

CSSCI 期刊評價與學術評價功能試識

楊九詮

1992年北京大學圖書館等單位《中文核心期刊要目總覽》首次出版，2000年中國社會科學院文獻信息中心印製出版《中國人文社會科學核心期刊要覽》，2000年南京大學中國社會科學研究評價中心主持研製的《中文社會科學引文索引（CSSCI）》光盤版正式出版發行。這也就是為學術界、期刊界和管理部門所通稱的“三大核心”。至此，中國人文社科學術期刊步入了以量化指標為主要工具的科學評價時代。與此同時，對“三大核心”的質疑從未停息。其中，尤以CSSCI最講究“科學”，影響日熾，漸成主要聚訟之源。2017年1月，《中文社會科學引文索引（CSSCI）來源期刊及集刊（2017—2018）目錄》公示，再度引起持續而激烈的爭議。

一、回到期刊評價與學術評價功能的事實

對CSSCI的紛爭不息，除了周知的事件性因素外，更具源始性的根由乃是作為“三大核心”尤其是CSSCI的理論資源、技術資源的“科學情報研究所”（ISI）的“引用索引”（SCI、SSCI和A&HCI）和“期刊引證報告”（JCR），簡言之，就是“影響因子”（IF）的功能、機制和文化。1990年代，是ISI的IF理論、技術和運作機制日趨成熟的時期，同時也是對它的質疑和批判日趨激烈的時期。CSSCI的影響力及其反作用力，同時胎自於此，具有學生關係。這可能是CSSCI始所未料的，致使日後面對質疑的解釋多了些自我辯護的沉味。

對CSSCI來說，新近幾個標誌性的事件顯得有些“流年不利”。一是2013年5月，78個科學組織的155位科學家簽署“舊金山宣言”，認為科學界應該停止使用影響因子評價科學家個人的工作。緊接著，諾貝爾獎得主蘭迪·謝克曼在《衛報》上發表文章，直言不諱地抵制三大國際期刊，引起了學術界和期刊界的世界性轟動。二是2016年7月，湯森路透公司宣佈將知識產權業務和科學信息業務以35.5億美元的價格出售給Onex Corp和霸菱亞洲投資。美國微生物學會（ASM）旋即宣佈旗下期刊不再支持影響因子。三是2017年2月26日，ISI創始人、“SCI之父”尤金·加菲爾德在美國賓夕法尼亞州去世。這一事件，總給人某種聯想和暗示。四是2017年初《中文社會科學引文索引（CSSCI）來源期刊及集刊（2017—2018）目錄》公示引發巨大爭議。

對CSSCI等機構的質疑與紛爭，基本上都是針對影響因子作為符號表徵對學術期刊評價及學術評價的真實性和準確性而來的，尤以對影響因子是否能夠用於學術評價為劇。面對質疑，各機構一方面給予了解釋，另一方面也不同程度的做出了不斷改進的表態。解釋與表態背後的動機晦暗不明，解釋可能是辯護性的，表態可能是策略性的。但質疑與紛爭也從外部推進了它們的不斷修正與完善。根據朱劍、仲偉民等學者的研究，在這一方面，“三大核心”中CSSCI用心最深、用力最勤，“近年來的改進有目共睹”。這與我的觀感也比較一致。

但是，“三大核心”在面對質疑時的自我開釋又顯得過於恭謹，在一味退避中形成了新的遮蔽，主要表現在反復強調只是期刊評價而不是學術評價，乃至CSSCI強調自己只是一個引文數據

庫，並非學術期刊的評價工具。事實上，影響因子，最早是用於圖書館制定文獻收藏計劃和經費預算，向讀者推薦優秀學術期刊。期刊評價，乃是影響因子的“幹細胞”——具有自我複製能力的多潛能細胞，可以說，影響因子各種功能要素的生長與擴張無不與此息息相關。CSSCI 等機構也許是為了息事寧人，退守到期刊評價，甚至退守到引文數據庫，遮蔽了真實存在也是合理存在的部分評價功能，更將現實中期刊界、學術界雖有質疑然又普遍用之的事實屏蔽於外。必須將期刊評價的功能與學術評價的功能，將期刊評價與學術評價的關係和盤托出，直面問題與事實，才能推進評價工作的持續改進和健康發展。

