学者合作网络外部影响力与学术成就的关系研究 *

闫晓慧[1,3] 杨杰[2,3] 杨文霞[2,3] 王昊[2,3] 邓三鸿[2,3]

[1] 山西警察学院侦查系 太原 030401

[2] 南京大学信息管理学院 南京 210023

[3] 数据工程与知识服务省高校重点实验室（南京大学） 南京 210023

摘 要：[目的/意义]传统计量指标在学者影响力评价时未考虑学术合作网络内外部的差异，本文提出合作网络外部被引量和外部h指数等学者合作网络外部影响力评价指标，为学者真实学术影响力评价提供依据。[方法/过程]以美国物理学会数据集的463 348篇研究论文及其234 086位消歧学者的数据作为样本，从全局合作网络抽取学者的自我中心合作网络，基于学者合作网络对内部与外部进行区分，提出外部被引量和外部h指数等学者合作网络外部影响力评价指标，结合基准计量指标进行一致性检验，基于诺贝尔奖、沃尔夫奖和狄拉克奖等物理学奖项数据证明合作网络外部影响力评价指标的有效性和优越性。[结果/结论]研究发现，学者合作网络内部与外部影响力分布存在显著差异，本文提出的学者合作网络外部影响力与基准指标一致性较强，收敛有效性分析表明外部被引量和外部h指数相比于基准指标能够更准确地识别获奖学者，表明合作网络外部影响力反映了学者的真实影响力，在学术成就中扮演着关键作用，本文提出的新指标对于构建未来人才评价体系和破“五唯”有着启示作用。

关键词：科学计量 人才评价 合作网络 外部影响力

分类号：G301 G25

CSTR：32305.14.CN11-1541.2025.02.009

DOI：10.13266/j.issn.0252-3116.2025.02.009

引用本文：闫晓慧，杨杰，杨文霞，等．学者合作网络外部影响力与学术成就的关系研究[J]. 图书情报工作，2025, 69(2): 96-107. (Citation: Yan Xiaohui, Yang Alex Jie, Yang Wenxia, et al. A Study on the Relationship Between Scholars' Collaboration External Impact and Academic Success[J]. Library and Information Service, 2025, 69(2): 96-107.)

1 引言 / Introduction

当前科研生态中，面对日益复杂化的需求和逐渐模糊的科技边界[1]，学者之间的合作变得越来越普遍[2]。科研合作已经逐渐成为科学研究的一种主要范式[3]，为促进科学的交流和多元发展，学者合作网络随之形成[4]。由此催生了合作网络研究、跨学科合作研究等诸多新兴研究领域。

近年来，大量研究针对学者合作网络的模式[5]、跨学科和跨组织性[6]、研究主题分布[7]、性别与种族分布[8]以及成员年龄[9]等要素展开了广泛而全面的讨论。然而，随着学者合作网络规模越来越大，目前对于个体学者的影响力在其合作网络内外部的分布则缺乏深入的探索、分析与表征方法，这在一定程度上阻碍了学界对科研合作和学者真实影响力的深刻认知。

为此，本文尝试基于学者合作网络提出一种识别科研合作网络的有效方法，用于深入剖析科研合作网络的内外部被引量现象。首先，以学术论文的共同作者为切入点，基于合作网络抽取出学者的自我中心合作网络，剖析学者合作网络的内外部被引量现象。接着，在考虑合作网络内外部被引量的基础上，提出合作网络外部被引量与外部h指数，以刻画学者的外部影响力。然后，基于美国物理学会的文献及其引文数据进行实证研究，讨论指标之间的一致性和收敛有效性。

2 相关工作概述 / Literature review

2.1 学者合作网络

近年来，学者之间的合作可以产生更多创新的想

*本文系国家社会科学基金项目“大数据环境下学术成果真实价值与影响的实时预测及长期评价研究”（项目编号：19BTQ062）研究成果之一。

作者简介：闫晓慧，讲师，博士；杨杰，博士研究生；杨文霞，博士研究生；王昊，教授，博士生导师；邓三鸿，教授，博士生导师，通信作者，E-mail：sanhong@nju.edu.cn。

收稿日期：2024-01-26 修回日期：2024-05-24 本文起止页码：96-107

版权所有©《图书情报工作》杂志社有限公司，未经许可不得转载（Copyrights © LIS Press Co., Ltd. Reproduction is prohibited without permission）

中国知网 https://www.cnki.net

法[10]、产生更高质量的研究成果[11]，更能够有效解决问题，因此学者之间的团队合作在科学研究中的重要性越来越明显。然而，有效的团队合作需要清晰的沟通[12]，利益冲突、权力动态和文化差异等都会阻碍合作进程，导致合作成果欠佳[13]。学者合作网络的组成和结构在研究成果质量等方面起着至关重要的作用[14]，一方面，更大和更多样化的学者合作网络产生的科研成果更具创新性和颠覆性[15-16]；另一方面，学者合作网络规模的增加也会带来合作摩擦和研究成果回报递减等消极影响[11]。尽管，近年来产生了大量关于学者合作网络与影响力的研究，然而，很少有研究区分学者合作网络内部和外部的影响力情况，并且对学者影响力在内部和外部的分布等情况探讨更少[17]。

学者合作网络与无形学院（invisible college）的概念类似，即是一种非正式的、具备较强凝聚力的“科学共同体”。A. Perianes-rodríguez 等[18]认为，在大部分学科中，学者合作网络是科研生态中的一个基本单元。合作网络中的学者通常会分享科研基金和科研材料，共同合作完成科研成果，需要注意的是，他们并不一定来自于同一机构。合作网络的学者之间通常相互熟悉、互相了解，久而久之形成一个个较为封闭的情理社会[19]。

先前，学者合作网络的识别是一项复杂而艰巨的任务，通过问卷调查、实地调研和访谈等方法能够对学者合作网络进行定性识别，但这种方式通常效率不高，比较难以广泛开展[20]。近年来，随着复杂网络的发展与推动[21]，基于科研成果合作情况的学者合作网络识别得到了充分发展，通常从共同署名的合著成果中识别学者的合作网络[22]。以合作网络中节点或边的相关指标为基础的社群检测方法已经成为识别学者合作网络的主流方法之一[23]。例如，Z. Hu 等[24]从论文层次的引证关系出发识别学者合作网络；B. Cai 等[25]提出了基于网络的节点密度、中心度和 K-means 区分不同的科研合作网络的方法；C. Carusi 等[26]通过计算发表于相同期刊上的学者相似度来识别出学者合作网络；H. Han 等[27]将合作网络与作者机构相结合，基于主题相似度方法挖掘出学术合作网络。

2.2 学者影响力

长期以来，学者影响力的评估一直是科学领域的一个重要问题，涉及到人才招聘、项目评估[28]、任期和绩效评估[29]以及奖励制度[30]等各个方面[31]。对学者及其影响力的准确评估不仅对学者的学术生涯发展至关重要，而且对整个科学领域的发展起着不可或缺的作用。正确的学者评价机制能够让政府和机构将更好地将资源和资金分配给在领域内做出重要贡献的学者[32]，以更好地促进科学发展。

学者影响力的测度一般是基于被引量、h 指数等指标[33]，学术论文之间的引用作为科学继承与进步的一种基本表现方式[34]，反映了“科学共同体”之间的学术认同和科技知识的流动与发展[35]。同时，引用存在多种模式，包括学者的内部引用、学者对其合作网络成员的引用、学者对于非合作网络学者的引用等。除了内部引用以外，目前对于另外两种引用模式尚未有明确的界定与研究。传统的定量指标，如发文量、总被引和 h 指数[36]等广泛用于评估学者的影响力。然而，这些指标不能完全表征学者真实的影响力[37]，尤其是无法精准测度学者在合作网络内部和外部的影响力。对于学者合作网络内外部的影响力的准确评估需要更加复杂和可靠的方法。

