Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Structural Mediation and Knowledge Integration in Interdisciplinary Research

Shiji Chen
Chinese Academy of Science and Education Evaluation (CASEE), Hangzhou Dianzi University
Yuewei Li
Chinese Academy of Science and Education Evaluation (CASEE), Hangzhou Dianzi University
Yunlong Yu
Chinese Academy of Science and Education Evaluation (CASEE), Hangzhou Dianzi University
Yang Li
liyang0815@hdu.edu.cn

Chinese Academy of Science and Education Evaluation (CASEE), Hangzhou Dianzi University

Article

Keywords: Knowledge integration, Structural mediation, Disciplinary diversity, Betweenness centrality, Citation impact, Knowledge diffusion

Posted Date: August 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10155287/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Interdisciplinary research is commonly conceptualized in terms of knowledge integration, emphasizing the diversity of disciplinary sources combined within a publication. However, integrating heterogeneous knowledge does not necessarily entail occupying a structurally mediating position in citation networks. This study analytically distinguishes between two dimensions often associated with interdisciplinarity: knowledge integration, measured by disciplinary diversity in references, and structural mediation, measured by paper-level betweenness centrality. Using 4,025 publications in scientometrics, we examine the empirical relationships among these dimensions and their associations with citation impact and cross-field knowledge diffusion. Results show little empirical association between knowledge integration and structural mediation, with only a very weak negative association observed for directed betweenness, indicating that integrative papers do not systematically function as structural bridges. Regression analyses further reveal that, within the field of scientometrics, knowledge integration is negatively associated with citation impact, whereas structural mediation is positively associated with citation counts. Despite their divergent effects on citations, both dimensions positively contribute to cross-field knowledge diffusion. These findings underscore that cognitive integration and structural positioning capture distinct organizational properties of scientific work, highlighting the importance of analytically distinguishing their roles when interpreting indicators in interdisciplinary research.

Introduction

In the evolving landscape of contemporary science, interdisciplinarity has become a prominent feature of knowledge production and innovation(Gibbons, et al., 1994). Complex societal and scientific challenges increasingly require the combination of perspectives, methods, and theoretical frameworks from multiple domains. As disciplinary boundaries become more permeable, the integration of heterogeneous knowledge sources has been widely regarded as an important characteristic of creative and transformative research. Consequently, understanding how interdisciplinarity manifests and relates to scientific outcomes has become a central concern in science studies.

A substantial body of research has operationalized interdisciplinarity through disciplinary diversity, typically measured by the distribution of referenced subject categories (Rafols, et al., 2012; Rafols & Meyer, 2010). Diversity-based indicators capture the heterogeneity of intellectual inputs and are often interpreted as reflecting knowledge integration. Numerous studies have examined how such diversity relates to innovation and citation impact, although empirical findings remain mixed across fields and contexts. Alongside diversity-based approaches, another stream of research has focused on structural positioning within citation networks. In this context, betweenness centrality has been interpreted as a structural indicator related to interdisciplinarity (Leydesdorff, 2007a, 2007b; Leydesdorff, et al., 2018). However, as noted by Leydesdorff, et al. (2018), high betweenness at the journal level may reflect multidisciplinarity in spanning disciplinary clusters rather than integrative processes within individual publications. Therefore, structural positioning in citation networks does not necessarily imply integrative knowledge recombination at the publication level.

Diversity-based indicators and betweenness centrality thus capture analytically distinct properties of scholarly organization—compositional knowledge integration and relational structural mediation—which may or may not coincide within individual publications. Despite their conceptual differences, these two properties have rarely been examined jointly at the paper level. Prior studies have either emphasized diversity as the primary operationalization of interdisciplinarity or employed network-based measures without systematically clarifying their relationship to knowledge integration.

As a result, it remains unclear whether highly integrative papers also occupy mediating positions in citation networks, and how these properties relate to research outcomes such as citation impact and cross-field diffusion. Addressing this gap requires analytically distinguishing compositional and structural features rather than assuming they represent a single underlying construct.

To this end, the present study investigates the relationship between reference-based disciplinary diversity and paper-level betweenness centrality using a dataset of publications in scientometrics. We treat knowledge integration and structural mediation as analytically distinct organizational properties and examine how each relates to citation impact and cross-field knowledge diffusion. Specifically, we address the following research questions:

(1) What is the empirical relationship between knowledge integration and paper-level betweenness centrality?

(2) To what extent do highly integrative papers also occupy structurally central positions in citation networks?

(3) How are these two organizational features differently associated with citation impact and cross-field diffusion?

By analytically distinguishing between knowledge integration and structural mediation, this study contributes to greater conceptual clarity about how different structural and cognitive properties are interpreted in interdisciplinary research.

2. Background

2.1 Knowledge Integration and Disciplinary Diversity

In bibliometrics, interdisciplinarity has often been operationalized through indicators of disciplinary diversity, which are commonly interpreted as reflecting knowledge integration. It is often operationalized through the disciplinary diversity of a paper's reference list or through the collaboration patterns of coauthors and their institutional affiliations (Wagner, et al., 2011). This line of research emphasizes the heterogeneity of intellectual inputs as the defining feature of interdisciplinarity.