事實是什麼？事實就是以 CSSCI 為代表的“三大核心”的主要功能是中國人文社會科學學術期刊的評價，而其對學術的評價功能，借用 CSSCI 機構負責人的話說，乃是期刊評價功能“衍生出來的一種附加功能”——所以說是“借用”，是因為他以某種退守的姿態，將期刊評價與學術評價均視為檢索工具的衍生功能，似乎真的“評價非我願”；然而，我們知道檢索工具所引得的正是以評價為主要功能的影響因子，亦即檢索工具乃是影響因子的挖掘和整理的技術手段。問題是什麼？問題是：一方面作為期刊評價，尚不周全和精準，需要不斷改進和完善；另一方面，作為學術評價，除了尚不周全和精準外，更主要的是超乎其“附加功能”所不能承受之重，越俎代庖成為學術評價制度中的僭主。所以至此，原因非常複雜，不全在評價機構，甚至從某種角度看主要不在評價機構。本文主要試著討論，如何理解和實現 CSSCI 的期刊評價與學術評價功能，使之如其所是、如其所能，對期刊發展和學術繁榮發揮應有作用。

二、用經典測量理論規約期刊評價功能

從 1920 時代洛特卡定律的“以量評質”，到 1930 年代布拉德福定律的“以刊評質”，再到 1960 年代加菲爾德建立的“影響因子”與引文分析理論，期刊評價步入了高度量化的實證世代。評價，始終是它們一路走來的第一推動力。也因此，設若遮蔽評價功能，CSSCI 檢索工具挖掘和整理的影響因子，就只是“實”的數據（practical score），而不是“真”的數據（true score），難以構成符號表徵系統，通俗一點說，也就是：不能說明什麼。

CSSCI 高度量化的檢索工具，屬於經典測量行為，必須受經典測量理論（CTT）的規約，其引文分析無法繞過效度與信度、標準參照與常模參照、方差與標準差等核心概念。當“三大核心”一致強調“二八定律”時，就是在用基於方差、標準差分析的正態分佈模型。如果“二八定律”未經正態分佈模型描述，只是經驗性理解，勢必減損檢索工具所引得的數據的價值。反之，則可以從方差和標準差發現很多問題，通過誤差值的檢核對效度與信度加以校驗。效度與信度，是經典測量理論的核心概念。CSSCI 影響因子的符號表徵主要是由“（數）值”而“質（量）”而不是由“（實）質”而“（數）值”，與教育測評等顯著不同，它的效度主要是通過信度來表達的，所以信度特別重要。這也是人們質疑 CSSCI 等缺乏學科支撐的理由所在，殊不知測評本身就其理論和技術來說就具有顯著的學術性和專業性。信度係數是由測量結果的一致性和穩定性決定的。許可範圍的誤差是可以通過技術手段加以平衡的，但如果誤差主要來源於對象的含混，致使測量在不同對象之間無法保持一致而大大降低信度係數，則應該也完全可以在正態分佈模型中通過方差和標準差分析將問題揭示出來。CSSCI 已經從一級學科進一步細化到二級學科，就是保證信度的切實舉措。比如，在教育學的 37 種 C 刊中，二級學科教育技術學的 7 個刊物的影響因子顯著高於其他二級學科刊物。一方面，這反映了教育學術研究的事實；另一方面，如果只就教育學一