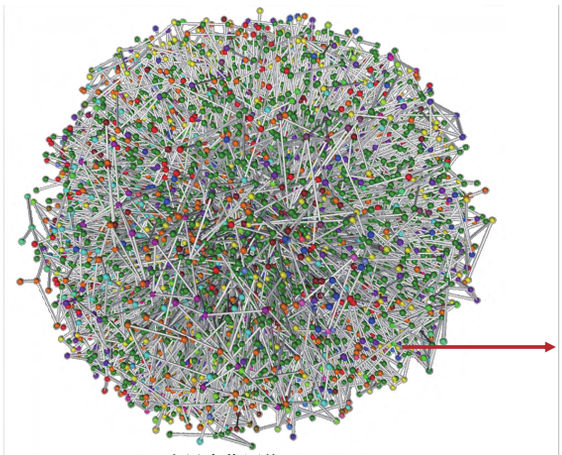
综上所述，当前的学者影响力评价指标并没有考虑影响力的来源，即，没有对学者合作网络内外部的影响力进行具体区分，缺乏最基本的学者合作网络外部影响力的定量表征方式，在一定程度上忽略了学者在合作网络内外部的情况的区别。基于此，需要开发更加全面的指标，用以深入分析学者影响力在其合作网络内外部的分布情况，来准确评估学者合作网络内外部的影响力及其科研成果的影响，为政策制定提供新的依据。

3 学者合作网络判定及影响力计算方法 / Methods for determining the collaborative network of scholars and calculating their influence

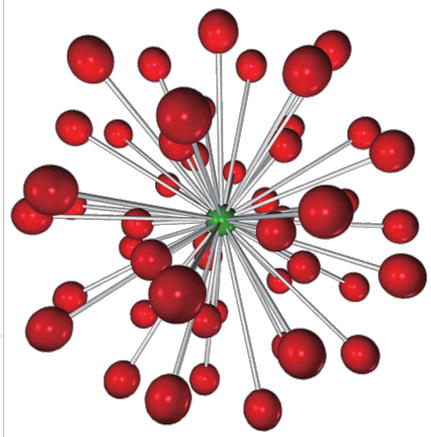
在一位学者漫长的科研旅程中，会有许多共同署名发表论文的合作者，但是他们未必都属于该学者的合作网络。按照 P. FONTANAROSA 等[38]和 L. HEIDI[39]等的定义，在学者合作网络中，各个学者之间是相互较为熟悉的，彼此是经验丰富的合作者，并且学者各自的自我中心合作网络应当展现出相似的拓扑特征。将所有学者共同存在的合作网络定义为全局合作网络 $G=(V, E)$，在考虑单个学者所在的科研合作网络时，从全局合作网络 G 中抽取出自我中心网络（ego-centric network）。在以自我中心网络为基础的合作网络中，所有节点都会链接到中心节点，各个节点与中心节点的边的权重表示他们的合作次数。如图 1（a）所示，考虑核心学者 S 的自我中心合作

中国知网 https://www.cnki.net

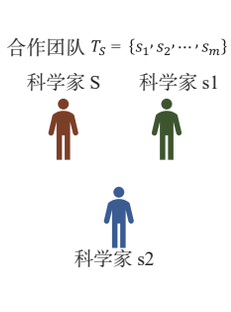
网络，将学者 S 的自我中心合作网络定义为 $G_S=(V, E)$，该网络中的节点集合 $V=\{v_1, v_2, \cdots, v_m\}$，边集合 $E=\{e_{i,j}, \cdots, e_{m,m} \mid i, j \leq m\}$。构建邻接矩阵 $M=(a_{i,j})_{m\times m}$，由于学者合作网络是一种有权重无向网络，因此 $a_{i,j}=a_{j,i}$。



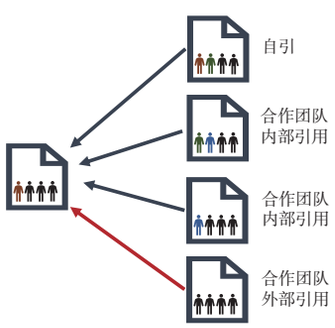
(a) 全局合作网络 $G=(V,E)$



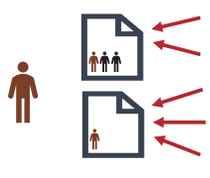
(b) 自我中心网络 ego network



(c)



(d)



$Eh(S)=max\{i \mid i \leq ECC(x_i)\}$

$EC(S)=\left|\{x \mid x \notin X_{T_S},\ x \in X_A,\ |a_{x,X_S}| \geq 1\}\right|$

(e)

图 1　学者合作网络及外部被引量示例

Figure 1　Illustration of the global collaborative network, ego-centric network, and external impact metrics

注：（a）APS 数据集中的全球学者合作网络，包含 20 多万个节点；（b）诺贝尔奖获得者杨振宁学者在 APS 数据集中的自我网络；（c）学者 S 的合作网络；（d）自引、网络内部被引量和外部被引量的说明；（e）$Eh(S)$、$EC(S)$ 和 $IC(S)$ 的计算说明

如图 1（c）—图 1（e）所示，考虑某学科领域的全部论文集合 $X_A=\{x_1, x_2, \cdots, x_n\}$；学者 S 署名的全部论文集合 $X_S \subseteq X_A$；学者 S 所在的合作网络 $T_S=\{s_1, s_2, \cdots, s_m\}$ 及该合作网络所有学者署名的论文集合 $X_{T_S} \subseteq X_T$，则学者 S 的网络内部被引量（Intra-team Citation, 以下简称为 IC）和外部被引量（External Citation, 以下简称为 EC）的计算方法分别如公式（1）和公式（2）所示：

$$IC(S)=|\{x|x \in X_{T_S}, x \in X_A, |a_{x,X_S}| \geq 1\}| \quad \text{公式（1）}$$

$$EC(S)=|\{x|x \notin X_{T_S}, x \in X_A, |a_{x,X_S}| \geq 1\}| \quad \text{公式（2）}$$

在此，若文章 x_i 引用了 S 的论文，即引文网络中的邻接向量 a_{x,X_S} 模大于等于 1，但如果 x_i 的作者属于 S 的合作网络，即 $x_i \in X_{T_S}$，则认为 x_i 不属于合作网络外引用；反之，可以认为其属于合作网络外部引用。合作网络外部学者引用学者 S 的全部论文的总数量即为学者 S 的外部被引量。

为平衡合作网络外引文量与发文量之间的关系[40]，本文进一步提出外部 h 指数（external h-index，以下简称为 Eh），考虑某学者 S 的全部论文，将其按照论文的外部被引量 (paper-level external citation，以下简称为 PEC) 降序排列，如公式（3）所示：

$$X_S=\{x_1, x_2, \cdots, x_r|PEC(x_1)<PEC(x_2)<\cdots<PEC(x_r)\} \quad \text{公式（3）}$$

则根据 X_S 中至多有 Eh 篇论文被引量大于或等于 Eh，计算出学者 S 的外部 h 指数，如公式（4）所示：

$$Eh(S)=max\{i|i \leq EPC(x_i)\} \quad \text{公式（4）}$$

4　数据与处理 / Data and preprocessing

4.1　数据来源

为保证引用数据的权威性和可靠性，本文采用美国物理学会（American Physical Society，APS）提供的文献及其引文数据。APS 数据集共包括 19 种期刊，涵盖了 1893 年至 2009 年的所有文献和引文数据，一共收集到 463 348 篇学术论文。具体的论文及被引情况见图 2。接着本文使用 igraph-0.10.2 构建大型引文网络有助于建立用于分析的大型邻接矩阵，同时使用 Python-scipy-1.10.1 等程序计算和评估相关指标。

