图书馆论坛
Library Tribune
ISSN 1002-1167,CN 44-1306/G2

《图书馆论坛》网络首发论文

题目：卓越引证指数：基于高被引论文引证效力的影响力评估方法
作者：宋欣雨，杨杰，彭兆祺，王左戎，邓三鸿
收稿日期：2026-06-13
网络首发日期：2026-07-27
引用格式：宋欣雨，杨杰，彭兆祺，王左戎，邓三鸿．卓越引证指数：基于高被引论文引证效力的影响力评估方法[J/OL]．图书馆论坛.
https://link.cnki.net/urlid/44.1306.G2.20260727.1508.004

网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

网络首发时间：2026-07-27 15:23:18　网络首发地址：https://link.cnki.net/urlid/44.1306.G2.20260727.1508.004

*本文系国家社会科学基金重点项目“基于领域集体智慧挖掘的颠覆性范式变革预测研究”（项目编号：25ATQ008）研究成果。

卓越引证指数：基于高被引论文引证效力的影响力评估方法*

宋欣雨，杨杰，彭兆祺，王左戎，邓三鸿

摘 要 针对传统文献计量指标在识别突破性研究和卓越科学家时忽略引用价值的问题，文章探索一种基于学术核心圈层关注度的影响力评估新方法——卓越引证指数（ECI），通过统计目标论文被高被引论文引用的次数，测度其学术影响力。基于微软 MAG 的 2.047 亿篇文献数据，以诺贝尔奖获奖者及其论文为高影响力样本，为其匹配同年、同刊、同领域的对照组，通过收敛效度分析和分布特征检验来评估指标有效性。研究发现，诺贝尔奖论文的 ECI 显著高于对照组，在识别诺奖得主方面，ECI 的收敛效度优于总被引次数、h 指数等传统指标；仅 7.78%的论文 ECI 大于 0，凸显学术影响力的金字塔结构；ECI 通过量化来自高被引论文的引用，为传统文献计量指标提供了增量信息，有助于推动基于证据的核心学术影响力评价。

关键词 引证指数；引文分析；学术评价

引用本文格式 宋欣雨,杨杰,彭兆祺,等.卓越引证指数：基于高被引论文引证效力的影响力评估方法[J].图书馆论坛,2026

Excellence Citation Index: Impact Evaluation Method Based on Citation Efficacy of Highly Cited Papers

SONG Xinyu, YANG Jie, PENG Zhaoqi, WANG Zuorong & DENG Sanhong

Abstract: Addressing the problem that traditional bibliometric indicators ignore citation value when identifying groundbreaking research and outstanding scientists, this paper explores a novel impact evaluation method based on attention from core academic circles. The Excellence Citation Index (ECI) is proposed, which measures the academic impact of a target paper by counting the number of citations it receives from highly cited papers. Using data of 204.7 million papers from the Microsoft Academic Graph (MAG), Nobel laureates and their papers are taken as high-impact samples, and a control group is matched for them by the same year, same journal, and same field. The validity of the indicator is verified through convergent validity analysis and distribution characteristic tests. The results show that the ECI of Nobel Prize-winning papers is significantly higher than that of the control group. In identifying Nobel laureates, the convergent validity of ECI is superior to traditional indicators such as total citations and h-index. Only 7.78% of papers have an ECI greater than zero, highlighting the pyramid structure of academic impact. By quantifying citations from highly cited papers, ECI provides incremental information to traditional bibliometric indicators and contributes to evidence-based evaluation of core academic impact.

Keywords: Citation Index; Citation Analysis; Academic Evaluation

0 引言

科学论文与科学家的评估是人才选拔[1]、项目评审及绩效评估[2]体系构建的核心环节。长期以来，这一议题始终是信息科学领域[3-6]和科学学领域[7-8]的关键问题。随着科研投入的持续增长和学术产出的爆炸式增加，如何高效、公正地识别具有突破性贡献的研究成果和卓越科学家，已成为各国科技政策制定者和管理者面临的紧迫挑战。学界开发了诸多基于引用的评估指标以衡量科学论文与科学家的影响力，如 h 指数[9]和 g 指数[10]，这些指标因数据易于获取、计算简单、便于比较而被广泛应用。然而，传统指标因评估一致性不足而受到批评[11]，尤其是在衡量里程碑式研究和卓越科学家的影响力时表现出显著局限性[12-13]。针对这些问题，学术界依据《旧金山宣言》《莱顿宣言》等倡导的原则，致力于推动评估体系的革新，

并呼吁构建更全面的评估指标[14-15]。

尽管存在局限性，基于引用的指标仍是评估科学论文和科学家最常用的方法，这主要归因于学术界对它们的长期使用与高熟悉度。学界已尝试通过高被引论文优化传统指标[16]。例如，2021 年 CWTS 莱顿排名摒弃 h 指数，转而关注前 10%的高被引论文，这一调整亦被 SCImago 机构排名采纳。研究表明，纳入高被引论文的数量或比例可作为基于引用指标的重要扩展[17]。然而，现有研究多聚焦于整体引文指标的优化，如 h 指数衍生模型，目标论文引证文献中高被引论文的数量占比及其对影响力评估的潜在价值尚未得到充分探讨。这一研究空白使得研究成果在学术核心圈层中的关注度难以有效捕捉，也限制了对突破性研究的早期识别能力。对此，本研究提出卓越引证指数，通过量化目标论文引证文献中高被引论文的数量，聚焦引用网络中来自高影响力文献的引用集中度，探索价值优先的影响力评估方法。

