言，學術乃指較為專門、系統的學習與探索，以區別於較為低級的基礎性學習（如幼兒、中小學教育）和以實用目的為主的學習（如各種技能性培訓學習），或較為浮泛、淺近的思考與意見表達。學術活動是對存在物（自然、社會、思維等）的性質（物性）及其變化趨勢（規律）的系統、深入的思索、探究與闡釋，旨在積累知識、探求新知、解決問題。學術與高等教育和研究機構息息相關。純粹的知識性學習與學術高度相關，但還不是學術。只有探究性的學習和思考、實驗等以驗證舊識、探索新知和解決問題為目的的活動，才是學術活動。由於研究對象複雜多樣，對其的探究非分門別類難以開展，因此學術往往以學科、專業和領域分殊的形式存在和展開。學術活動因其專門和深入，往往需要不同程度的“規訓”和“建制”。隨著人類知識積累日益繁富，研究對象日益深入，社會生活日益豐富，經濟日益發展，技術日益進步，學術活動呈現為日益專門和深入、針對性越來越強的樣態。另一方面，由於研究對象的複雜性和解決問題的難度日益增加，“跨學科”探究變得日益必須和普遍。

關於學術，尤可注意“學”與“術”的區別與聯繫。梁啟超曾專門為文對之做過討論，^④指出“學”主要是對真理的探求（“發明”），而“術”則是用掌握的科學知識（“發明之真理”）解決實際問題（“致諸用者”）：“……則學也者，觀察事物而發明其真理者也；術也者，取其

發明之真理而致諸用者也。”^⑤梁啟超強調“學”是“術”的基礎，“術”是“學”的應用：“學者術之體，術者學之用。二者如輔車相依而不可離。學而不足以應用於術者，無益之學也；術而不以科學上之真理為基礎者，欺世誤人之術也。”^⑥嚴復亦抱有與梁啟超大體相似的看法：“蓋學與術異，學者考自然之理，立必然之例；術者據既知之理，求可成之功。學主知，術主行。”^⑦

梁、嚴之區分“學”與“術”，猶如今之區分基礎科學與應用科學，或“科學”與“技術”。兩者聯繫緊密而又絕不相同，由是對其評價既緊相聯繫而又應該有所區別，對此我們在後面將着重予以討論。

還需予以說明的是，學術研究，尤其是人文社會科學學術研究，並非純然的理性求知過程，而是一個包含了情感、態度、旨趣、利益、立場等非理性而現實的因素在內的複雜過程。這既體現為人文社會科學研究本身的複雜性，也無疑增添了人文社會科學學術評價的複雜性。

2. 評價

一般而言，評價乃指評論價值大小或高低。價值往往指積極作用。在哲學上，不同的思想視域和思想方式對價值有不同的理解。如價值倫理學的主要創立者、知識社會學的先驅、德國哲學家舍勒把價值理解為客觀的實在，它所反映的是存在事物的等級狀態。人們還可以從人與對象物的關係的思想視域中理解價值現象，即價值可以指人根據自身的需要、意願、興趣或目的對與他生活相關的對象物賦予的某種好或不好、有利或不利、可行或不可行等特性，也可以指對象物所具有的滿足人的各種需要的客觀特性。而人類生活經驗的不同領域以及不同需要又規定了價值類型的多樣性，如經濟價值上的有用與無用，倫理價值上的善與惡，審美價值上的美與醜等。^⑧價值是相對於人而言的，是對象物對作為主體的人的期望、欲求的滿足，滿足程度之大小，決定其價值之高低。這是價值的主要表現形態，具有很強的功利、效用色彩。價值既不純然是主觀的，也不純然是客觀的，而是主觀與客觀的契合、統一。價值對象之實用（功利、效用）價值，在其多方面價值中，無疑居於基礎和優先的地位，是評價者首先要評價的。至於評價對象的道德價值或審美價值，雖然其與功利性漸行分離甚而完全超越，但並非價值的主要和普遍形態。“倉廩實則知禮節，衣食足則知榮辱”，道德價值須在實用需求得到很大滿足後漸次呈現。至於一物之審美價值，則更需審美主體衣食無憂而生閒情逸致，有條件感受（享受）“沒有利害關係的愉悅”，方可卓然而至。實用價值、道德價值和審美價值次第而生，前者更為重要而緊迫，後者更顯超越與崇高，但後者以前者得到相當程度之滿足為基礎和前提。當然，一個衣食無憂之人，其德性提升和審美追求的價值欲求會相應變得重要和強烈。

明確價值的分殊和優先序，有助於我們更為全面、準確地對學術開展多層次、多維度的評價。

價值具有相對性與多元性。由於價值是價值主體與價值客體的契合，首先表現為價值客體對價值主體的效用。一方面，評價不會是純然的“無中生有”，而總有某種形式的客觀依據在。另一方面，作為價值主體的人，在立場、態度、人生經歷、知識積累、所處環境等方面都是獨特的，是“個體化”的“這一個”，因而使評價又不可避免地表現出不同程度的相對性，表現出明顯的“因人而異”情態。“世界上不存在兩片完全相同的葉子”（萊布尼茨），更不可能存在兩個完全相同的人。由於價值主體的獨特性，同樣的存在物對不同的人，其效用、好壞、美醜便不可能完全相同。“一個人的美食是另一個人的毒藥”，“情人眼裡出西施”，即此之謂。這種評價的差異與多元，同樣也會反映、體現在學術評價上。甚至因為學術評價的相對複雜性，評價結果的差異將會更形突出和明顯。

3. 人文學科

人文學科是對人性的狀貌及其養成、生活的意義和價值的系統思索與探討，意在展示和判斷人性，探究人性形成、完善的條件，提高人的素養，精神、文化為其核心內容，價值導向為其顯著特色，歷史性、批判性、主體性、相對性、多樣性為其學科特徵，理解、解釋、體驗（內省和移情）、想像、語言分析和文本釋義為其主要方法，以人豐富的知情意為其主要內容的心智、情感世界為其主要材料，教化、昇華、涵養、“成人”為其探究主旨。“人文科學致力於從人內在的精神世界和作為人的內在精神之客觀表達的文化傳統、文化體系的統一中去總體地探索和把握人類的價值、理想和追求。”^⑨人文學科“追求人性的和諧和完善，追求人生的整體和內在意義，並且以人、人性、人生為視角，從總體上反思和關照外在於人的社會和自然”。^⑩人文科學（學科）主要探究人的精神活動及其品性養成，以文學（語言文字學、文學批評、文學史等）、歷史學、哲學為其主幹。