中国知网 https://www.cnki.net

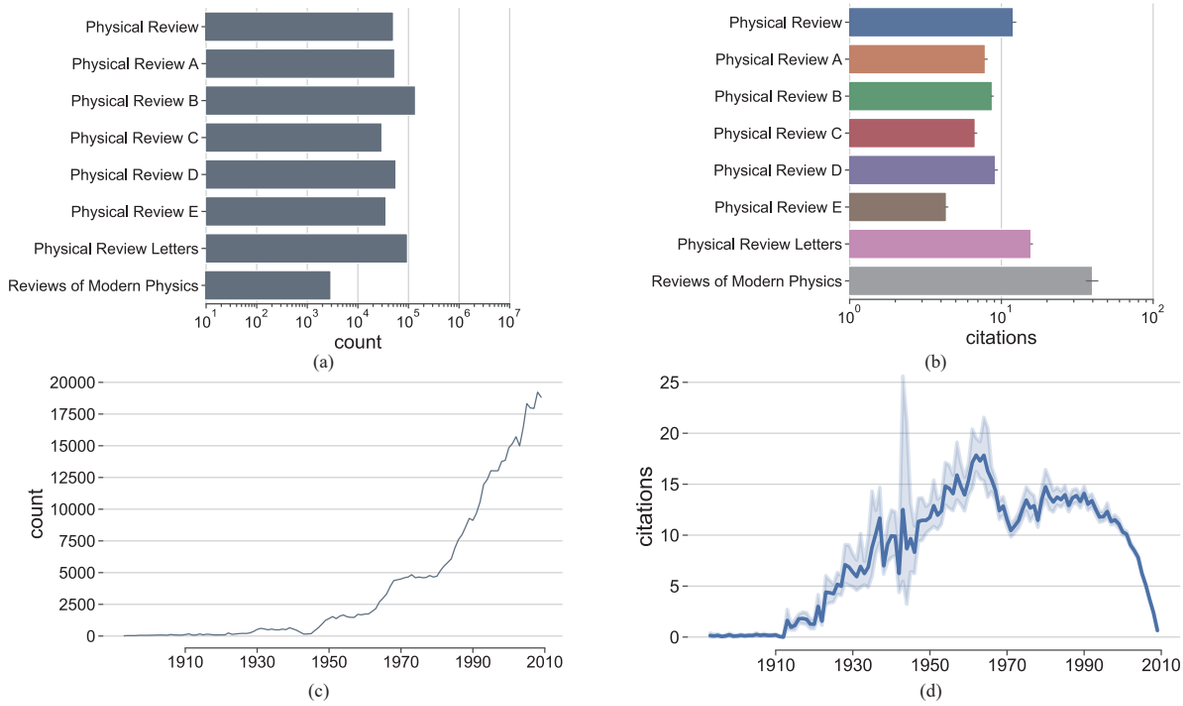


图 2　APS 数据集情况

Figure 2　APS dataset description

4.2　学者信息消歧

在进行大型数据集处理时，消除姓名歧义是定量测量和评估学者的关键且具有挑战性的步骤之一[41]，同一学者可能以不同的名字出现，或者不同的学者可能具有相同的名字。为了解决这个问题，本文遵循了 R. Sinatra 等[42]的方法完成学者消歧工作。

具体来说，本文从 APS 数据集中提取了所有的学者姓名对，并使用正则表达式根据姓氏、缩写和中间名相似性以及引文和机构隶属关系进行匹配。为了进一步消除名称不明确的学者的歧义，本文同时比较了物理树（physic tree）数据集。最终，本文确定了 234 086 名不同的学者，本文根据 D. Hicks 等[43]的方法，为所有论文的学者赋予分值。

4.3　获奖学者数据

为了评估本文所提出的指标在识别杰出物理学家方面的有效性，本文收集了多个著名国际奖项的获奖学者的数据，包括诺贝尔奖（Nobel Prize）、沃尔夫奖（Wolf Prize）和狄拉克奖章（Dirac Medal）。这些奖项被公认为科学界的最高荣誉，反映了该学科最高水平的创新和重大贡献[44]。

APS 中消除歧义的学者姓名通过使用各种来源与这些获奖学者进行匹配，具体包括获奖学者的主页、维基百科和物理树传记信息。为了确保高水平的学术活动，本文只考虑在 APS 数据集中发表至少 5 篇论文的学者，最终确定 130 位获奖学者作为评估拟议指标识别成功和高影响力的学者的基准。

5　结论与讨论 / Conclusion and discussion

5.1　描述性统计

首先，本文对学者的合作网络内外部影响力情况进行分析，主要集中在学者的网络外部被引量和 h 指数两个方面，具体见表 1。学者的合作网络外部平均引用量为 46.3 次，约占总引用量的 75.6%。这表明，学者合作网络外部的被引量比网络内部的被引量对于学者的影响力的贡献大很多。

表 1　学者合作网络内外部影响指标的描述性统计

Table 1　Descriptive statistics of internal and external impact indicators of scholar collaboration networks

指标	样本数量	平均值	标准差	最小值	中间值	最大值
成果数量	234 086	5.9	12.5	1	2	491
Citations	234 086	61.2	234.0	0	9	159 06
EC	234 086	46.3	182.6	0	6	149 67
h-index	234 086	2.5	3.2	0	1	58
Eh	234 086	2.0	2.5	0	1	52

J. Chu 等研究发现，除去合作网络后的学者评价更能够被普遍接受[45]，这一发现对于学者评价研究具有一定的积极意义。相比于传统的引文计算方法，本文构建的指标 EC 和 Eh 能够简单直接地测度学者的合作网络外部影响力，强调了学者在合作网络外部被认可的重要性。此外，从表 1 还可以看出，学者的 h 指数和网络外部 h 指数之间的平均值差异为学者

中国知网　https://www.cnki.net

h 指数的 20%。这个差异比总引用次数的差异要小，表明网络外部 h 指数的能够平衡引用和成果数量之间的关系，为全面评价学者影响力提供新的视角。

为了更加细致地了解学者产出和影响力之间的分布情况，采用 Python 中的 PowerLaw 库 [46] 来拟合每个指标的幂律函数 $N(x)\sim x^{-\alpha}$，具体情况见图 3。从图 3 可以看出，这些指标具有很强的幂律分布特征，显示了学者中存在着论文和影响力的显著异质性。这种异质性可能和学者研究领域、研究经历和资源可获得性有关系，进一步体现了马太效应 [47]、科学界的竞争性和发展不平衡性 [44]。这些指标的幂律分布表示，在本文选择的学者中，少数学者的学术影响力远远超过其比例，而大多数则具有相对较低的影响力。

学者的总被引量（图 3（b））和合作网络外部被引量（图 3（e））之间以及 h 指数（图 3（c））和合作网络外部 h 指数（图 3（f））之间的分布和幂律拟合参数 α 之间存在显著的相似性。这一观察结果表明总引用和合作网络外部被引量之间、h 指数和外部 h 指数之间存在强有力的相关性。这种相关性表示合作网络外部被引量在确定学者整体学术影响力方面具有显著的权重。但需要注意的是，网络外部被引量（EC）和网络内部被引量（IC）的分布存在着明显的差异，相较于网络外部被引量，拥有较低的网络内部被引量的学者的比例比较高。在学者影响力评价中，网络外部被引量（EC）和网络外部 h 指数（Eh）可以作为有效的补充指标，来更加全面地反映学者的学术影响力。本文提出的网络外部被引量（EC）和网络外部 h 指数（Eh）比较擅长捕捉学者合作网络外部的影响力，可以为学者评价研究提供更加全面的视角。

这一结果凸显了学者之间的异质性，网络外部被引量在决定整体学术影响力方面发挥着重要作用。本文所提出的指标 EC 和 Eh 为传统基于引用的指标提供了一种补充视角，有助于更细致地理解学者在其合作网络内外部的影响。