1 文献综述

1.1 传统引文评价指标的演进与局限

学术评价的核心挑战在于将复杂的学术贡献抽象为可量化、可比较的指标。传统文献计量指标经历了从“数量主导”向“影响兼顾”的演进过程。早期科研评价资源有限，数量成为最易获取的显性标准。论文发表数量直接统计学者发表的论文总数，用以反映学术生产力规模。然而，过度依赖发文量作为评价标准会带来一系列问题。首先，不同学科在知识生产模式、合作强度与发表周期上存在根本差异，导致论文产出效率悬殊[18]，基于发文量的简单跨学科比较缺乏合理性。其次，普遍存在的发表压力，如部分机构的非升即走制度在一定程度上助长了学术功利化风气，使得学者倾向于追求短平快的研究，而放弃重要但周期长或冷门的方向[19]。被引次数的引入是评价从“产出”转向“影响”的关键一步，它体现了同行对研究成果的重视程度[20]。但被引次数仅是对引用行为的简单计数，无法区分正面引用、批判性引用与工具性引用之间的价值差异[21]。h 指数作为衡量学者学术影响力的核心指标，自 2005 年提出以来被广泛采用。它同时考量学者的论文数量（至少 h 篇）和影响力（每篇至少被引 h 次），在一定程度上避免了单纯追求高产或单篇高引的偏颇。此外，旧论文持续被引可推动 h 指数缓慢上升，鼓励学者追求长期学术价值。然而，h 指数存在一个关键局限：只要 h 核内论文的被引频次增长未能推动 h 核边界（即第 h+1 篇论文的被引次数突破 h）向外扩展，h 指数就保持不变。这意味着，一篇 h 核内的论文即使从被引 10 次跃升至 1,000 次，只要其他论文的被引没有相应提升，h 指数仍然不变，从而导致突破性贡献被系统性低估[22]。g 指数是对 h 指数的改进，更关注高被引论文集的整体影响力[23]，能够对 h 指数相同的学者进行更精细的区分[24]。但 g 指数在特定计算情境下存在争议：为辅助计算而引入零被引文章或虚构论文（如某些实现中将不足 g 篇的部分以 0 被引补齐），这种操作偏离了基于有效学术成果统计的初衷，在处理低产学者时可能产生系统性偏差[25]，缺乏收敛效度[26]。可见，尽管传统指标在演进中不断尝试提升评价的科学性和区分度，但它们均不同程度地忽视研究的原创性深度、理论或方法论的突破性贡献、跨学科融合价值以及特定领域的独特贡献，难以有效识别和衡量真正具有变革性意义的卓越工作。评价体系应更注重考察学者的原始创新能力，并聚焦于挖掘其突出的代表性成果，这正是近年来兴起的代表作评价制度的核心逻辑[27-28]。突破性成果的识别与评价面临量化支撑不足的困境，如何从大量论文中客观筛选出具有核心影响力的代表作，如何在同行评议存在主观偏差时获得可靠的辅助计量证据，仍有待解决。这一困境的实质在于，传统指标将每一次引用视为等价，无法区分来自领域内核心研究者的反复引用与大量边缘研究的提及，二者的学术分量截然不同，但在传统指标结果上可能并无差异。这为本研究提供了切入点，若能区分引用来源的价值层级，使来自高水平研究的引用获得更高权重，便可为代表作和突破性成果的识别及评价提供更聚焦的计量支撑。

1.2 引用价值测度的尝试

针对传统指标将引用等量齐观的缺陷，学界尝试从引用动机、引用强度等维度引入“引

用价值”概念。在引用动机层面，早期研究通常依赖人工标注，通过分析引文上下文中的情感词汇或修辞结构，将引用划分为正面、负面等维度[29]。人工方法准确度较高，但受限于人力成本，难以在大规模文献集上实施。在此基础上，后续研究尝试利用引用情感分析揭示学科交叉中的知识内化规律，通过图传播算法和词向量模型对引文情感进行半自动化测度[30]。然而，此类方法仍高度依赖全文数据的获取，且情感分类器的准确性受训练语料质量的影响较大，跨领域迁移时性能往往下降。因此，在缺乏大规模、高质量全文数据库的情况下，基于引用动机的测度难以成为常规评价工具。在引用强度层面，部分学者根据引用在文中的位置（如引言、方法、讨论）或引用长度赋予不同权重，认为出现在方法部分的引用多为工具性引用，学术价值低于出现在引言或讨论中的概念性引用[31]。这一假设在部分学科中得到了一定验证，但在引用习惯不同的领域，引用位置的学术价值含义可能完全不同，且同一位置引用的动机未必一致，引用强度指标往往需要按领域定制，通用性较差。此外，也有研究尝试结合同行评议或专家判断来识别重要引用，例如获取专家对引用属性的重要性权重，构建引用质量评分指标[32]，专家判断具有权威性，但其成本高昂、主观性强、难以实时更新，无法应用于大规模、常态化的科研评价场景。即使采用德尔菲法等收敛专家意见，仍面临周期长、领域依赖性强等问题。综上所述，已有方法试图从动机、情感、位置、强度等引用行为的内部维度区分引用价值，这一思路固然精细，却因依赖全文数据和专家经验，受限于数据源的获取难度与计算的复杂性，可操作性与跨领域通用性不足。本研究转而采取外部视角，不以引用行为本身为分析对象，而以引用者自身的学术层级为判别依据，将对每条引用内容的微观解读，转换为对引用来源的宏观识别，从而规避了对全文获取与情感分类的高度依赖，为大规模、跨学科的核心影响力研究筛选提供了可行路径。