級學科統測，再以“值”評“質”，信度就會出現重大偏差。為什麼教育技術學刊物的影響因子那麼高呢，原因在於教育技術發展高速，加之其知識累積具有顯著的覆蓋性特徵，學科發展必須唯新是求，所以論文會大量引用最新論文的成果。綜合刊之所以難以把握，就是因為雖是綜合但各有側重，其中的學科成分非常複雜，測量結果在對象之間缺乏一致性，信度係數偏低。這次，CSSCI 公示目錄將幾家高校學報從“C 刊”降為“C 擴”，引發爭議，也許即緣乎此。這裡只是猜測，最終解釋要通過嚴格的數據描述，而不是意見之爭甚至意氣之爭。

需要說明的是，凡測量都有誤差，誤差源可能非常複雜，經典測量理論處理誤差的技術非常有限，有鑑於此，克倫巴赫等人提出了概化理論（GT），設計了一套方法去系統辨明多種誤差方差的來源。那麼，CSSCI 是否要引入概化理論呢？沒有必要。因為概化理論將情境因素收納在內，而經典測量理論則需要盡可能將情境因素作為干擾剔除在外。CSSCI 出現的一些不周全不精準的問題，完全可以在經典測量理論框架內加以解決。解決的關鍵，一是信度問題，而信度的關鍵是對象的區分與確定，如果必要，可以突破現有的學科劃分；二是效度問題，效度的關鍵是合理區分引文類別，使不同類別引文的分值通過轉換達到大致等值，再綜合不同年段以及影響因子半衰期等因素形成科學的複合性指標。

補充一點，CSSCI 也部分借助於同行評議，但基於經典測量理論的要求，同行評議作為額外的情境因素只能夠“從旁”評議；如果非得採分，同行評議的效用必須先得通過影響因子或加法或減法、或弱化或強化、或分列或組合的調整來驗證和表達，否則有可能會帶來人為的數據污染。

三、學術評價功能的限度與樣態

“以刊評文”與“以文評刊”的關係從來就存在。在經驗世界中，學者選擇自己的論文發佈在何種期刊，已就是“以刊評文”與“以文評刊”的互動。一般來說，作者總是希望好的論文發表在好的期刊上，這與將被引期刊等級考慮在內的特徵因子指標具有邏輯的一致性。所以，CSSCI 由期刊評價功能衍生出學術評價功能是很自然的事。現在的問題是，一方面 CSSCI 本身還存在不周全不精確的地方，需要進一步改進；另一方面，各評價主體在使用上存在過度與失當的地方。

首先，為行政權力所誤用。典型的是教育部學科評估，明確規定有 C 刊的論文才算成果。說誤用，是因為行政做了不該它做的事。行政權力與學術權力各有邊界，行政權力是外部的、強制性的、法律授予的，其用於學科評估，也只能限制在諸如所撥付經費使用等方面，而不應該進入學科發展規劃、學術水平和教學質量評估等學術內部事務。CSSCI 機構屬於典型的第三方評估組織，對使用者來說，具有天然的選擇性。按照其鮮明的工具性、自律性、市場性（即可選擇性），CSSCI 可以歸入哈耶克所說的“自發秩序”中，參與了學術期刊和學術研究生態的構建。也就是說，CSSCI 功能與行政權力是不兼容的。行政權力採信於它，既扭曲了行政權力，也可能耗損它的公信力，而公信力對社會組織尤為重要，是行政權威無法替代的。行政權力的徵用，是我國“影響因子崇拜”文化形成的主要根由。

其次，以一般識別特殊。期刊評價遷移到學術評價的功能，主要是在公約數上保障學術水平平均線的穩定與提升。也就是說，作為常例，有較高的適用性；但作為特例，它很可能是不適用的，尤其是對一些研究領域與研究範式極具開創性的學術成果，可能更是如此。這是由經典測量理論所決定的。只有在共性範疇，標準參照與常模參照才是“合法”的；非標準的、特殊性的，則需