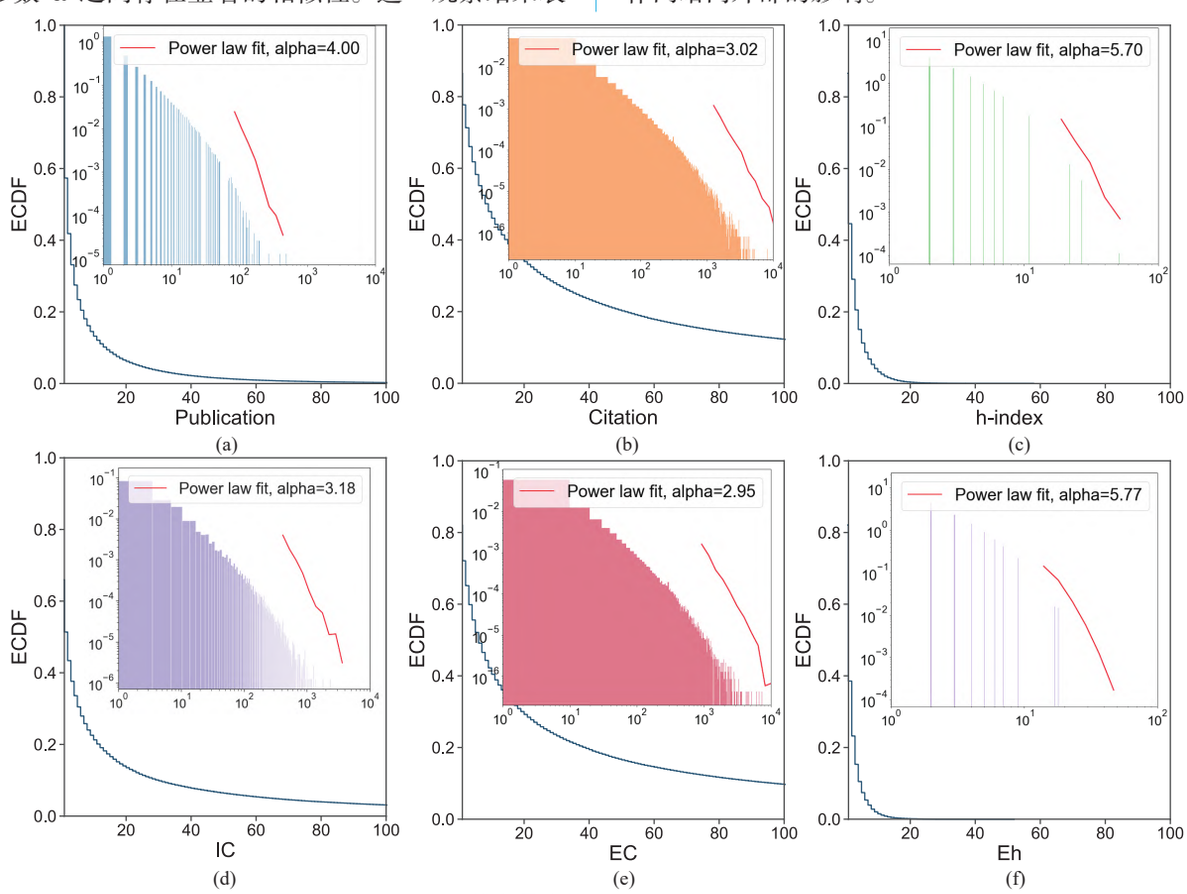


图 3 （ECDF）曲线和具有拟合 Alpha 的学术指标的幂律分布

Figure 3 Empirical cumulative distribution function (ECDF) curves and power-law distribution with fitting Alpha for academic metrics

5.2 一致性检验

如图 4 所示，通过计算网络外部指标与基准指标的经验分布函数（empirical cumulative distribution function，ECDF）和核密度估计（kernel density estimation，KDE），本文发现二者具有相似的分布，这表明二者在分析和评价过程中有较强的一致性。

中国知网 https://www.cnki.net

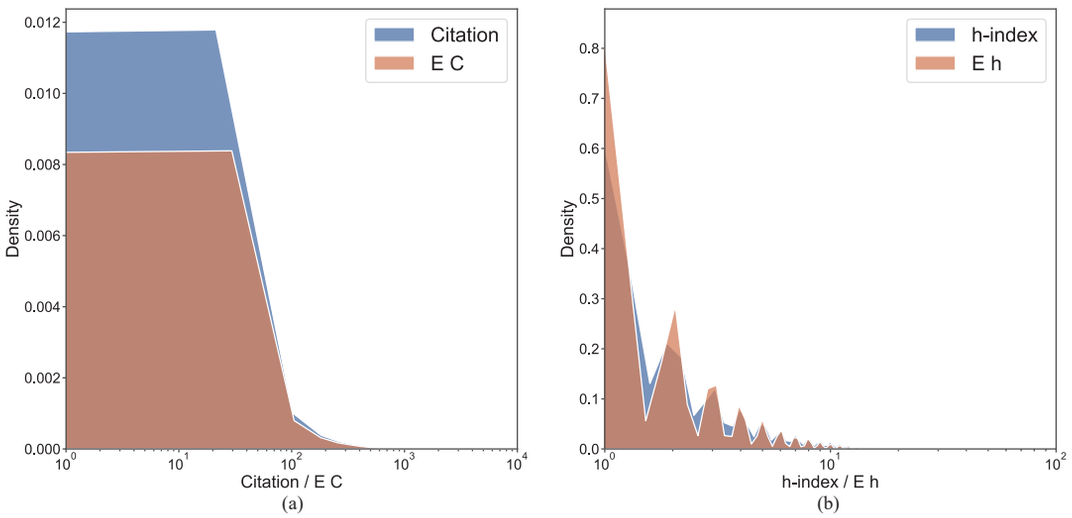


图 4　核密度估计（KDE）密度

Figure 4　Kernel density estimation (KDE) density plot

本文进一步选取肯德尔秩系数（Kendall's tau）作为指标测度一致性的衡量标准，肯德尔秩系数可用于离散数据的相关性分析，适用于本文的数据，这个部分主要检验用于评估学者影响力的不同指标之间的一致性[48]。具体而言，本文选取肯德尔秩系数来确定本文中新提出的指标（即外部被引量（EC）、外部 h 指数（Eh）和内部被引量（IC））与传统基准指标（如总引用、总发表量和 h 指数）之间的相关性。在图 5 中，不论是获奖学者还是非获奖学者的被引量和 h 指数都存在着很强的相关性，究其原因，主要是 h 指数的计算规则就兼顾了成果数量和被引情况。同时，同获奖学者比较，非获奖学者的被引量和网络外部被引量之间的相关性要更强，同样 h 指数和网络外部 h 指数之间的相关性也是非获奖学者要更加明显一些。此外，还可以观察到在两组对比中，网络内部被引量（IC）和被引量和 h 指数都是中等程度的相关，然而同网络外部被引量的相关性就弱很多，故可以得出，虽然合作网络内部的引用很重要，但是与学者总影响力的相关性没有网络外部引用强。

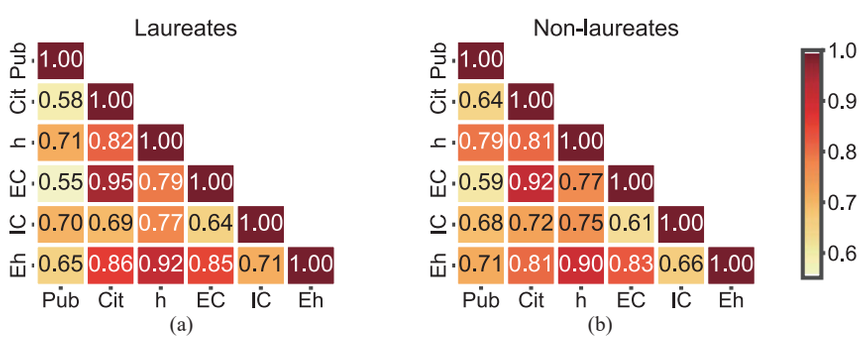


图 5　诺贝尔奖获奖学者和非获奖学者在每个指标上的肯德尔秩系数

Figure 5　Kendall's tau correlation coefficients for Laureates and non-Laureates across each metric