4. 社會科學

社會科學是以社會現象為研究對象的科學。其任務是描述、展示各種社會現象，揭示社會變化趨勢（或曰發展規律）。其概念的首次提出是在啟蒙運動時期，是法國重農主義者杜爾哥和他的朋友孔多塞侯爵首先使用了“社會科學”（*La science sociale*）這一概念。相對於人文學科和自然科學，社會科學相對晚出，興起於 18 世紀中後期，在 19 世紀中葉以後方形成建制。“嚴格說來，社會科學是在 19 世紀才出現的。”^⑪美國社會科學史家彼得·曼尼克斯指出，如今學科林立的社會科學是 19 世紀末才以一種公認的現代形式出現的，是學術界對工業化時代的挑戰所作出的回應。^⑫美國學者 D. W. 卡爾霍恩也曾指出：“社會科學主要是技術革命以及隨之發生的社會變化的結果。”^⑬由世界體系思想家華勒斯坦擔任主席、一批國際知名學者共同組成的古本根重建社會科學委員會在他們的集體研究成果《開放社會科學——重建社會科學報告書》中指出：“我們能夠對人類的本性、人類彼此之間的關係、人類與各種精神力量的關係以及他們所創造並生活於其間的社會制度進行理智的反思，這一想法至少同有記載的歷史一樣古老。普通的宗教文獻探討這類問題，我們通常所說的哲學文獻同樣也探討這類問題。除此之外，還有各種世代相沿的口傳智慧，它們往往在歷史上的某一時期獲得書面定型。這些智慧果實要麼以啟示的形式呈現出來，要麼就是對某些永恆真理所作的理性演繹。……我們今天所說的社會科學正是這種智慧的後裔，不過它乃是一個關係疏遠的後裔，或許經常還是一個忘恩負義的後裔，因為社會科學有意識地給自己規定了一個任務，那就是去追尋超越於任何公認的或演繹的智慧之上的真理。社會科學是近代世界的一項大業，其根源在於，人們試圖針對能以某種方式獲得經驗確證的現實而發展出一種系統的、世俗的知識。”^⑭他們明確表示：“在 1850 至 1945 年期間，人們對一系列的學科進行了界定，這些學科共同構成了一個可以‘社會科學’名之的知識領域。”^⑮他們還特地指明：“……社會學、經濟學和政治學構成了一個以國家為中軸的三一體，從而鞏固了它們作為核心社會科學（以研究普遍規律為主旨）的地位。”^⑯正如有學者指出的：“總的說來，社會科學作為一個獨立的學科體系產生於 19 世紀，是近現代社會結構化的產物，是適應大工業生產、城市等大規模社會結構的管理需要而產生的。它把近代以來產生的結構化的或大規模的社會組織、社會群體、社會關係作為研究對象。社會科學的主要學科，如經濟學、社會學和政治學的研究對象都是近代以來才得以形成和發展的社會現象。這些學科的產生本質上是近現代社會變化的結果。……離開了近代以來市場的發育和社會結構的複雜化、市民社會的相對獨立性，就無法理解現代社會科學

的產生和發展。作為社會科學的研究對象的社會本體，不僅決定著現代社會科學的產生，而且規定著社會科學的社會功能。現代社會科學產生和發展的根本意義，就是對客觀的社會變化進行說明和預測。”^⑩在分別界定、簡要介紹了人文學科與社會科學之後，還需簡要說明的是，人是社會的人，社會由人組成。由於“人”與“社會”的密不可分、複雜關聯，導致了人文學科與社會科學的複雜關聯。人文學科與“社會”高度相關，社會科學與“人”形影不離，導致“人文學科”與“社會科學”的區分只具有相對價值，是就其“人文性”和“社會性”的強弱、多少相對而言的，而難以判然劃分。如以創制和維護社會運行秩序為旨歸的法學，其探究目的、內容均與人性和社會緊相關聯，既有很強的“人文”色彩，又與“社會”息息相關，但其探究問題更具“社會”特色，因此我們將其主要歸為“社會科學”。

二、學術評價的主體、目的、內容與方法

關於學術評價的主體、目的、內容（要素）與方法，所有從事相關活動的人肯定都會默有所知，但多不完整、準確和清晰。為使對當前中國學術評價中存在的一些問題的討論更加準確，更有針對性，需要對之先做一些辨析與討論。如同前面的“關鍵概念界定”一樣，這也是我們明晰學術評價理念、探討學術評價實踐的基礎性工作。

1. 學術評價的主體

學術評價的主體是而且只能是人或人的集合（機構、團體等由人形成的組織）。當今世界，人工智能技術進展迅速，取得了許多令人震驚的成就。“掃地機器人”日益走進普通人的生活；AlphaGo 和 Master（大師）的驚人棋藝，極大地激發了大眾對人工智能的進一步關注。完全可以想像，在不遠的將來，我們可以把學術評價部分或全部地交給機器人來做。但目前還不能。而且我們同樣也可以想像，即使在不久的將來，學術評價可以交給機器人來完成，或主要通過人工智能方式進行，但某種程度的弊端不可避免，以致仍需有血有肉的人來糾偏。到目前為止，學術評價的主體只能是人和各種人的集合體。凡是人，皆可成為學術評價的主體。但作為學術評價主體的人，當然存在專業與外行、教育程度、知識和經驗的顯著差異。相對於確定的學術評價對象，具有相應專業知識、受過特定規訓的“專業人士”的評價，無疑更為精準、更具價值，這樣的專業人士更應充當學術評價的“主體”。