要通過表現性評價去描述和分析。有些諾獎得主的論文並沒有驕人的影響因子。1963 年度物理學諾獎得主約翰內斯·詹森 1961 年 3 篇論文的 SCI 引用僅有 4 次。2002 年化學諾獎得主田中耕一的獲獎成果是一篇會議論文，由於未被 SCI 收錄沒有任何影響因子。用特例可以警示我們影響因子的限度；另一方面，用特例來否定常例，也是不妥當的。也要看到，特例會轉化為常例，這個過程一般會伴隨著影響因子的遞增；如果使用得當，CSSCI 也可以預測和推進新領域和新範式的形成。比如，愛因斯坦相對論方面的論文發表後缺少反響，但多年以後逐漸出現了引用高峰。

第三，用技術性評價替代實質性評價。正如前文所示，CSSCI 影響因子的符號表徵主要是由“值”而“質”，屬於技術性評價，其表達方式是高度形式化的數學符號表徵系統。穩定性與通用性，是技術性評價的優勢所在。但是，學術研究的主體與對象、過程與結果，充滿了複雜性和不確定性；而學術研究也正因其複雜性自具了豐富性的特質，正因其不確定性自具了創新性的可能。過於依賴技術性評價，一個標準評天下，必然損傷學術研究的豐富性和創新性。形式化的技術評價之外，還需要基於學科內容質量的實質性評價。學術評價中的“代表作制度”，就屬於實質性評價。我們注意到，一些機構在職稱評審中，參照 CSSCI 等評價機構的指標制定參評標準，並要求申請者自薦幾篇代表作並對代表作在學術、政策、實踐諸方面的重要性、創新性、應用性和影響力等加以陳述，評審專家綜合全部申報材料進行評議和評判，這一過程一定程度上兼顧了技術性評價與實質性評價。但實質性評價往往受制於條件和成本，難以經常性和大規模運用，尤其是對外部評價來說。也因此，人們往往偏好於用技術性評價的簡易來替代實質性評價的變易，致使學術評價越來越偏離學術研究的本相。對此，我們應該記取澳大利亞教育學者科林·馬什的警示：“真實性源於評價什麼是最重要的，而非評價什麼是最方便的。”

第四，用外部評價替代內部評價。CSSCI 對大學和科研機構來說，是外部評價。學術研究是高深知識運行和發生的行為，具有高度的專業性和自主性。學術自治，是學術健康發展的重要保障。對學術研究來說，基於學術共同體的內部評價更為重要，其本身就是研究人員學術生活和學術活動的組成部分。內部評價更多使用實質性評價，但並不同於實質性評價。小微規模的外部評價有時也採用實質性評價的方法。內部評價是自我為了自我通過自我對自我的評價。內部評價的良性機制是專業、良知、真誠友好的人際關係，其功能和目的是改進而不是甄別，在過程而不在結果。CSSCI 用於內部評價，不應該是直接的，更不應該是全部的，其權重應該是係數性的——一定程度地使用，其姿態應該是參數性的——參考與提示。以外部評價替代內部評價，將造成組織內部學術活動的分裂與衝突、工具化與功利化，造成學術生產基因異化與衰變。作為大學和科研機構不宜用公約數特徵的 CSSCI 來規約組織和成員的發展。有不少機構都有自己的核心期刊目錄，這是好的現象。這些目錄，往往在參考某個或多個學術期刊評價指標的基礎上有所增益。儘管有些目錄不盡合理，但作為內部事務，還是要從組織發展願景、科研規劃和人才培養要求出發，通過內部民主的方式加以調適。這還遠遠不夠。將 CSSCI 帶入內部評價，最適宜的途徑是將衍生的學術評價功能轉換為學術評論功能，借助 CSSCI 數據庫，進行為我所用的引文分析，對影響因子重新賦能。比如通過被引的數量與質量，考察組織內部學術生產的代際狀況，優化人才結構，調整考核策略；通過引文流變線索和分佈狀況，勾畫知識譜系、繪製知識圖譜，制訂和調整學術發展規劃；通過引文範圍的異動趨向，發現和追蹤交叉學科和新興學科；等等。這一方面，CSSCI 應該可以提供全方位、定制化的知識管理和學術服務等業務。

（作者係《華東師範大學學報（教育科學版）》主編，編審） [責任編輯 劉澤生]