1.3 高被引论文作为评价基准的原理与实践

学术引用分布呈现显著的金字塔结构，少数高影响力论文及其作者占据了绝大部分引用份额和学术话语权[33-34]。这一现象在量化层面表现为：2000—2015 年间，全球前 1%卓越科学家所占引用份额从 14%攀升至 21%，其基尼系数高达 0.70，远超各国收入差距水平[35]。换言之，学术影响力的集中程度甚至超过了经济财富的分配不均，这为基于精英层的评价范式提供了实证基础。主流的文献计量评价体系，如中科院分区和 JCR 分区，其设计逻辑本身即是对这种学术影响力分层化的制度化回应。在此背景下，高被引论文（通常指同领域、同年度被引频次排名前 1%的论文）成为重要的评价基准。其设计原理在于基于相对排名设定阈值，旨在缓解学科差异、文献规模及发表时间等因素带来的偏差，从而增强跨学科评价的可比性。这一指标已被广泛应用于高校和科研机构的评价实践，并催生了多样化的奖励机制（如高被引科学家评选、论文高被引奖励），显著引导着学者的研究行为。然而，现有基于高被引论文的评价范式存在一个关键局限：它聚焦于目标论文本身是否达到高被引阈值，即关注被引用的数量与排名，却未能充分挖掘引用来源的层级差异对影响力评估的价值。具体而言，一篇论文获得的总被引次数中，可能包含大量来自普通论文的引用，也可能包含少量但来自卓越学者或另一篇高被引论文的引用。前者代表广泛但可能浅层的关注，后者则意味着该研究进入了高影响力论文的引用视野。但现有指标无法区分这两种情况。这种“只计被引次数，不论来源身份”的评价逻辑，可能高估热门但非突破性的工作，而低估小众但具有变革性的研究。上述局限表明，高被引论文作为评价基准的潜力尚未被充分挖掘，现有范式仅将其作为衡量目标论文影响力的标尺，却忽略了它同时可以作为识别引用来源的参照。基于这一认识，本研究不再关注目标论文本身是否达到高被引阈值，转而考察其引证文献中高被引论文的引用频次，以此构建基于核心圈层关注度的影响力评估方法。

2 研究方法

2.1 卓越引证指数

卓越引证指数（Excellence Citation Index，ECI）是一种用于评估科学出版物和科学家影

响力的新型指标。该指标的核心思想是：一篇论文的学术价值不仅取决于其被引总次数，更取决于谁在引用它。具体而言，ECI 依赖于引用目标论文的“高被引论文”数量，其中高被引论文定义为特定领域和年份中被引次数排名前 1%的论文。前 1%阈值的选取依据是文献计量学领域对高被引论文的通行定义。这一标准已被 ESI 等主流学术评价数据库所采用，其合理性建立在学术引用分布的偏态特征之上，引用高度集中于极少数论文，前 1%的论文通常占据了该领域总被引次数的相当比例，构成了学术核心圈层的主体。前 1%能够有效捕捉核心圈层关注信号，维持ECI的价值筛选效力。单篇论文的ECI定义为该论文的引证文献集合中，属于高被引论文的篇数（图 1A），而科学家的 ECI 则通过累加其所有论文引证文献中的高被引论文总数确定（图 1B）。以 2016 年诺贝尔物理学奖获奖论文及获奖者为例展示了 ECI 的计算逻辑（图 1C）。

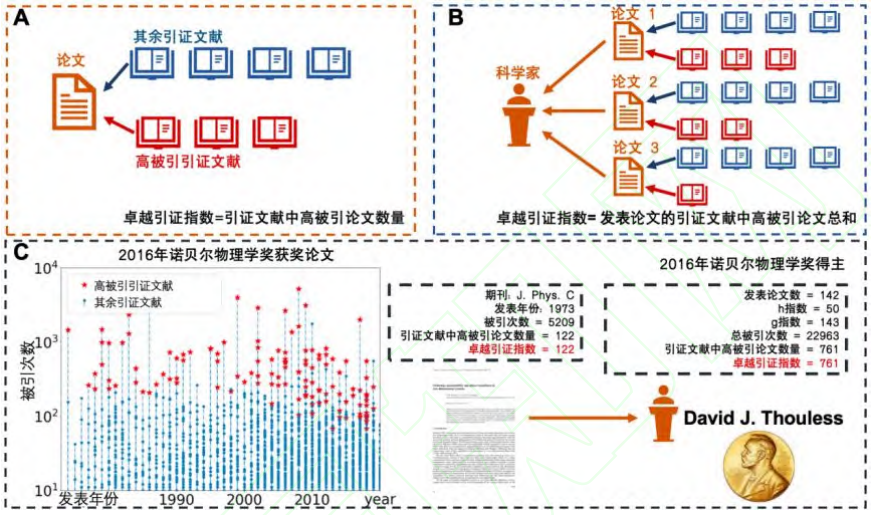


图 1 论文及科学家卓越引证指数的说明与示例

ECI 在评估研究影响力方面的设计逻辑使其在以下方面区别于传统指标。

第一，构建“数量—价值”协同评估的双轨模式。传统被引次数将所有引用等量齐观，而 ECI 只计入来自学术核心圈层（高被引论文）的引用，从而过滤掉边缘性的引用干扰。一篇论文进入领域前 1%，反映的是其被引总量大，但可能源于话题热度而非学术深度。ECI 衡量的则是引用者中有多少篇高被引论文，反映该研究在高影响力成果中的被提及程度。二者可能一致，也可能显著背离，当背离发生时，ECI 提供总被引次数所不能反映的额外信息，即该研究所获引用究竟来自普通引用者还是高影响力学者。这一特性为识别突破性研究（如诺奖级成果）提供辅助计量证据，能揭示研究在学术最前沿群体中的实际关注度。

第二，更强的抗引用操纵性。提升总被引次数可通过自引、互引联盟等策略实现，但提升 ECI 要求引用者自身发表过高被引论文，而高被引论文的发表门槛高，这在逻辑上增加了通过非学术手段操纵的成本与难度。

第三，缓解不同学科的引证生态差异。由于高被引论文的判定在同一学科、同一年度内进行，这使得引证文献筛选环节已对领域和发表时间进行了控制，有助于减少传统指标因学科引用习惯不同等导致的评估偏差，为跨学科研究比较提供了相对公平的计量基准。

第四，体现学术影响力的累积与扩散规律。一项突破性成果发表后，可能需要数年时间才能被领域内高影响力学者引用并写入其高被引论文。ECI 会随着时间推移逐步增长，但其增长曲线反映了影响力向核心圈层渗透的过程，而非普通引用量的线性堆积。这为追踪研究的长期影响提供了新视角。