學術評價主體具有多元化特徵，既可以是具有不同生存處境、知識背景、專業旨趣、利益立場、文化心理特徵的個人，也可以是由這樣的人組成的目標各異、水平不一的多種多樣的類、群、單位、組織等集合體，如教學研究機構、專門的科研管理單位、各種各樣的評審委員會，數量眾多的知識生產、人才使用組織，還有數量日益增多的專門的學術評價機構。

2. 學術評價的目的

馬克思曾以蜜蜂築巢與人的活動的區別，說明人的活動通常是有意識、有目的的。學術評價活動亦復如是。目的引導、制約手段和結果。在我們集中討論當前中國學術評價活動中存在的一些問題及解決之道之前，對學術評價的目的略做討論，當不為無益之舉。

學術評價的目的可以從抽象與具體，宏觀與微觀角度加以界分。學術評價的微觀（具體）目的，主要是評定學術成果提供者的學術水平，從而為對學術研究人員的聘用、晉升、學術資助等決定提供依據，或者是對學術成果提供者給予某種獎勵（評獎），旨在對學術成果提供者的勞動（也包括天賦等因素）給予某種承認或激勵，從而達到維護與學術活動有關的機構的良性運轉，

鼓勵、督促學術活動從業者積極向上的目標。學術評價的宏觀（抽象）目的，是通過對學術成果提供者的承認和激勵，推動學術事業的發展，進而增加人類知識積累，提高人類認識能力、道德情操和審美情趣與能力，解決與人類物質和精神生活息息相關的種種現實或長遠問題（工程技術、應用類社會科學學術成果）。

3. 學術評價的內容

學術評價即特定主體對學術成果（或學術成果提供者）的水平高低、影響和貢獻大小所做的判斷。學術評價主體可以是個體或機構，或因特定目的而聚合起來的“共同體”；學術評價的對象可以是一項或多項學術成果、學術成果提供者（個人、機構、載體等）、特定時段（朝代、世紀等）或空間（單位、地區、國家等）的學術狀況或水平。

無論學術評價的對象如何多元與複雜，皆由一項項學術成果累積而成。這種單項的學術成果，通常以文章、著作的形式存在。種種科研項目，最後也多以文章或著作的形式“結項”（當然，學術成果並非只有文章和著作，這也是必須明確的）。因此，對文章、著作的評價便成為學術評價的基礎性工作。其中，對文章的評價又更為基礎和普遍。因此我們後面對學術評價的內容的探討，便集中在對學術文章評價的探討上。

我們認為，對學術文章的評價，包括水平、影響力、社會貢獻等多種維度，這些維度各有不同的呈現與指向，不可混為一談。

（1）一篇文章的水平，主要表現為其開新（觀點、材料、方法、問題）的多少與程度（深度與難度），論證是否充分而簡潔，無或甚少語法與邏輯錯誤，標點符號、數字用法、註釋體例是否規範等。當然，如果一篇文章除了在上述方面都做得很好之外，文字表達還很流暢，甚而有音韻之美，讀之讓人產生較強或很強的閱讀快感，那我們就認為這樣的文章是更有水平的。換種方式來講，一篇文章之水平高低，可以主要從兩個方面去衡量：一是開創性如何，二是論證表達水平如何。

（2）一篇文章的影響力，主要取決於其被多少人關注到，被什麼樣的人關注到（所謂“影響有影響力的人”），關注程度如何。如果一篇文章被更多的人閱讀，甚至被他人引用（由是有期刊的“影響因子”一說），那毫無疑問是更有影響的。當然也需注意，與“水平”之分高低、社會貢獻之分大小（兩者均以“零”為極限）不同，影響除了有“大小”之分外，還有“好壞”之分。一篇文章如果在價值導向上有違人類、族群、“共同體”比較一致、能推動社會進步的價值取向，甚至反其道而行之，那這樣的影響就是“壞”的、“負面”的。

（3）一篇文章的社會貢獻，主要看其對人類知識積累、文化傳承或解決實際問題所發揮的作用之有無或大小。考量一篇文章的社會貢獻，尤重其對解決現實問題的直接幫助或參考作用。因為其所“解決”或“幫助解決”的問題是“現實”的、“緊迫”的、“關係重大”的，與人們日常生活息息相關，如醫學問題、環境問題、社會治安問題、財政金融問題等，所以，人們對一項學術成果的社會貢獻的評價，往往在很大程度上受到“即時性”、“緊迫性”的影響。這對於那些一時用處不明顯，而在以後的長時間中發揮了或將發揮重大作用的學術成果，難免“有失公平”。但這幾乎是不可避免的，因為人們的評價往往基於即時的感受和認知，而那並未降臨的“未來”功效，畢竟難以被人當下普遍、強烈感知。

一篇文章的水平、影響力和社會貢獻彼此之間通常是正相關的，但不一定線性、顯著正相關，而是呈現為種種複雜情形。

4. 學術評價的方法

大抵與學術研究方法相類似，人們在開展學術評價時，依據學科性質、對象特質和評價目的等因素，在方法上可以有質性評價、量化評價與綜合評價之別。

(1) 質性評價：主要是對學術成果的創新特質、探究深度與難度、結論可靠性和價值、論證水平進行的品質判定，相對忽略學術成果以篇次、長度等表現出來的量的特性，重點關注評價對象相對而言的新、精、深程度。

(2) 量化評價：主要是運用數學方法，對學術成果及其提供者進行以公開發表的文字作品的篇冊數、各種層次的項目獲得和通過數、各種形式的資助金額多少等進行量的統計，按“多者為優”的原則確定評價對象的相應優先序。現代計算機技術、特別是互聯網技術的高度發展，使相關統計變得及時和準確，為開展學術成果的量化評價提供了很大的方便。

(3) 綜合評價：即為上述兩種方法的結合，既對評價對象的水平高低，又對學術成果的數量多少進行評價，力圖取長補短，相互糾偏，以求取評價結果的客觀和公正。此種方法運用恰當，其長處十分明顯。但顯然存在著在結合過程中何者為先、如何賦權的棘手問題，處理不好，則可能結合的是兩者之短，而無一者之長。這是需要評價者採用此種方法時特別注意的。