5.3　学术成就

学术成就是一个很复杂的概念，有多种方式进行评价。在本节中，将物理学领域的诺贝尔奖、沃尔夫奖和狄拉克奖等著名奖项的获奖学者作为评估所提出的指标——外部被引量（EC）、外部 h 指数（Eh）和内部被引量（IC）的基础，以理解影响学者影响力和合作网络内外部之间的关系。表 2 和表 3 呈现了具有最高 EC 和 IC 的前 5 位学者之间的显著差异。结果显示，具有最高 EC 的学者引用计数明显高于具有最高 IC 的学者。例如，具有最高 EC 的顶级学者 Walter Kohn 的总引用次数为 15 906，几乎是具有最高 IC 的学者 Manuel Cardona（8 810）的一倍。这些发现表明，合作网络外部被引量可能是决定学者整体学术影响力的一个关键因素，而合作网络内部被引量则与学者真实影响力的关系较弱。

具有最高外部被引量（EC）的学者都是诺贝尔奖获得者或获得过狄拉克奖章和沃尔夫奖等著名奖项的人，是这些学者卓越的学术影响力和认可的重

中国知网 https://www.cnki.net

要表现之一。相反，具有最高 IC 的学者却较少获得过这类学术荣誉。这意味着，虽然学术网络内部合作和引用对科学研究至关重要，但它们可能对整体学术成就和认可的影响不如学术网络外部的影响和引用显著。

表 2　EC 值排名前五学者的情况

Table 2　Top 5 scientists with the highest external citation (EC)

学者	Pub	h	Cit	EC	IC	E h	Prize
Walter Kohn	115	36	15 906	14 967	939	31	Nobel
Philip W. Anderson	138	55	13 920	12 946	974	48	Nobel
Steven Weinberg	133	58	12 435	11 562	873	52	Dirac
Patrick A. Lee	203	57	13 564	11 186	2 378	46	Dirac
Bertrand I. Halperin	185	56	12 470	10 398	2 072	45	Wolf

表 3　IC 值排名前五学者的情况

Table 3　Top 5 scientists with the highest intra-team citation (IC)

学者	Pub	h	Cit	EC	IC	E h	Prize
Manuel Cardona	491	41	8 810	4 704	4 106	29	none
Gen Shirane	385	53	10 494	6 602	3 892	33	none
Yoshinori Tokura	348	46	8 794	5 114	3 680	31	none
Robert J. Birgeneau	256	47	8 260	5 225	3 035	31	none
Allan H. Macdonald	307	44	7 846	4 815	3 031	30	Wolf

接着用本文提出的两个新型指标，即网络外部被引量（EC）和外部 h 指数（Eh），比较了获奖学者和非获奖学者的外部影响。图 6 中的结果揭示了影响学术成就和认可的因素，并显示了获奖学者和非获奖学者的外部影响存在显著差异，其中对于非获奖学者而言，外部被引量（EC）更能成为整体学术影响的良好指标。这一观察结果表明，非获奖学者可能更加依赖外部合作和资源来实现学术成就和认可。此外，还对获奖学者和非获奖学者之间的 EC 和被引量分布进行了考察，发现非获奖学者具有更高比例的网络内部被引量。这种差异可能是由非获奖学者对外部资源和合作机会的获取有限造成的。另一方面，获奖学者可能更容易获得网络外部的影响和资源，从而导致外部影响更加均衡分布。

所提出的指标 EC 和 Eh 为学者的合作动态和学术影响力提供了更全面的视角，增强了对科学界内研究影响和网络动态的理解。通过引入网络外部被引量的维度，可以加深对科学界内研究影响和网络动态的理解。此外，本文的发现强调了在评估学者的学术影响力时考虑网络外部的影响和引用的重要性，特别是对于可得资源有限的非获奖学者而言。

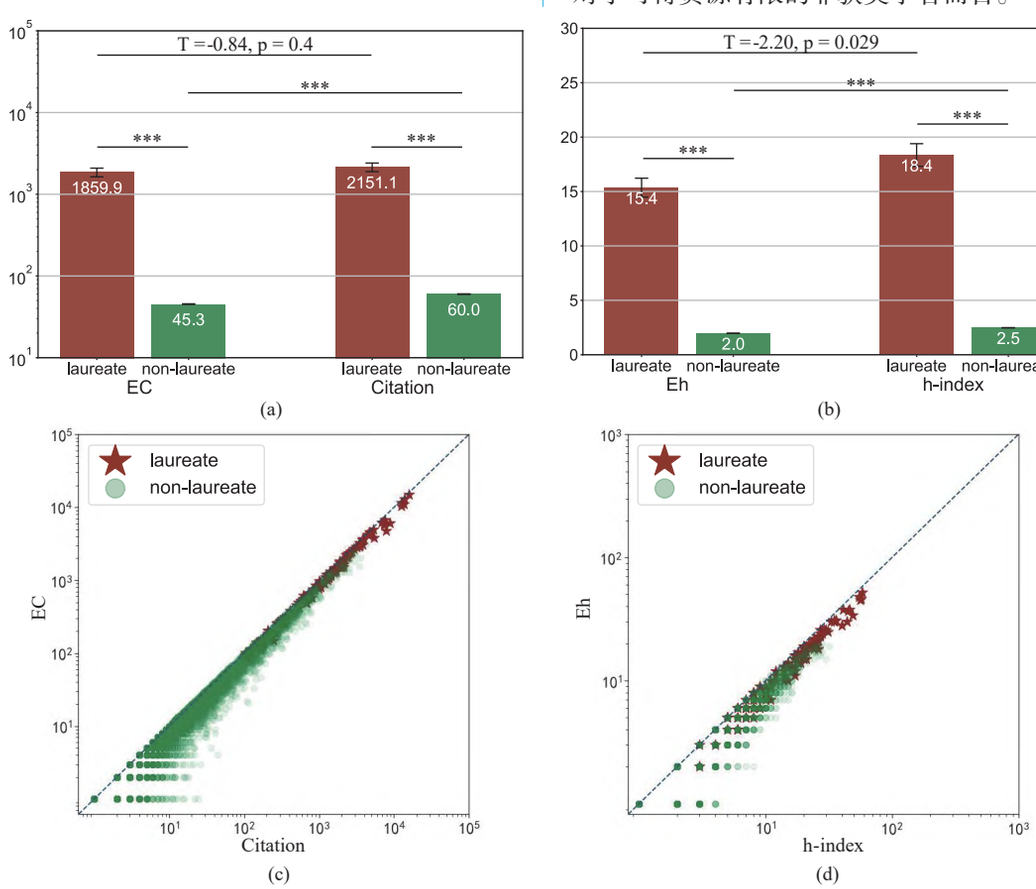


图 6　获奖学者和非获奖学者的网络外部被引量和网络外部 h 指数比较

Figure 6　Comparison of external citation and external h-index of Laureates and non-Laureates