2.2 MAG 数据集

本研究使用微软学术图谱（Microsoft Academic Graph, MAG）[36]计算 ECI 及传统指标。

MAG 作为大型科学文献计量数据库，其历史规模超过 Web of Science 和 Scopus，后继者 OpenAlex 延续了开放学术图谱的理念，已发展成为全球最大的开放学术数据库。选取 MAG 数据集中 1800—2021 年间发表的 2.047 亿篇文献，涵盖期刊论文、会议论文和预印本等。数据集还包括 2.071 亿已消歧的作者记录及其学科归属信息，为系统性解析 ECI 的分布特征及有效性提供实证基础。图 2B 展示 MAG 数据集中所有论文的描述性统计结果。

2.3 诺贝尔奖论文与获奖者

Li 等[37]提供了物理学、化学和医学领域诺贝尔奖获奖论文及获奖者的数据，从中识别出 874 篇诺贝尔奖获奖论文和 361 位诺贝尔奖获奖者作为研究样本，所有获奖论文的期刊、年份、卷期等信息均与 MAG 数据集中的记录进行了逐一匹配，最大程度控制期刊声望、出版时间窗口及学科领域对引用积累的影响，使 ECI 的组间比较更加纯净。同时，在作者层面选取发文量最高的对照作者，是为了避免因低产作者引入随机偏差，确保对照组与诺奖组在学术活跃度上具有可比性。参考 Min 等[38]使用的方法为诺贝尔奖论文和获奖者选择对照组，如图 2A 所示。具体步骤包括：一是正样本信息提取，从 MAG 数据集中提取每位诺贝尔奖获奖者所撰写的获奖论文，获取其发表年份、期刊名称、卷号、期号以及 OECD 二级学科分类代码等关键元数据；二是论文层面匹配，从 MAG 数据集中选择与诺贝尔奖获奖论文同年、同卷、同期发表且属于同学科类别的论文作为对照组论文；三是作者层面匹配，提取对照组论文的所有作者，并从 MAG 数据集中检索其发表信息，为控制作者学术产出规模的潜在混杂影响，对每位对照作者按照其历史发文总量进行降序排序，选取发文量最高的 10 位作者纳入对照组科学家集合。

上述匹配策略在控制关键混杂因素的同时，其设计特征也决定了后续比较的性质。在论文层面，同年、同刊、同领域的匹配方式控制了发表时间、期刊声望和学科差异对引用指标的干扰，但这也意味着对照组论文整体刊载于与诺奖论文同层次的高水平期刊，其引用表现本身可能已处于较高基线。在作者层面，选取发文量最高的对照作者确保了学术活跃度的可比性，但同样使对照组集中于高产研究者群体。该策略构成了一个条件较为严格的对照设计，对照组并非普通研究者，而是与诺奖得主发表平台相近且产出活跃的研究者。这一设计对 ECI 识别效果的检验是偏保守的，在对照组本身已具备较强学术表现的情况下，ECI 的组间差异可能被低估。若在此条件下 ECI 仍能呈现显著优势，则为其识别效力提供了更稳健的证据支撑。

图 2C 展示了诺贝尔奖获奖论文及其对照组的描述性统计结果，图 2D 呈现了诺贝尔奖获奖者及其对照组的描述性统计结果。

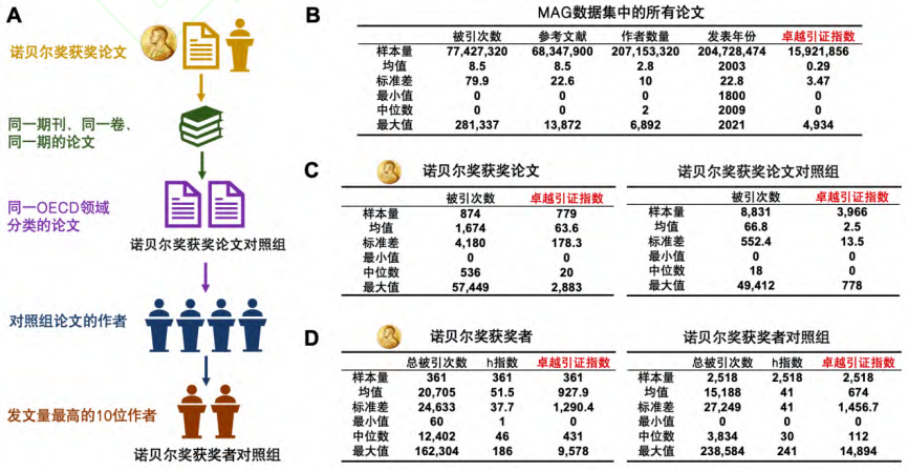


图 2 对照组匹配流程与样本统计

3 结果与分析

3.1 论文卓越引证指数

为探究论文 ECI 的分布特征，将诺贝尔奖获奖论文的 ECI 与对照组论文进行对比，对照组采用同年、同刊、同领域的匹配策略。图 3A 展示了 MAG 数据集中所有论文的 ECI 整体分布。结果显示，在 2.046827 亿篇论文中，仅有 1592.1856 万篇（占比 7.78%）的 ECI 大于 0，呈现显著的长尾分布特征，仅有极少数研究能被领域内前 1%的高被引论文引用，反映高影响力学术引用的稀缺性。ECI 大于 0 的论文意味着研究成果已引起领域内高影响力学者的关注，而绝大多数研究的影响力局限于边缘引用网络。这一分布特征从两个角度界定了 ECI 的适用条件：一方面，ECI 通过聚焦前 1%高被引论文的引用，能够对已进入核心圈层的高影响力研究实现有效筛选；另一方面，对于 ECI 等于 0 的论文，该指标缺乏进一步的区分能力。因此，ECI 的筛选效力本身即意味着它更适合作为识别卓越成果的补充工具，而非面向全体论文的普适性评价指标。