必須指出的是，沒有絕對客觀公正的學術評價，沒有任何一種評價模式和方法“放之四海而皆準”，不存在完美的單個學術評價方法。任何一種評價方法，在稟有長處的同時，又難免存在短板，“取長補短”也絕非處處可行。所謂矛盾無時不有，無處不在，當為我們應該具備的辯證法直覺與信仰。而在種種矛盾、衝突的方法、工具中，我們只能“兩害相權取其輕”，尋求相對更為合理、客觀、公允的方法。對單獨一種方法的取捨是如此，對方法的綜合運用亦復如此。保持理智的謙虛與態度的勤勉，綜合運用不同的學術評價方法，相互取長補短，則無疑可以提高評價的準確度，使我們的評價變得更客觀、全面和公正。

三、當前中國學術評價中特別值得注意的幾個問題

在前面梳理了關涉本文主旨的“學術”、“評價”、“人文學科”和“社會科學”概念，簡要論述了學術評價主體、目的、內容和方法後，接下來我們集中討論當前中國人文社會科學學術評價^⑩中的幾個特別值得注意的問題。這些問題在學理上並不高深，但在實踐中卻普遍存在，且“為害不淺”並難以“解決”，因而尤有深入辯證、仔細擘畫的必要。

(一) 如何看待學術成果的質與量

一段時期以來，“重量輕質”成了學術評價中最為人詬病的問題，甚至有學者稱量化評價是學術事業的毒藥，激憤之情令人咋舌。而批評歸批評，激憤歸激憤，在當代中國學術場域中，甚至在許多並非典型的學術場域（如中小學校、醫院、工程設計等）中，重量輕質、量化評價仍大行其道。由此帶來的弊端十分明顯：低水平甚至無水平可言的所謂學術文章大量發表，以致學術資源（各種來源的人力、物力和財力）大量浪費；甚至即便在自然科學、工程技術這樣以純然、客觀地追求真知或致力於解決問題的領域，為了“學術成果”而大肆造假之舉，亦時有所聞。“與真理為伍”的學術天地，竟成了騙子時有出沒的地方。那麼究竟應該怎樣看待學術成果質與量的關係，究竟應該怎樣對待量化評價呢？對此，筆者的基本看法是：以質為主，質在量先，質量結合。

1. 在學術成果質與量的關係上，毫無疑問應重質輕量，質在量先。辯證法學說主張，使一事物區別於他事物的是該事物的質，而非其量；在一定程度和範圍內，某物的量變並不能改變其質。

作為優化人性、探求真知或解決重大現實生活問題的學術活動，質比量重要、優先考慮質乃天經地義，也尤其需要我們加以明確認知和強調。回望源遠流長的中國文化，我們亦可從中獲得啟示。在中國文化流變中，儒釋道無疑為學術、思想中影響最顯著者。在儒家學說中，影響最大的“學術成果”毫無疑問當推孔子的《論語》^⑨，而《論語》不過近一萬六千字；老子的《道德經》^⑩毫無疑問是道家的主要經典，亦只“五千餘言”，至有《老子五千文》之稱；《心經》僅有260個字，因便於持誦，在佛教信眾中極為流行。由是觀之，思想、學術之廣泛流傳和持久影響，並不取決於字數多、篇幅大等“量”的規定性，而其思想意旨等“質”性規定顯然更具意義。

2. 當然，同時也要看到，量化評價只要客觀公正，方式、程序得當，在學術成果評價方面也具有參考價值，有的還很重要。假如兩個評估對象的學術成果在單獨項上“質”相同而“量”不同，自然應該以量多者為勝。但其前提是“質”相同。同時還要看到，一些機構所做的量化評價，是針對機構或刊物而非針對個人的，其在評價機構和刊物上將有重要參考價值，但不能作為對個人的評價。一個機構所具有的學術品質，並不為其每一個組成人員所具有；一個刊物所具有的總體學術水平，並不為刊發其中的每一篇文章所具有。人們對一個機構、一個刊物學術水平的感知或評價，是就其整體狀況相較而言的，是對其“平均值”、對某種水平（高或低）的文章所佔比例大小的評價，屬概率性評估，而非對每一個作者、每一篇文章的精準評價。因此，針對某個體、某篇文章的學術評價，還需具體而為，由專家同行進行客觀、公正、細緻的水平鑒別，而不能以發表在某種級別的學術平台的學術成果量的多少簡而化之地加以評價。回望歷史我們亦可得到啟示。“唐宋八大家”之為“唐宋八大家”，即在於他們相較於“八大家”之外的同時代人有更多的好文章。^⑪與“八大家”大致同時而不在“八大家”之列的一些人，如劉禹錫、杜牧、范仲淹、司馬光等人，就單篇文章而論，其水平足與位列“八大家”者的文章相頡頏，但由於其相同水平的文章要少一些，便不在“八大家”班列之中。由此當可體會“量”的重要性。那種視量化評估為學術評價之毒藥的觀點，我們認為不免失之偏激。不能一味否定量化評價的作用。應該看到，在多數情況下，量化統計的結果均具有一定的參考作用。但學術成果數量多少與水平高低並非線性正相關，適當參考、重視量化評價方法和結果，不應改變“質在量先”的取向，更不應將量化評價方法奉為標準。

（二）關於同行評議

長時間以來，在職稱評定、崗位評聘、績效獎勵、學術名譽評比等與學術有關的活動中，“行政化干預”、“行政化傾向”、“外行評價內行”等現象頗受詬病，實行“同行評議”的呼聲可謂此起彼伏、不絕於耳。這種傾向就整體而論，無疑是方向正確的。由於學科、知識的領域分殊，特別是現代社會“知識爆炸”，探究程度越來越深，專業分工越來越細，那種百科全書式的“通才”已經越來越少見了，甚至是根本不可能存在了（當然，人與人之間在知識廣博、興趣廣泛程度方面的不同，還是普遍存在的），有的只是各種類型的“專家”（“專業工作者”）。考評學術成果（提供者）的開創性、影響力、社會貢獻這樣專業性較強或很強的問題，只有同領域的專家才能更加精準。