中国知网 https://www.cnki.net

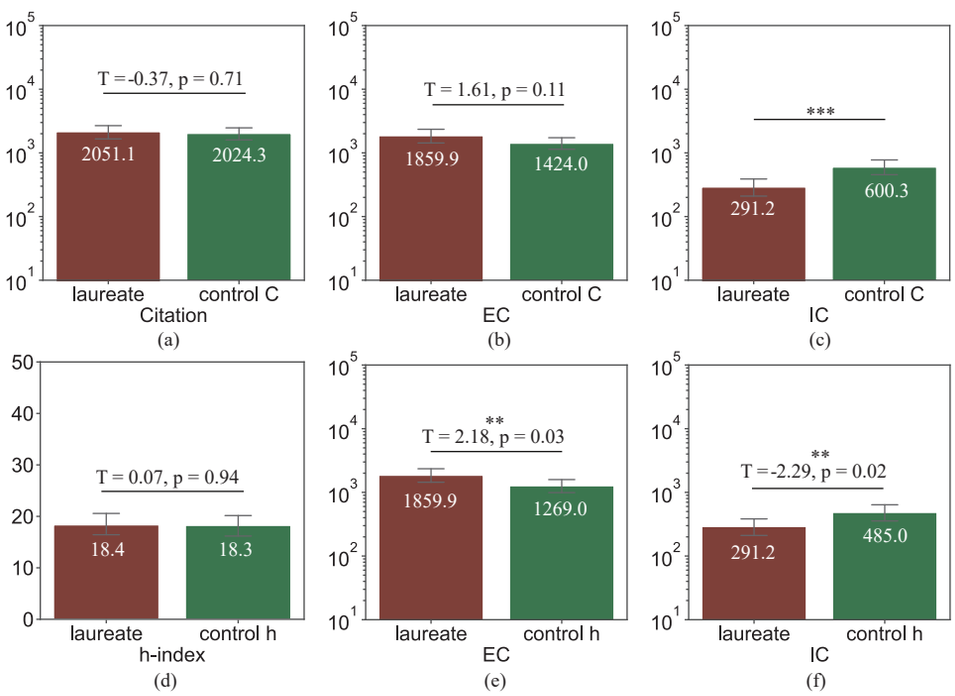
总的来说，本文使用所提出的指标对获奖学者和非获奖学者进行比较，突显了网络的外部影响和引用在确定学者整体学术影响和认可方面的重要性。这些发现为深入了解学术成就的因素提供了一定的见解，并为传统基于引用的指标提供了互补的衡量方式，有助于更细致地理解学者在合作网络内外部的影响。

5.4 收敛效度

为了进一步验证指标的科学性和有效性，本文对网络外部影响力指标和学者学术成就的收敛效度进行分析。指标的收敛效度指的是指标对目标对象的范围、大小和效率的有效衡量能力[49]，换言之，确定本文提出的指标是否能够捕捉到学者真实的学术影响力。

（1）对照分析。为了进一步证实学术成就与所提出的指标之间的关系的有效性，本文通过选择三组非获奖学者进行了对照组分析，如图 7 所示，这些学者具有与获奖学者相似的总被引量（图 7（a）—图 7（c））或 h 指数分数（图 7（d）—图 7（f））。第一对照组 C 与获奖学者具有相同的引用，t 检验结果显示获奖学者和对照组 C 之间的引用没有显著差异。然而，在 EC 和 IC 指标上存在明显的差异，获奖学者的平均 EC 高于对照组 C，而平均 IC 显著低于对照组（p value<0.01）。第二对照组 h 与获奖学者具有相似的 h 指数，再次，获奖学者和对照组 h 之间的 h 指数没有显著差异。然而，在 EC 和 IC 方面存在显著差异，获奖学者的平均 EC 显著高于对照组（p value<0.01），而平均 IC 显著低于对照组（p value<0.01）。

图 7 中的结果表明获奖学者在学术网络外部和内部方面具有与其他学者不同的影响力分布。与对照组相比，获奖学者的平均 EC 较高表明网络外部的影响和合作是其学术成就的一种因素。相反，获奖学者的平均 IC 较低表明合作网络内部引用并不能反映学者真实的学术影响力。



注：误差棒表示 95% 置信区间，*p<0.05, **p<0.01, ***p<0.001

图 7 对照分析情况

Figure 7 Control group analysis with Laureates

（2）平均排名和识别比例。获奖学者本文进一步基于诺贝尔奖、沃尔夫奖和狄拉克奖等获奖学者数据，利用平均排名（AR）和识别比例（IP）这两种分析手段对所提出指标的收敛效度做进一步深入研究。

首先，本文计算了样本中所有获奖学者的平均排名以评估每个指标的收敛效度。这种方法确定了获奖学者[50]，获奖学者在整个样本数据中使用给定评价指标的平均影响力排名。如图 8（a）所示，较低的获奖学者平均排名表明该指标在区分获奖学者和其他学者方面更为有效。相反，较高的获奖学者平均排名表明该指标可能在捕捉获奖学者的真实影响力方面不够有效。

其次，本文使用识别比例（IP）来评估指标辨别获奖学者的收敛有效性。通过使用给定指标对样本数据中的所有学者进行影响力排名，并选择前 n% 的样

中国知网 https://www.cnki.net

本来确定获奖学者的识别比例[51]。本文对于 n 从 1% 到 40% 的识别比例平均值作为最终的识别比例，具体计算如公式（5）所示：

$$IP(m,n)=\frac{\Sigma f_n(m,i)}{N},\ for\ n\ from\ 1\%\ to\ 40\% \quad 公式（5）$$

其中，$IP(m, n)$ 是在阈值 n%，处度量 m 的识别比例，N 是样本大小，$f_n(m, i)$ 是一个二进制变量，如果获奖学者 i 在由指标 m，排名的前 n% 中，则该变量等于 1，否则为 0。然后，计算对于 n 从 1% 到 40% 的平均 $f_n(m)$ 作为最终的识别比例。

平均排名和识别比例分析的结果显示在图 8 中，展示了获奖学者在各种指标（包括 EC、Eh、IC、引用、h 指数和发表物）下的平均排名和识别比例计算结果。如图 8 所示，EC 具有最高的平均排名和识别比例，其次是总被引量、Eh、h 指数、IC 和成果。EC 的识别比例为 0.951，表明它能够以较高的精确度识别获奖学者。类似地，获奖学者的 EC 平均排名为 2.23%，表示 EC 是识别获奖学者最有效的指标。

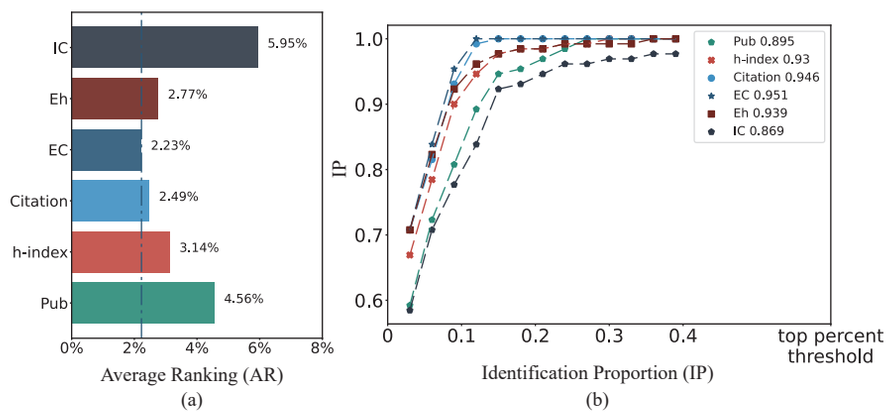


图 8　评估指标的收敛效度

Figure 8　Convergent validity of evaluation metrics

平均排名和识别比例分析的结果表明，与传统的人才评价指标相比，学者合作网络外部影响力评价指标在识别物理学领域的获奖学者方面更为准确。这强调了在评估学者的学术影响时，需要考虑学者的学术网络和内部“互用”的影响。学者的学术网络或研究团队的内部引用占据了学者被引量的重要部分，但本文的结论表明，来自学者学术网络成员的引用并不能反映学者的真实影响力，而是作为“引文泡沫”，对于学者真实影响力测度和人才评价产生了一定的阻碍作用。

6　总结与展望 / Summary and prospect

本文基于学者合作网络的拓扑相似性抽取科研合作网络，在综合考虑学者合作网络内部被引量的基础上，提出合作网络外部被引量（EC）与合作网络外部 h 指数（Eh），以刻画学者的合作网络外部影响力。通过对物理学的 463 348 篇学术论文和 234 086 位消岐学者的数据进行实证研究，本文发现学者影响力在合作网络内部和外部两个方面均存在明显差异，学者合作网络外部的引用在影响力方面发挥着重要作用。对物理学获奖学者和非获奖学者指标的分析进一步证明了学者合作网络外部的影响力更有助于学者获得学术成就。本文提出的网络外部被引量（EC）和网络外部 h 指数（Eh）能够较为合理地反映学者合作网络外部的影响力，可以为学者评价研究提供更加全面的视角，也为政策制定提供新的启示。