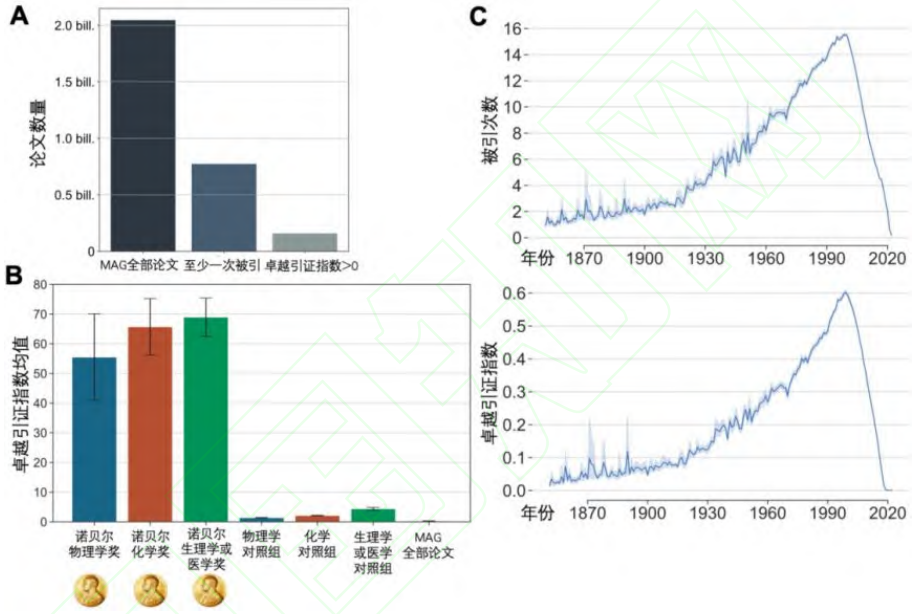


图 3 论文卓越引证指数分布

如图 3B 所示，在物理学、化学、生理学或医学 3 个类别中，诺贝尔奖获奖论文的 ECI 均显著高于其对照组论文，诺奖成果作为公认的突破性研究，在 ECI 上的显著优势，证实了该指标对里程碑式研究的区分效力，其区分度在高影响力论文群体中尤为突出。

此外，考察了 ECI 随时间的分布特征，将其与总被引次数的时间分布进行对比，结果显示两者在宏观趋势上呈现高度相似性（图 3C）。这表明 ECI 并非完全独立于传统被引次数，而是在引用来源层面增加了区分度。高被引论文的引用行为与总被引次数的累积趋势一致，但更聚焦关键引用者的贡献。若某研究在发表初期即被多篇高被引论文引用，其 ECI 会先于总被引次数体现影响力潜力。然而，由于高被引论文的认定需要累积足够的被引窗口期，若引用该研究的高被引论文本身发表时间较晚，ECI 的响应将相应滞后。因此，短期内 ECI 等于 0 未必等同于影响力不足，在应用于新近发表论文的早期识别时需结合时间因素解读。一旦经过这一时间窗口，ECI 即可有效追踪研究影响力的演进过程，既能捕捉因最新高被引论文引用而产生的短期爆发性影响，也能反映持续被领域经典研究引用的长期学术价值，从而为关键突破节点识别及学者职业生涯动态评估提供可参考的时间序列数据。

综上所述，图 3 构建了 ECI 的“三维验证体系”：图 3A 从整体分布证明其筛选功能，图 3B 通过诺奖案例验证其对高影响力研究的识别效力，图 3C 从时间维度确认其与传统指标的互补性及动态评估价值。

3.2 科学家卓越引证指数

图 4A 显示了 ECI 与传统科学家影响力测度指标（发表论文数、总被引次数、h 指数和 g 指数）之间的 Kendall's Tau 相关系数。科学家 ECI 是衡量研究影响力的累积指标，定义为该科学家所有出版物的 ECI 之和。数据显示，无论是获奖者还是对照组，ECI 与总被引次数的相关性最高，与发表论文数的相关性最低，说明该指标对学术产出数量相对不敏感，而是更关注研究成果被学术界采纳的程度。诺贝尔奖获得者的 ECI 与其他传统衡量指标之间的相关性整体低于对照组，该现象可能由多种因素共同导致：一方面，样本量差异、组内异质性及分布偏态等统计因素可能产生一定影响；另一方面，对于诺奖得主这类卓越学者，其影响力往往集中在一些突破性工作上，而非均匀分布在所有论文中，这提示 ECI 可能提供了不同于传统指标的信息增量，同时也反衬出传统评估指标在刻画卓越学者特征时的局限性。