但由於歷史、文化、人性以及無人可免的認知缺陷等原因，同行評議也存在這樣那樣的問題。一是人情因素的干擾。如果評價者與被評價者在“熟”與“疏”的程度上存在區別，甚或區別不小，或者有各種程度的熟人打招呼，評價者就難免受親疏關係的影響，做出偏離真實水平、抬熟貶疏的評價。二是以觀點、取向同異判高低，對與己觀點、取向相同或相近者給予高評價，反之

則給予差評。喜同厭異，本也為人之常情，非德才兼備者不能免。恰如宋代理學家程頤所言：“人心所從，多所親愛者也。常人之情，愛之則見其是，惡之則見其非，故妻孥之言雖失而多從，所憎之言雖善為惡也。”^②清末民初的經學大師廖平也有“門戶既異，矛盾肇興”^③之語。三是評價者與被評價對象對問題的興趣和熟悉程度存在差異，難以“視界融合”，因此亦難以做出準確評價。面對同一個學術成果（文章或著作）做出截然不同甚至完全相反的評價，可謂司空見慣，中外皆然。針對專家評議中現實或可能存在的種種問題，應採取適當措施予以減少或避免。具體建議如下。

首先，在評審專家的選聘上，應該挑選真正適合充任相應評價工作的人員。作為同行評議的專家，應該德才識兼具，要有崇尚公正、容納異己的寬廣胸懷，要有泛覽精研、觸類旁通的淵博知識，要有撥雲見日、恰當精準的鑒別能力。這樣遴選專家，時常有較大難度，但亦需知難而進，盡力而為。鑒於知識領域日益廣泛、學科門類日益紛雜、專業分工日益細密，充任評價工作的專家應是真正的同行，最好是“小同行”。同時，為盡量避免“人情”等情感因素對評價工作的影響，要建立健全必要的隨機、迴避的評價專家遴選機制，讓熱衷於跑關係、託人情的被評價對象難於下手。對存在諸如親緣、師承、工作等特定關係的評價對象，要明確相應評審專家的迴避制度。

其次，要建立健全透明有效的責任擔當與監督機制。面對同行評議中較難避免的“人情”與“胸懷”問題，相對有效的辦法或許就是一些學者提到過的“記名投票”，增加評價的透明度。在讓投票者“擁有權力”的同時，適當承擔壓力，更好地發揮輿論監督與壓力的作用。要創造條件，努力探索，逐步建立、推行學術評價工作全程留痕、全程公開、全程監督的制度和措施。

再次，要建立、暢通適當的質疑、申訴和回溯機制。要允許被評價者採取適當程序對不利評價結果進行“質疑”，並能夠採取便當的途徑進行有理有據的“申訴”，甚至對評價中的嚴重不當行為進行“舉報”。評審者有義務針對這種“質疑”、“申訴”和“舉報”做出解釋。要有方便、可靠的方法，保證對評價過程可以進行必要的回溯，查核造成其不利評價結果的每一位評審人的評審結論及理由。要建立可操作的糾偏機制，對經過質疑、申訴、舉報、回溯發現的重大失當行為，予以必要的補救。

最後，要建立適當的專家責任追究機制。參與學術評價的每一位評價者均須對其評價後果“負責”。有必要大力探究、採用如何讓評價者對其評價行為負起責來的切實有效的方法。要以可靠、有效的方式建立學術評價的“不良記錄”制度。如評審者在履職過程中被確證有失職行為，當以公開透明的“不良記錄”等形式施以懲戒。對經多方確認、確有嚴重不當評價行為的專家，要給予“不良記錄”等多種形式的告誡與譴責。要採取類似於許多領域已經採取的“行業禁入”的方法，對在參與學術評價的過程中有過兩次以上“不良記錄”或其他形式的告誡或譴責的評價者，應在一定年限內甚至終身取消其參與任何形式的學術評價的資格。

綜合以上分析，可以看到，在學術評價上，一方面，我們應該堅持專家評議、同行評議的立場，相信相對於大眾評議、“外行”評議，專家、尤其是同行專家的評議更為精準。同行評議可以在很大程度上解決“外行評價內行”問題，糾偏對量化數據的不當使用（對創新性和探究深度反映模糊甚至“遲鈍”，忽視學術研究的學科與地區差異，以特定方式操控數據等），應該成為學術評價的優選方法。另一方面，也要對同行評議保持必要的“謙抑”，針對同行專家評議中存在的種種問題，多方面採取措施予以預防和糾正，盡可能保證評價工作的獨立性、公正性和評價結果的科學性、客觀性。

（三）關於分類評價

因為我們認為學科劃分與學術評價有甚大關聯，故在前面對其做了較為細緻的探討。回到本文探究的主題——學術評價上來，我們認為，對不同學科的學術成果，應有不同的評價取向，應採取、堅持分類評價的辦法：1）對於人文學科（文史哲、藝術學等），主要評價學術成果的學術水平，適當評價其影響力，淡化其短期社會效益評價；2）對於社會科學（經濟學、政治學、社會學、法學等）學科，重點評價其社會影響力和社會貢獻度，適當評價其學術水平；3）對於应用型社會科學學科（如應用經濟、市場營銷、新聞廣告、編輯出版等學科）學術成果，重點評價其社會貢獻，適當評價其學術影響，適當淡化其學術水平評價。

當然，評論任何學術成果，都有學術水平評價問題。任何學術成果，一經發表，只要有了讀者，也就會發生不同範圍、程度的影響。任何學術成果，至少在增進人類知識積累的意義上，也都有有一定的學術社會貢獻。我們上面所做的區分，只是就其優先序、相對重要程度言之，不可絕對而論。

分類評價的另一個維度是根據評價對象數量多少、領域大小而區分為微觀、中觀和宏觀評價。學術成果微觀評價主要是針對個體或人數很少的組織機構進行的學術評價；學術成果中觀評價主要是針對較小範圍的區域（地區、市縣）、較大規模的機構（一所或少量幾所大學、一個或少量幾個研究機構）、一個或少數刊物等進行的學術評價；學術成果宏觀評價則主要是針對較大範圍（國家、國家集合體、洲、全球）、領域（一級學科等）進行的學術評價。對不同的評價對象，應分別或優先採取不同的評價方法：微觀評價以質性評價為主，適當參考量化評價結果；中觀評價重視質性評價與量化評價的綜合、平衡；宏觀評價以量化評價為主，適當輔之以質性評價。