本文也存在一些局限性：首先，本文仅基于美国物理学学会的论文数据，尽管有着数据集质量高和可消歧的优势，但该数据集不能体现学者的整个学术网络，且在学科方面仅限于物理学这一领域；其次，本文并未考虑不同的学术合作模式，仅依靠共同署名来确定学者的学术网络，这一方法虽然计算简便、拓展性强，但是在实践中可能存在一些不合理之处；最后，本文虽然基于诺贝尔奖、沃尔夫奖和狄拉克奖等物理学奖项数据证明了合作网络外部影响力评价指标的有效性和优越性，但是这一评价模式是否能真正应用于我国广大科研人才的评价，还需要进一步的研究。

未来研究可以进一步挖掘学者在不同研究领域、地理区域和职业阶段中合作模式的潜在变化，并验证所提出指标在其他学科领域中的实用性。此外，未来研究还可以对本文提出的外部被引量和外部 h 指数等学者合作网络外部影响力评价指标进行优化，例如对学术团队设置阈值、考虑时序模型等。最后，

中国知网 https://www.cnki.net

未来还可以将学者合作网络外部影响力评价指标与我国现行人才评价政策相结合进行研究，以进一步构建面向我国未来科技人才评价的新模式。

参考文献 /References：

[1] LAZER D, PENTLAND A, WATTS D, et al. Computational social science: obstacles and opportunities [J]. Science, 2020, 369(6507): 1060-1062.

[2] 吴利俊，侯鑫鑫，张嵬 . 多要素融合视域下的高校合作贡献度指数构建及实证研究 [J]. 图书情报工作，2024, 68(7): 113-122. (WU L J, HOU X X, ZHANG W. Construction and empirical research on the university cooperation contribution index from the perspective of multi-factor integration [J]. Library and information service, 2024, 68(7): 113-122.)

[3] FONTANAROSA P, BAUCHNER H, FLANAGIN A. Authorship and team science [J]. Jama-journal of the American Medical Association, 2017, 318(24): 2433-2437.

[4] LEDFORD H. Team science [J]. Nature, 2015, 525(7569): 308-311.

[5] 魏雅慧，张琳，黄颖 . 合著关系与研究主题双重视角下的师承合作模式研究 [J]. 图书情报工作，2024, 68(6): 30-47. (WEI Y H, ZHANG L, HUANG Y. Research on the mentorship collaboration patterns from the dual perspectives of co-authorship relationships and research themes[J]. Library and information service, 2024, 68(6): 30-47.)

[6] FONTANA M, IORI M, MONTOBBIO F et al. New and atypical combinations: an assessment of novelty and interdisciplinarity [J]. Research policy, 2020, 49(7): 104063.

[7] VASILYEVA E, KOZLOV A, ALFARO-BITTNER K, et al. Multilayer representation of collaboration networks with higher-order interactions [J]. Scientific reports, 2021, 11(1): 5666.

[8] YANG Y, TIAN T Y, WOODRUFF T K, et al. Gender-diverse teams produce more novel and higher-impact scientific ideas [J]. Proceedings of the National Academy of Sciences of the United States of America, 2022, 119(36): e2200841119.

[9] LIU M, JAISWAL A, BU Y, et al. Team formation and team impact: the balance between team freshness and repeat collaboration [J]. Journal of informetrics, 2022, 16(4): 101337.

[10] SUN Y, LIVAN G, MA A, et al. Interdisciplinary researchers attain better long-term funding performance [J]. Communications physics, 2021, 4(1): 263.

[11] JONES B, WUCHTY S, UZZI B. Multi-university research teams: shifting impact, geography, and stratification in science [J]. Science, 2008, 322(5905): 1259-1262.

[12] PARISI G. Balance research funds across Europe [J]. Nature, 2016, 530(7588): 33-34.

[13] ZHOU J, ZENG A, FAN Y, et al. Identifying important scholars via directed scientific collaboration networks [J]. Scientometrics, 2018, 114(3): 1327-1343.

[14] WAGNER C, WHETSELL T, MUKHERJEE S. International research collaboration: novelty, conventionality, and atypicality in knowledge recombination [J]. Research policy, 2019, 48(5): 1260-1270.

[15] 唐旭丽，李信 . 科研团队多样性对学术颠覆性创新的影响研究 —— 以人工智能领域为例 [J]. 情报学报，2023, 42(1): 43-58. (TANG X L, LI X. Effect of the diversity of scientific teams on disruptive innovation in academia: a case study in the field of artificial intelligence [J]. Journal of the China Society for Scientific and Technical Information, 2023, 42(1): 43-58.)

[16] 吕冬晴，阮选敏，李江，等 . 跨学科知识融合对 D 指数的影响 [J]. 情报学报，2022, 41(3): 263-274. (LYU D Q, RUAN X M, LI J, et al. Research on the relationship between the interdisciplinarity and radicalness of papers from the perspective of knowledge integration [J]. Journal of the China Society for Scientific and Technical Information, 2022, 41(3): 263-274.)

[17] FRONCZAK A, MROWINSKI M, FRONCZAK P. Scientific success from the perspective of the strength of weak ties [J]. Scientific reports, 2022, 12(1): 5074.

[18] PERIANES-RODRÍGUEZ A, OLMEDA-GÓMEZ C, MOYA-ANEGÓN F. Detecting, identifying and visualizing research groups in co-authorship networks [J]. Scientometrics, 2010, 82(2): 307-319.

[19] LI E, LIAO C, YEN H. Co-authorship networks and research impact: a social capital perspective [J]. Research policy, 2013, 42(9): 1515-1530.

[20] MARTIN-ALCAZAR F, RUIZ-MARTINEZ M, SANCHEZ-GARDEY G. Assessing social capital in academic research teams: a measurement instrument proposal [J]. Scientometrics, 2019, 121(2): 917-935.

[21] 吴柯烨，孙建军，谢紫悦 . 基于专利文本挖掘的细粒度技术机会分析 [J]. 情报学报，2023, 42(10): 1199-1212. (WU K Y, SUN J J, XIE Z Y. Research on fine-grained technology opportunity analysis based on patent text mining [J]. Journal of the China Society for Scientific and Technical Information, 2023, 42(10): 1199-1212.)

[22] 张金年，罗艳 . 基于内容的作者研究相似度与潜在合作网络分析 —— 以图书馆学期刊为例 [J]. 情报科学，2021, 39(8): 86-93. (ZHANG J N, LUO Y. Content-based author research similarity and potential collaboration network analysis: taking library science journals as an example [J]. Information science, 2021, 39(8): 86-93.)

[23] PAUL A, DUTTA A. Community detection using local group

中国知网 https://www.cnki.net

assimilation [J]. Expert systems with applications, 2022, 206(Nov.): 117794.

[24] HU Z, LIN A, WILLETT P. Identification of research communities in cited and uncited publications using a co-authorship network [J]. Scientometrics, 2019, 118(1): 1-19.

[25] CAI B, ZENG L, WANG Y, et al. Community detection method based on node density, degree centrality, and k-means clustering in complex network [J]. Entropy, 2019, 21(12): 1145.

[26] CARUSI C, BIANCHI G. Scientific community detection via bipartite scholar/journal graph co-clustering [J]. Journal of informetrics, 2019, 13(1): 354-386.

[27] HAN H, ZHAI X, HAN J, et al. Discovering research teams from scientific papers and patents[C]// Proceedings of the 3rd international conference on computer science and application engineering. New York: ACM, 2019: 1-5.