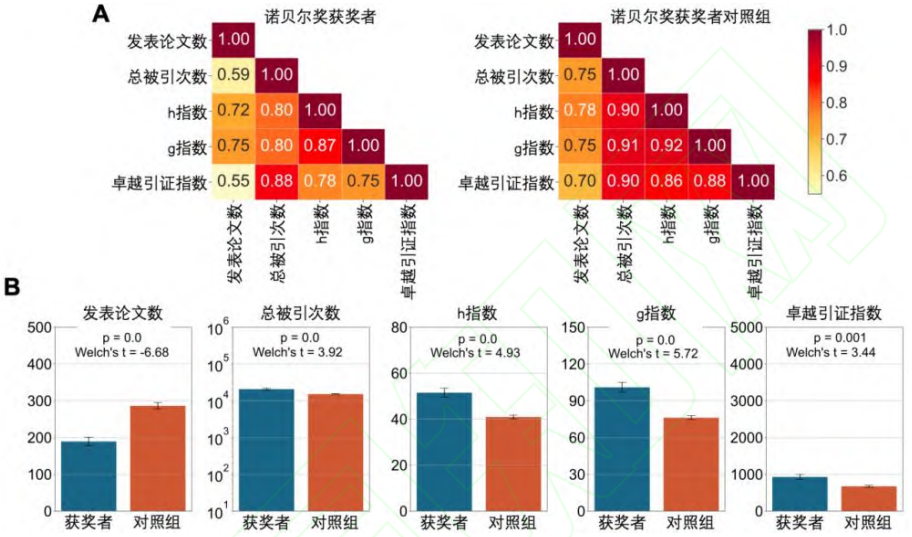
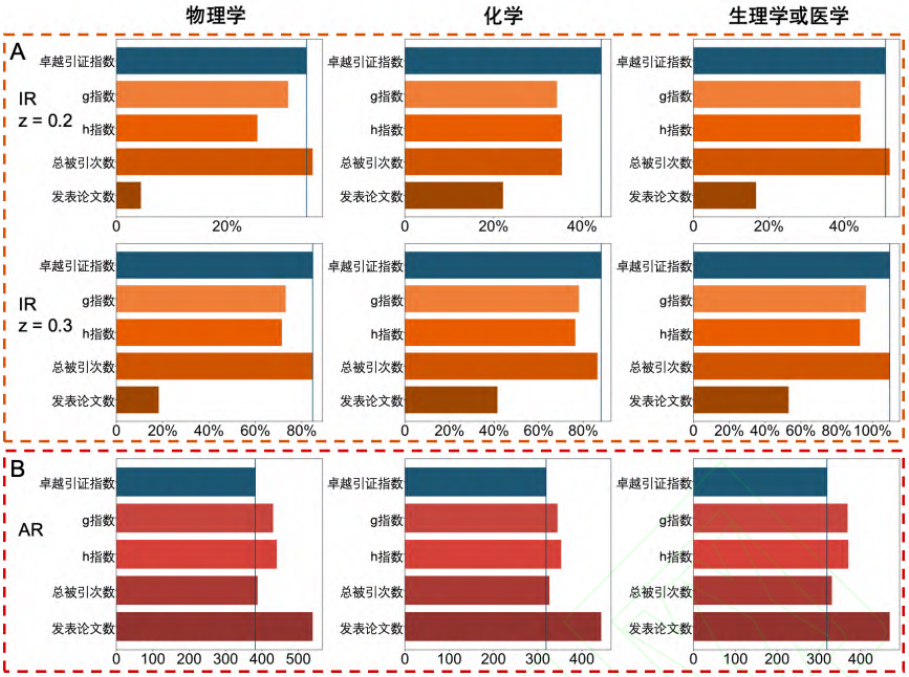


图 4 科学家卓越引证指数与传统指标的对比

为进一步探讨诺贝尔奖获得者与非诺贝尔奖获得者在研究影响力方面的差异，使用 Welch's t 检验对两组的各项指标进行了均值比较，如图 4B 所示，诺贝尔奖获得者的论文发表数量显著低于对照组，但 ECI 和其他指标显著更高，这一结果印证了诺奖得主并非以数量取胜，而是凭借少量但极具影响力的成果脱颖而出。

收敛效度可用于衡量某一指标对目标对象影响范围及效率的捕捉能力[39]。为量化指标的收敛效度，设计两种模型。第一种为识别率（IR），即采用 Xu 等[40]的方法，计算按指标排序前 z 比例科学家中诺贝尔奖得主的占比。IR 值越高，表明指标的收敛效度越强。选取前 20%和 30%的科学家作为测量范围计算 IR 值。第二种模型为所有诺奖得主在该指标下的平均排名（AR）。不同于关注召回率的 IR，AR 衡量所有诺奖得主在指标排序中的平均位置，AR 值越低，收敛效度越强[41-42]。

采用 361 位诺贝尔奖获得者和 2,518 位对照科学家作为研究样本，IR 和 AR 分析结果分别如图 5A 和图 5B 所示。总体而言，除在 z=0.2 水平识别诺奖得主时，ECI 的 IR 略低于总被引次数外，在其余 IR 条件和全部 AR 分析中均优于传统指标，在识别科学界高影响力科学家方面展现更高的区分度和排序敏感性。图 4、图 5 的交叉验证，为构建聚焦于突破性贡献的学术评价体系提供了初步的数据支撑。需要指出的是，上述结论基于以诺奖得主为金标准的验证框架，ECI 在识别尚未获得顶级荣誉但同样具有卓越贡献的科学家时，其区分效力是否同样稳健，仍需进一步检验。



注：识别率(IR)越高，收敛效度越高；平均排名(AR)越低，收敛效度越高。

图 5 卓越引证指数和其他指标的收敛效度

4 总结与讨论

4.1 主要贡献

本研究评估了 ECI 相较于传统指标在衡量高水平研究影响力方面的有效性。通过大规模实证数据证实，诺贝尔奖获奖论文的 ECI 显著高于对照组论文，在识别诺贝尔奖获奖者时，ECI 的收敛效度优于总被引次数、h 指数、g 指数等传统指标。这表明 ECI 是评估高水平研究影响力的一个有价值的文献计量补充指标，为评价论文及科学家提供了可供参考的量化工具。

在理论层面，传统引用指标的核心缺陷在于将所有引用等同视之，难以区分普通引用与高影响力引用。ECI 通过筛选引证文献中前 1%的高被引论文（即学科内具有影响力的核心研究），构建了引用金字塔的顶端识别机制。其评估逻辑在于：引用行为的价值并非由次数决定，而取决于引用者的学术地位。高被引引证文献的聚集程度，反映了研究进入学术核心圈的速度与深度，避免普通引用对评估结果的稀释。ECI 正是通过区分引用来源层级，为这一筛选逻辑提供了可量化的工具。此外，如果某篇论文总被引次数不高，但其引证文献中包含大量高被引论文，仍可证明其在学术脉络中的关键节点地位，这为识别冷启动型的创新研究提供了新路径。基于上述逻辑，ECI 本质上是在评价体系中嵌入基于引用来源层级的证据，作为传统指标或专家判断的补充，为同行评议注入更聚焦、可验证的数据支撑。