可喜的是，在與學術評價關係緊密的職稱評定、科研業績評價等領域，分類評價正得到政府、社會的廣泛認同，並以不同方式加以強調。2012年中共中央、國務院印發的《關於深化科技體制改革加快國家創新體系建設的意見》中，強調要深化科技評價和獎勵制度改革：“根據不同類型科技活動特點，注重科技創新質量和實際貢獻，制定導向明確、激勵約束並重的評價標準和方法。基礎研究以同行評價為主，特別要加強國際同行評價，着重評價成果的科學價值；應用研究由用戶和專家等相關第三方評價，着重評價目標完成情況、成果轉化情況以及技術成果的突破性和帶動性；產業化開發由市場和用戶評價，着重評價對產業發展的實質貢獻。”^⑩中共中央辦公廳、國務院辦公廳印發的《深化科技體制改革實施方案》（中辦發〔2015〕46號）強調，“對從事基礎和前沿技術研究、應用研究、成果轉化等不同活動的人員建立分類評價制度。”^⑪在2013年下發的《教育部關於深化高等學校科技評價改革的意見》（教技〔2013〕3號）中，針對“分類評價實施不到位”的現實，即強調要“建立導向明確、激勵約束並重的分類評價標準和開放評價方法”，實施科學的分類評價。“基礎研究項目要以原始創新性成果和創新性人才培養為評價重點，著重評價科學價值。應用研究項目要以關鍵技術和核心技術突破、自主知識產權成果、經濟社會效益等為評價重點，著重評價目標完成情況、成果轉化情況以及技術成果的突破性和帶動性。產業化開發項目要以技術、產品的成熟度和市場反應為評價重點，著重評價對產業發展的實質貢獻，促進科技成果資本化、產業化。”^⑫為深入落實《意見》精神，增強分類評價的可操作性，教育部科學技術委員會並制定了相對明確具體的《高等學校科技分類評價指標體系及評價要點》。“指標體系”對科技人員、創新團隊、創新平台、科技項目等對象進行了分類；採用立體化架構、二維分類方式，提供了百餘個具體指標，涵蓋創新質量、服務貢獻、科教結合、創新文化等要點；

強調“在評價實踐中，要根據同一類對象的崗位、類型、特色，選擇抽取相應的指標要點，組合形成分類指標體系，實施分類評價。”^②同樣地，在中共中央辦公廳、國務院辦公廳印發並發出通知要求各地區各部門結合實際認真貫徹落實的《關於深化職稱制度改革的意見》^③中，要求“科學分類評價專業技術人才能力素質”，強調針對範圍廣泛、層次豐富的職稱評定，應“把握不同領域、不同行業、不同層次專業技術人才特點，分類評價”；同時強調“基礎研究人才評價以同行學術評價為主，應用研究和技術開發人才評價突出市場和社會評價，哲學社會科學研究人才評價重在同行認可和社會效益”。

這些出自權威部門的關於科技與人才評價的意見和主張，雖然在針對對象、評價目的和側重方向上與本文討論的學術評價並不相同，但都與學術評價高度相關，而且與我們分類評價學術成果的主張，在理念上無疑是高度契合的。

同時也要看到，時至今日，分類評價的要求在現實生活中並沒有得到廣泛理解和執行。在與崗位聘用、職稱評定、業績獎懲、職務升遷關係密切的種種評價中，“重數量輕質量、重形式輕內容、重短期輕長遠的現象依然存在；評價指標單一化、評價標準量化、評價結果功利化等傾向沒有得到根本扭轉；分類評價實施不到位……”^④在自然科學、工程技術領域如此，在遠更複雜、多元的人文社會科學領域，對分類評價的陌生和漠視則更為普遍和嚴重。

四、附論：《舊金山宣言》、《萊頓宣言》及其中國鏡鑒

學術評價本為世界性難題，其莫衷一是之態，非為中國所獨有。當今世界，全球化、信息化引發的矛盾日益加劇，其波及廣度、影響深度更為世人所矚目。特別是在人文社會科學領域，對由評價的地域性、主觀性和社會文化差異引發的諸多問題本就難以達成共識，而量化評價之對深度和差異的漠視和忽略，更加引發了人們對據之得出的評價結果的不認同。面對由跨領域、以量化評價為主的學術評價引發的越來越嚴重的問題，國際學術界已經做出若干努力。其中最為引人注目的，當為《舊金山宣言》（*The San Francisco Declaration on Research Assessment, DORA*）和《萊頓宣言》（*Bibliometrics: The Leiden Manifesto for research metrics*）的簽署與發表。

針對資助機構、學術機構和其他團體改進科研成果評價方式的迫切需要，2012年12月美國細胞生物學學會年會期間，一群學術期刊的編輯和出版者提出了一套關於科研評價的建議，在此基礎上形成了《舊金山宣言》。2013年5月，78個科學組織的155位科學家簽署了這份宣言。《舊金山宣言》引發了世界科學界的廣泛關注，《科學》雜誌撰寫社論支持《舊金山宣言》，認為影響因子最重要的危害是可能妨礙創新，它引導科學家關注發高影響因子的文章，追逐所謂的“熱點”，而不是潛心科研創新。^⑤最新的消息是，2017年4月27日，作為擁有《自然》等知名學術期刊的出版巨頭，施普林格·自然出版集團旗下的“自然科研”正式簽署了《舊金山宣言》。