[28] 张靖雯 , 闵超 . 引文扩散视角下论文学术影响力和社会影响力比较研究 —— 以生物医学为例 [J]. 情报学报 , 2023, 42(1): 31-42. (ZHANG J W, MIN C.A comparative study of the academic and social impact of papers from the diffusion perspective: taking biomedicine as an example [J]. Journal of the China Society for Scientific and Technical Information, 2023, 42(1): 31-42.)

[29] DING Y, CRONIN B. Popular and/or prestigious? measures of scholarly esteem [J]. Information processing & management, 2011, 47(1): 80-96.

[30] YIN Y, WANG Y, EVANS J A, et al. Quantifying the dynamics of failure across science, startups and security [J]. Nature, 2019, 575(7781): 190-194.

[31] MA A, MONDRAGÓN R J, LATORA V. Anatomy of funded research in science [J]. Proceedings of the National Academy of Sciences, 2015, 112(48): 14760-14765.

[32] YIN Y, GAO J, JONES B F, et al. Coevolution of policy and science during the pandemic [J]. Science, 2021, 371(6525): 128-130.

[33] WALTMAN L. A review of the literature on citation impact indicators [J]. Journal of informetrics, 2016, 10(2): 365-391.

[34] KUHN T. Historical structure of scientific discovery [J]. Science, 1962, 136(3518): 760-764.

[35] MIN C, BU Y, WU D, et al. Identifying citation patterns of scientific breakthroughs: a perspective of dynamic citation process [J]. Information processing & management, 2021, 58(1): 102428.

[36] 韩毅 , 伍玉 , 申东阳 , 等 . 中文科研论文未被引探索 I: 外部特征影响研究 —— 以图书馆情报与文献学为例 [J]. 图书情报工作 , 2018, 62(4): 6-13. (HAN Y, WU Y, SHEN D Y, et al. PartI of the exploration on uncited papers in Chinese: the influences of external features of papers: a case study of library and information science [J]. Library and information service, 2018, 62(4): 6-13.)

[37] BRITO R, NAVARRO A. The inconsistency of h -index: a mathematical analysis [J]. Journal of informetrics, 2021, 15(1): 101106.

[38] FONTANAROSA P, BAUCHNER H, FLANAGIN A. Authorship and team science [J]. JAMA, 2017, 318(24): 2433-2437.

[39] HEIDI L. Team science [J]. Nature, 2015, 525(7569): 308-311.

[40] MOHAMMED S, MORGAN A, NYANTAKYI E. On the influence of uncited publications on a researcher's h-index [J]. Scientometrics, 2020, 122(3): 1791-1799.

[41] YIN Y, WANG Y, EVANS J A, et al. Quantifying the dynamics of failure across science, startups and security [J]. Nature, 2019, 575(7781): 190-194.

[42] SINATRA R, WANG D, DEVILLE P, et al. Quantifying the evolution of individual scientific impact [J]. Science, 2016, 354(6312): aaf5239.

[43] HICKS D, WOUTERS P, WALTMAN L, et al. The Leiden Manifesto for research metrics [J]. Nature, 2015, 520(7548): 429-431.

[44] WAY S, MORGAN A, LARREMORE D, Larremore D, et al. Productivity, prominence, and the effects of academic environment [J]. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116(22): 10729-10733.

[45] CHU J, EVANS J. Slowed canonical progress in large fields of science [J]. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118(41): e2021636118.

[46] ALSTOTT J, BULLMORE E T, PLENZ D. Powerlaw: a python package for analysis of heavy-tailed distributions [J]. Plos one, 2014, 9(1): e85777.

[47] MERTON R. The Matthew effect in science. The reward and communication systems of science are considered [J]. Science, 1968, 159(3810): 56-63.

[48] ZHAI L, YAN X, ZHANG G. Bi-directional h-index: a new measure of node centrality in weighted and directed networks [J]. Journal of informetrics, 2018, 12(1): 299-314.

[49] BORNMANN L, TEKLES A. Convergent validity of several indicators measuring disruptiveness with milestone assignments to physics papers by experts [J]. Journal of informetrics, 2021, 15(3): 101159.

[50] WANG R, ZHOU Y, ZENG A. Evaluating scientists by citation and disruption of their representative works [J]. Scientometrics, 2023, 128(3): 1689-1710.

[51] XU S, MARIANI M S, LU L, et al. Unbiased evaluation of

中国知网 https://www.cnki.net

ranking metrics reveals consistent performance in science and technology citation data [J]. Journal of informetrics, 2020, 14(1): 101005.

作者贡献说明 /Author contributions：

闫晓慧：论文撰写，分析数据，进行实验；
杨杰：设计研究方案，分析数据，论文修订；
杨文霞：数据收集；
王昊：论文修订；
邓三鸿：提出研究思路，论文起草，论文最终版本修订。

支撑数据：

本文中使用的引文数据可以通过合理的要求获得。APS 引文数据集可在 https://journals.aps.org/datasets 上访问获取，而物理树数据可以从 https://academictree.org/physics/ 检索。有关获奖学者的信息可以从诺贝尔物理学奖（https://www.nobelprize.org/prizes/physics/）、诺贝尔化学奖（https://www.nobelprize.org/prizes/chemistry/）、沃尔夫奖（https://wolffund.org.il/home-page/）和狄拉克奖章（https://www.ictp.it/prize/dirac-medal）的官方网站上获得。
支撑数据由作者自存储，E-mail：yan_xiaohui@126.com。

A Study on the Relationship Between Scholars' Collaboration External Impact and Academic Success*

Yan Xiaohui[1,3] Yang Alex Jie[2,3] Yang Wenxia[2,3] Wang Hao[2,3] Deng Sanhong[2,3]

[1]Investigation Department, Shanxi Police College, Taiyuan 030401

[2]School of Information Management, Nanjing University, Nanjing 210023

[3]Key Laboratory of Data Engineering and Knowledge Services in Provincial Universities (Nanjing University), Nanjing 210023

Abstract: [Purpose/Significance] Traditional bibliometric indicators fail to capture the distinctions between the internal and external dimensions of academic collaboration networks. This paper introduces new metrics, such as external citations and the external h-index, for evaluating the external impact of scholarly collaboration networks, providing a foundation for a more authentic assessment of scholars' academic influence. **[Method/Process]** Utilizing data from the American Physical Society, which includes 463,348 research papers and 234,086 disambiguated authors, this study extracts scholars' ego-centric collaboration networks from the global network. It distinguishes between internal and external influences within these networks and introduces metrics such as the external citation count and the external h-index. These are tested for consistency against traditional bibliometric indicators and validated using data from prestigious physics awards such as the Nobel Prize, the Wolf Prize, and the Dirac Prize, to demonstrate the effectiveness and superiority of the external impact metrics. **[Result/Conclusion]** The study reveals significant disparities in the distribution of internal versus external influence within scholarly collaboration networks. The proposed external impact metrics introduced are highly consistent with traditional indicators. Convergence validity analysis shows that the external citation count and external h-index are more precise in identifying award-winning scholars compared to traditional metrics. These findings suggest that the external influence of collaboration networks reflects the true impact of scholars and plays a crucial role in academic achievements. The new metrics proposed here offer insightful implications for constructing future talent evaluation systems and breaking away from the "five only" criteria, providing a pathway for reform in academic assessments.

Keywords: scientometrics scientists' evaluation collaborative network external impact

*This work is supported by the National Social Science Fund of China project titled "Real-Time Prediction and Long-Term Evaluation of the True Value and Impact of Academic Achievements in the Big Data Environment" (Grant No. 19BTQ062).

Author(s): Yan Xiaohui, lecturer, PhD; Yang Alex Jie, doctoral candidate; Yang Wenxia, doctoral candidate; Wang Hao, professor, PhD, doctoral supervisor; Deng Sanhong, professor, PhD, doctoral supervisor, corresponding author, E-mail: sanhong@nju.edu.cn.

Received: 2024-01-26 Revised: 2024-05-24 Pages: 96-107

中国知网 https://www.cnki.net