在实践层面，ECI 具有多方面的应用价值。其一，面向评价对象，提供更具针对性的识别工具。ECI 通过量化高质量引用网络的连接强度，可辅助识别高影响力研究与学者，为科研资助分配、人才引进、合作机会筛选等决策提供客观的数据支撑。其二，面向研究方法，提供可推广的分析范式。本研究为诺贝尔奖论文和获奖者设计的对照组匹配方法，可推广应用于学者职业生涯轨迹分析、颠覆性成果的早期识别、不同资助模式的绩效评估等更广泛的场景。其三，面向评价生态，引导评价视角的转向。由于 ECI 仅计入来自高被引论文的引用，普通的自引、互引乃至引用联盟等操纵行为难以显著提升 ECI 值，从而在一定程度上减少人为操纵引用的可能性，引导学术共同体将注意力从追求引用量转向创造实质性学术贡献，从

指标设计的内在机理来看，ECI 所依托的高被引论文阈值判定基于学科内相对排名，这一机制在引用来源的识别环节已对学科和年份等因素进行了控制，理论上可适配不同语言和地域的学术数据库，具备向中国本土情境迁移的潜力。更为关键的是，ECI 从“被引数量”到“引用者身份”的视角转换，与我国破五唯、推行代表作制度的内在逻辑存在共通之处：二者均旨在推动评价从看重数量转向看重质量，从看重产出规模转向看重标志性贡献。在此框架下，ECI 有望为识别真正进入核心学术视野的代表性成果提供一种可量化的参考视角。

4.2 局限与展望

本研究的验证工作主要基于自然科学领域诺贝尔奖级成果及其国际数据库，ECI 在更广泛学科与情境中的适用性仍存在以下局限，也由此构成了未来研究的重要方向。

第一，学科维度的适用边界。本研究结论对自然科学基础研究具有较好的解释力，但在人文社会科学、工程技术、交叉学科和新兴领域中的适用性有待检验。在人文社会科学领域，知识生产模式与自然科学存在差异。人文社科研究引用动机更为多元，引经据典式的权威引用、商榷式的批判引用以及工具性的方法引用混杂其中。更重要的是，人文社科的引用周期显著长于自然科学，一项成果可能需要更长时间才能进入学术核心圈层。这导致基于短期引用窗口的高被引论文阈值判定，在捕捉人文社科研究的真实影响力时出现系统性滞后。在工程技术领域，学术影响力与技术影响力存在分流。工程技术研究的重要产出不仅包括论文，还涵盖专利、技术标准、产业化应用等形式的贡献。这些贡献的引用痕迹不完全体现在学术论文的参考文献中，导致 ECI 可能低估那些在产业界具有巨大影响力但学术引用表现相对平淡的研究。而对于交叉学科与新兴领域，其往往横跨多个学科分类体系，引用来源分散在不同领域，前 1%的阈值在单一学科内划定可能失之偏颇，导致跨学科影响力被系统性低估。此外，不同学科的高被引分布形态存在显著差异，部分学科引用高度集中，前 1%门槛极高，部分学科引用相对分散，前 1%门槛较低，这种分布差异使得跨学科比较时 ECI 的绝对数值含义不尽相同。

第二，情境维度的可迁移性。本研究对 ECI 的检验主要依托外文文献数据库，其评价基准与引用生态内嵌于国际学术话语体系。中国已形成较为完备的中文学术期刊体系，其学术共同体在组织方式、评价惯例和引证行为上可能呈现出与国际学术圈不同的特征。更为根本的是，中国科研评价强调研究应服务于国家战略需求与本土实践，这种价值导向使得单纯以国际数据库为参照的指标设计逻辑面临深层适配问题。中国学术共同体内部精英层的形成机制与国际学界不完全一致，前 1%高被引论文这一阈值在中国语境下的有效性与代表性能否等值，仍有待系统验证。

面向上述局限，未来研究可从以下路径展开。针对学科异质性问题，可分别选取不同学科的权威验证样本，系统检验 ECI 的区分效度：在人文社科领域，可引入长引用窗口或动态阈值策略，以缓解引用周期过长导致的识别滞后；在工程技术领域，可构建融合专利引用、技术标准采纳等多元证据链的复合型 ECI 指标，弥补单一学术引用维度对技术影响力的盲区；在交叉学科与新兴领域，可尝试基于多学科分类体系进行阈值加权或引用来源归一化，以校正因学科边界模糊导致的评估偏差，并探索不同高被引分布形态下 ECI 的分学科标准化方案，使跨学科研究更具可比性。针对中国学术情境的可迁移性问题，需开展基于中国高质量期刊数据库的 ECI 实证检验，系统评估 ECI 在中国学术生态下的识别效力与调适策略，并探索其与中国自主评价体系协同融合的可能路径，为该指标从国际情境的有效性走向中国情境的有用性提供实证基础。此外，未来可开展 ECI 在不同高被引阈值设定下的稳健性检验，考察前 5%、前 10%乃至学科特定阈值下 ECI 的分布特征与收敛效度变化，明确阈值选择对评价结果的影响程度与敏感性，为不同应用场景下的阈值调优提供经验依据。