《舊金山宣言》指出，科研產物多種多樣，包括報道新知識、新數據、新反應物和新軟件的研究文章，知識產權，以及得到高度訓練的年青科學家。資助機構、聘用科學家的機構和科學家自身都有對科研產物的質量和影響進行評價的願望和需要。對科研產物予以準確度量、明智評估和周到使用乃當務之急。期刊影響因子常被用作衡量個人和機構科研產物的主要參數。由湯森·路透計算得出的影響因子，原本是用於幫助圖書館人員確定購買期刊的工具，而不是對一篇文章的科研質量的評價。關鍵是要明白期刊影響因子作為學術評價的工具存在許多廣為人知的缺陷。這些局限包括：1) 期刊中文章被引用的分佈情形是很不均衡的；2) 期刊影響因子具有專業領域

特色；它由多種多樣、高度差異的文章組成，包括原創研究文章和評述文章；3）影響因子可以為編輯策略所操縱（或“被遊戲”）；4）用以計量影響因子的數據既不透明，也非為公眾可開放獲取。

為改進評估研究產物實際品質的方式，《舊金山宣言》提出了大量建議措施。《宣言》指出，在評價研究效果的過程中，種種產物而不是研究文章的重要性在未來將會增加，但是經過同行審議的論文仍將是標示研究評價特色實質的首要研究產出。《宣言》因此建議首先還是要集中關注發表在經過同行審議的期刊中的研究文章，但是，通過對諸如數據集、重要的研究產出這樣的其他產出的認可，可以改善和擴展評價行為及其結果。

貫穿建議方案的幾個主題是：在提供資助、考慮委任或提升職務時，需要清除使用諸如影響因子這樣基於期刊的衡量標準；需要就其本身的長處而不是基於該研究發表於其中的期刊的影響等因素來對研究進行評價；需要對在線出版提供的機會（諸如放鬆對文章字數、圖表和參考文獻不必要的限制，探索關於重要性和影響力的新指標）加以利用。

聯繫到本文探討的主題，《舊金山宣言》中如下內容尤應予以關注：《宣言》重點強調了對期刊的評價與對文章、對個體科研工作者的評價的不同，不能用對期刊的評價去替代對文章和個體科研工作者的學術水平、貢獻的評價；要具體、綜合評價研究成果，重視其實質性影響，而不能簡單代之以發表其成果的出版物水平；要杜絕或減少對評價標準的不恰當操控；在採用、比較標準時，要考慮文章類型、學科領域的差異；要對採用的標準保持開放和透明。這些主張，與我們關於同行評議、分類評價的主張，顯然是高度一致或類似的。

2014年，在荷蘭萊頓召開的一次國際會議上，美國佐治亞理工學院公共政策教授迪安納·希克斯（Diana Hicks）等人提出了合理利用科學評價指標的七條原則，後來擴充為十條，並於2015年4月發表在《自然》雜誌上，此即為《萊頓宣言》。《萊頓宣言》中提出的引導科研評價的10條原則主要內容如下：1）量化評估應支持質性專家評價；2）評估方法、措施、標準應基於研究機構、團體或研究者的研究使命發揮作用；3）保護卓越的地區化研究；4）保持數據收集和分析過程公開、透明和簡單；5）允許被評估者驗證資料和分析；6）考慮發表和引用行為中的學科差異；7）對單個研究者的評價應基於對其一套申報材料的質性評判；8）避免評估指標不當的具體性和虛假的精確性；9）認識到評估和指標的系統影響；10）定期仔細檢查指標和及時更新指標。

其中特別值得注意的是第1、2、3條和第6、7條，分別強調了學術評價中質性評價的優先性和量化評價的輔助性，研究目的（使命）對評價的影響，在學術評價中應盡量考慮、照顧、保護地區、學科和個人差異。^⑩這些原則與我們前面主張的質主量輔、同行評議和分類評價無疑是高度契合的。

《萊頓宣言》和《舊金山宣言》更多的是針對自然科學和工程技術領域學術評價存在的問題，但其中包含的內容，無疑對我國當前的人文社會科學評價具有顯著參考價值。其所針對的問題，如量化評價優先、以發表成果的出版物水平評價具體科研成果的研究質量和貢獻、忽視學科或地域差異等，在我國當前人文社會科學學術評價領域廣泛存在，甚至更形嚴重。其所倡導的學術評價理念與方法，如淡化量化評價，重視質性評價，強調專業、綜合評價，弱化發表成果的出版物評價，重視學術成果的專業、地區差異等，與我們的相關理念完全契合，對解決當前中國人文社會科學領域聚訟紛紜的學術評價問題，具有很高的借鑒價值。

最後要說的是，學術評價是一個“知易行難”的問題，對於其應該如何進行，在理念上人們

是比較容易達成某種程度的共識的，而且會認為非如此不可，並無他路可尋。但在實踐操作層面，人們往往又棄理念於不顧，而按飽受詬病的操作方式進行。其間的文化、心理、社會原因似乎更值得探究。不過，理念上的撥亂反正、正本清源仍然是必要的。多數理念都難有完全實現之日，但只有有了清晰、正確的理念，點點滴滴的進步才會發生，而且一定會發生，最終收“積跬步以至千里”之效。

①就我國現行專業技術職稱評定而論，就包含工程、衛生、農業、經濟、會計、統計、翻譯、新聞出版、電、藝術、教師、科學研究等29個系列，其涉及面之廣，是其他國家或地區不多見的。而在現行評價體系中，多數系列的職稱都與“學術”高度關聯。

②參見《後現代性與公正遊戲——利奧塔訪談、書信錄》，談瀛洲譯，上海：上海人民出版社，1997年，第151頁。

③參見張暉：《科技名詞，就在你身邊——從首批科學技術新詞發佈試用說起》，北京：《光明日報》，2014年7月3日。嚴濟慈原文載1935年出版的《東方雜誌》第32卷第3號。

④該文發表於1911年6月26日出版的《國風報》第15期，署名“滄江”，後收入《飲冰室合集·文集》之二十五（下）。

⑤⑥梁啟超：《學與術》，《飲冰室合集》第3冊，文集之二十五下，北京：中華書局，1989年影印版，第11~12、12頁。

⑦嚴復：《〈原富〉按語》第58節。見王栻主編《嚴復集》第4冊，北京：中華書局，1986年，第885頁。

⑧參見《辭海》（第六版縮印本），上海：上海辭書出版社，2010年，第876頁。

⑨⑩朱紅文主編：《人文社會科學導論》，北京：教育科學出版社，2011年，第326、325頁。

⑪簡明不列顛百科全書編委會：《簡明不列顛百科全書》，北京：中國大百科全書出版社，1986年，第6卷，第121頁。

⑫Manicas, *A history and philosophy of the social science*, Oxford & New York: Basil Blackwell, p. 7. 轉引自朱紅文主編：《人文社會科學導論》，北京：教育科學出版社，2011年，第181頁。

⑬卡爾霍恩：《變革時代的社會科學》，李述一等譯，北京：社會科學文獻出版社，1989年，第46頁。

⑭⑮⑯華勒斯坦等：《開放社會科學——重建社會科

學報告書》，劉鋒譯，北京：生活·讀書·新知三聯書店，1997年，第3、31、31頁。

⑰朱紅文主編：《人文社會科學導論》，北京：教育科學出版社，2011年，第190~191頁。

⑱在接下來我們對學術評價的討論中，主要針對的是當代中國人文社會科學學術評價，有明確的時空和學科界線，這是特別需要提請注意的。後不一一說明。

⑲《論語》並非完全意義上的孔子作品，而為孔子弟子及再傳弟子關於孔子言行的記錄，內容有孔子談話、答弟子問及弟子間相與談論。現為研究孔子學說的主要材料。除此之外，“述而不作”的孔子更無其他獨立學術著述流傳於世。相傳其從事過整理《詩》《書》等古代文獻，並把魯史官所記《春秋》加以刪修，但這似乎更近今日所謂“編輯”工作，而非獨立著述。

⑳《道德經》一書是否為老子所著，歷來亦有爭論。不過一般認為書中所述，基本反映了老子的思想。

㉑或許有人認為，“八大家”之文為文學文章，以之為例說明文章“質”與“量”的關係不盡恰當。我們認為，就文章質量與影響力的關係而論，以之為例，當不致大謬。

㉒《周易程氏傳》卷二《隨傳》，見《二程集》（下冊），北京：中華書局，2004年2月第2版，第785頁。

㉓廖平：《公羊驗推補證凡例》，上海：《國粹學報》，第2年第7期，1906年8月，“學篇”，第6b頁。

㉔參見中共中央、國務院《關於深化科技體制改革加快國家創新體系建設的意見》。來源：新華網，2012年9月23日。

㉕參見中央政府門戶網站（www.gov.cn）。來源：新華社，2015年9月24日。

㉖㉗參見《教育部關於深化高等學校科技評價改革的意見》（教技〔2013〕3號），引自http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_784/201312/160920.html

⑳參見《關於發佈〈高等學校科技分類評價指標體系及評價要點〉的函》（教技委〔2014〕4號），引自教育部政府門戶網站，http://www.moe.gov.cn/s78/A16/s8213/A16_sjhj/201507/t20150703_192413.html

㉑2016年11月1日，中央全面深化改革領導小組召開中央全面深化改革領導小組第二十九次會議，審議通

過了《關於深化職稱制度改革的意見》，自2016年11月1日起實行。

㉒㉓參見嚴蔚剛：《科研評價應遵循什麼基本原則——“A類期刊”與萊頓宣言、舊金山宣言的對話》，北京：《光明日報》，2016年7月12日。

參考文獻：

- [1]馬克斯·韋伯：《學術與政治》，馮克利譯，北京：生活·讀書·新知三聯書店，1998年。
- [2]恩斯特·卡西爾：《人文科學的邏輯》，關子尹譯，上海：上海譯文出版社，2013年。
- [3]伊曼紐爾·沃勒斯坦：《反思社會科學——19世紀範式的局限》，劉琦岩、葉萌芽譯，北京：生活·讀書·新知三聯書店，2008年。
- [4]華勒斯坦等：《開放社會科學——重建社會科學報告書》，劉鋒譯，北京：生活·讀書·新知三聯書店，1997年。
- [5]華勒斯坦等：《學科·知識·權力》，劉健芝等編譯，北京：生活·讀書·新知三聯書店，1999年。
- [6]伊曼紐爾·沃勒斯坦：《知識的不確定性》，王昂等譯，濟南：山東大學出版社，2006年。
- [7]Alexander Rosenberg, *Philosophy of Social Science*. Third Edition, Boulder, CO: Westview Press, 2008.
- [8]卡爾霍恩：《變革時代的社會科學》，李述一等譯，北京：社會科學文獻出版社，1989年。
- [9]希拉·斯勞特、拉里·萊斯利：《學術資本主義》，梁驍、黎麗譯，北京：北京大學出版社，2014年。
- [10]朱紅文主編：《人文社會科學導論》，北京：教育科學出版社，2011年。
- [11]李維武：《人文科學概論》，北京：人民出版社，2007年。
- [12]袁方主編：《社會研究方法教程（重排本）》，

北京：北京大學出版社，2013年重排本。

[13]風笑天：《社會研究方法》（第四版），北京：中國人民大學出版社，2013年8月第4版。

[14]劉明：《學術評價制度批判》，武漢：長江文藝出版社，2006年。

[15]劉大椿等：《人文社會科學研究成果評價體系研究》，北京：經濟科學出版社，2009年。

[16]邱均平、譚春輝、任全斌等：《人文社會科學評價理論與實踐》（上下冊），武漢：武漢大學出版社，2012年。

[17]任全斌：《人文社會科學成果評價研究》，北京：中國社會科學出版社，2010年。

[18]李向玉主編：《總編視角——華文學術期刊發展趨勢國際研討會論文集》，北京：社會科學文獻出版社，2016年。

[19]嚴蔚剛：《科研評價應遵循什麼基本原則——“A類期刊”與萊頓宣言、舊金山宣言的對話》，北京：《光明日報》，2016年7月12日。

[20]張凱、劉英傑：《〈萊頓宣言〉的核心觀點及借鑒意義》，北京：《世界農業》，2016年第9期。

作者簡介：喻陽，《新華文摘》總編輯、編審。
北京 100706

[責任編輯 劉澤生]