参考文献

[1]陈翀,王嘉怡,王一凡,等.科研人员学术专长的属性度量与应用研究[J].图书情报工作,2025,69(21):83-93.
[2]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.
[3]胡泽文,李慧敏,谢少珂.理工类学科论文零被引率与文献特征之间演变关系的计量比较[J].图书馆论坛,2026,46(5):39-52.
[4]颜兆萍,项波,石进.基于学术和社会维度特征融合的论文影响力评价研究[J].情报科学,2026,44(1):194-201.
[5]王左戎,邓三鸿,宋欣雨,等.全评价视域下面向我国学术成果评价工作的学术知识图谱构建方案研究[J].图书馆杂志,2025,44(2):35-46.
[6]牌艳欣,索传军,肖玥,等.学术论文模糊评价方法研究[J].图书情报工作,2024,68(1):13-26.
[7]AZOULAY P,GRAFF-ZIVIN J,UZZI B,et al.Toward a more scientific science[J].Science,2018,361(6408):1194-1197.
[8]FORTUNATO S,BERGSTROM C T,BORNER K,et al.Science of science[J].Science,2018,359(6379):eaao0185.
[9]HIRSCH J E.An index to quantify an individual's scientific research output[J].Proceedings of the National Academy of Sciences,2005,102(46):16569-16572.
[10]EGGHE L.Theory and practise of the g-index[J].Scientometrics,2006,69(1):131-152.
[11]BRITO R,NAVARRO A R.The inconsistency of h-index: A mathematical analysis[J].Journal of Informetrics,2021,15(1):101106.
[12]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.
[13]WANG R,ZHOU Y,ZENG A.Evaluating scientists by citation and disruption of their representative works[J].Scientometrics,2023,128(3):1689-1710.
[14]CAGAN R.The San Francisco declaration on research assessment[J].Disease Models & Mechanisms,2013,6(4):869-870.
[15]HICKS D,WOUTERS P,WALTMAN L,et al.Bibliometrics: the Leiden Manifesto for research metrics[J].Nature,2015,520(7548):429-431.
[16]胡泽文,崔静静,徐榕,等.社会科学领域潜在高影响力论文的多维指标测度与深度学习预测研究[J].情报资料工作,2026,47(3):16-23.
[17]WALTMAN L.A review of the literature on citation impact indicators[J].Journal of Informetrics,2016,10(2):365-391.
[18]陈思齐,杨承罡,章成志.基于大语言模型的同行评审意见分歧测度[J].图书馆论坛,2026,46(5):107-118.
[19]徐苏兰,余利川.被误用的“非升即走”:功能异化、反噬效应及消解路径[J].江苏高教,2024(9):37-44.
[20]GARFIELD E.Citation indexes for science:A new dimension in documentation through association of ideas[J].Science,1955,122(3159):108-111.
[21]熊回香,杨丹丹,周明洁,等.成长期学者引用动机演化分析[J].现代情报,2024,44(8):110-123.
[22]杨杰,孔嘉,张艺炜,等.融合论文颠覆性与巩固性的学者二元影响力测度[J].情报学报,2023,42(12):1412-1423.
[23]高丹,李秀霞,臧思思.Gm 指数:基于论文被引离散度的作者学术影响力评价新指标[J].情报科学,2020,38(4):57-61.
[24]AHANGAR H G,SIAMIAN H,YAMINFIROOZ M.Evaluation of the scientific outputs of researchers with similar h index: a critical approach[J].Acta Informatica Medica,2014,22(4):255.
[25]ZHANG C T.Relationship of the h‐index, g‐index, and e‐index[J].Journal of the American Society for Information Science and Technology,2010,61(3):625-628.
[26]DE VISSCHER A.What does the g‐index really measure?[J].Journal of the American Society for Information Science and Technology,2011,62(11):2290-2293.
[27]包翔,王世鹏,刘桂锋.社会科学学者中文论文代表作抽选的影响因素[J].图书馆论坛,2024,44(1):86-93.
[28]张逸勤,杨杰,宋欣雨,等.基于代表作的科学家学术影响力评价模式研究——以诺贝尔奖得主为例[J].情报理论与实践,2025,48(6):61-68.
[29]MORAVCSIK M J,MURUGESAN P.Some results on the function and quality of citations[J].Social Studies of Science,1975,5(1):86-92.
[30]姜霖,顾继光.基于引用情感的跨学科知识内化规律——以我国图书情报学科为例[J].情报杂志,2024,43(4):200-207.
[31]王莉军,刘茹,杨波,等.开放科学视角下的数据论文引用意图研究——以生物医学领域为例[J].情报学报,2025,44(3):296-308.
[32]YANIASIH Y,BUDI I.Weighting in-text citation location and context meanings for proposing citation quality score metric[J].Performance Measurement and Metrics,2025,26(1):34-48.
[33]MERTON R K.The Matthew effect in science:The reward and communication systems of science are considered[J].Science,1968,159(3810):56-63.

[34]COLE J R,COLE S.The Ortega Hypothesis:Citation analysis suggests that only a few scientists contribute to scientific progress[J].Science,1972,178(4059):368-375.
[35]NIELSEN M W,ANDERSEN J P.Global citation inequality is on the rise[J].Proceedings of the National Academy of Sciences,2021,118(7):e2012208118.
[36]WANG K,SHEN Z,HUANG C,et al.Microsoft academic graph:When experts are not enough[J].Quantitative Science Studies,2020,1(1):396-413.
[37]LI J,YIN Y,FORTUNATO S,et al.A dataset of publication records for Nobel laureates[J].Scientific Data,2019,6(1):33.
[38]MIN C,BU Y,SUN J.Predicting scientific breakthroughs based on knowledge structure variations[J].Technological Forecasting and Social Change,2021,164:120502.
[39]BORNMANN L,DEVARAKONDA S,TEKLES A,et al.Are disruption index indicators convergently valid?The comparison of several indicator variants with assessments by peers[J].Quantitative Science Studies,2020,1(3):1242-1259.
[40]XU S,MARIANI M S,LU L,et al.Unbiased evaluation of ranking metrics reveals consistent performance in science and technology citation data[J].Journal of Informetrics,2020,14(1):101005.
[41]MARIANI M S,MEDO M,LAFOND F.Early identification of important patents:Design and validation of citation network metrics[J].Technological Forecasting and Social Change,2019,146:644-654.
[42]MARIANI M S,MEDO M,ZHANG Y C.Identification of milestone papers through time-balanced network centrality[J].Journal of Informetrics,2016,10(4):1207-1223.

作者简介 宋欣雨，南京大学信息管理学院博士研究生；杨杰，南京大学信息管理学院博士后；彭兆祺，王左戎，南京大学信息管理学院博士研究生；邓三鸿（通信作者，sanhong@nju.edu.cn），博士，南京大学信息管理学院教授、博士生导师。
收稿日期 2026-06-13

（责任编辑：刘洪）