Rafols and Meyer (2010) proposed that disciplinary diversity consists of three fundamental components: variety, balance, and disparity. Variety refers to the number of distinct disciplines involved; balance

captures how evenly contributions are distributed across these disciplines; and disparity measures the cognitive distance among them. These dimensions have informed a wide range of composite indicators aimed at capturing the integrative character of research outputs.

Earlier operationalizations of interdisciplinarity tended to emphasize either variety or balance. Examples include Citation Outside of Category (Porter & Chubin, 1985), Shannon entropy (Ricotta & Szeidl, 2006), the Brillouin Index (Chang & Huang, 2012), and Herfindahl's diversity index (Porter & Rafols, 2009). More recent measures have sought to incorporate all three components of diversity. The Rao–Stirling (RS) index represents one of the first comprehensive approaches integrating variety, balance, and disparity into a unified framework (Rafols & Meyer, 2010). Nonetheless, Rousseau (2018) noted that RS does not fully satisfy certain formal criteria of "true diversity," particularly regarding the monotonicity of balance.

Subsequent developments have refined diversity-based measures. The Leinster–Cobbold diversity indices (LCDiv), also referred to as ${}^{2}D^{s}$, have been proposed as more theoretically consistent formulations of interdisciplinarity (Mugabushaka, et al., 2016; Zhang, et al., 2016). Leydesdorff (2018) introduced the DIV indicator, which treats variety, balance, and disparity as analytically distinct constructs before combining them (Leydesdorff, 2018; Leydesdorff, et al., 2019a, b). Building on entropy theory, Mutz (2022) reconceptualized the components of diversity as "entropy masses," culminating in the DIVe indicator.

At the same time, recent studies have begun to examine how well bibliometric diversity indicators correspond to other representations of interdisciplinarity. For example, Aksnes, et al. (2026) show that commonly used metrics such as Shannon entropy, ${}^{2}D^{s}$, and DIV* correlate only weakly with authors' self-assessed interdisciplinarity, suggesting that diversity-based indicators may capture only certain dimensions of cross-disciplinary research.

Across these approaches, the dominant analytical focus remains the same: interdisciplinarity is inferred from the heterogeneity of referenced knowledge domains. In this sense, diversity-based indicators primarily capture the extent to which a paper integrates knowledge from multiple disciplinary sources.

While most studies rely on backward-looking reference diversity, some scholars have examined interdisciplinarity from the perspective of citing documents. Carley and Porter (2012) proposed a forward diversity index based on the disciplinary composition of citing papers, thereby shifting attention toward the diffusion of research across fields. This perspective highlights that integration (input heterogeneity) and diffusion (output dispersion) represent analytically distinct processes.

Recent work has extended this perspective by incorporating structural and semantic information into diversity-based approaches. For example, Kong, et al. (2025) propose a citation structural diversity metric that combines citation network structure with semantic features, illustrating how network-level diversity can capture aspects of knowledge diffusion not evident in reference diversity alone. More broadly, van den Besselaar (2025) critiques the limitations of traditional indicators and calls for greater

methodological pluralism in measuring interdisciplinarity, including approaches based on field mapping and dynamic citation relations.

Together, these developments highlight that interdisciplinarity may manifest differently at the stages of knowledge production and knowledge dissemination.

2.2 Structural Mediation in Citation Networks

Citation network analysis offers a structural perspective on the organization of scientific knowledge. Within this framework, betweenness centrality is widely used to characterize the positional significance of nodes in a network. As defined by Freeman (1977), betweenness centrality measures the extent to which a node lies on the shortest paths connecting pairs of other nodes. Nodes with high betweenness values occupy brokerage positions through which information or citations are more likely to flow.

In science mapping and bibliometric research, betweenness centrality has been applied to identify structurally influential actors. Chen (2006) used this metric to detect potentially transformative junctures in scientific development. At the journal level, Leydesdorff (2007a, 2007b) explicitly proposed betweenness centrality as an indicator of interdisciplinarity, arguing that journals bridging otherwise weakly connected citation clusters occupy structurally integrative positions. Subsequent analyses of Journal Citation Reports (JCR) data demonstrated that the interpretive validity of betweenness centrality depends on normalization procedures and local citation contexts (Leydesdorff, 2007a; Leydesdorff & Rafols, 2011a). More recent network analyses continue to employ betweenness centrality to identify actors that bridge fragmented sub-communities within scientific collaboration and citation networks (Kohus, et al., 2025).

Building on this line of inquiry, several studies have revisited the interpretation of betweenness centrality as an indicator of interdisciplinarity. Leydesdorff, et al. (2018) distinguished betweenness centrality from diversity as alternative operationalizations of interdisciplinarity. They suggested that betweenness centrality primarily reflects multidisciplinarity—that is, structural spanning across disciplinary clusters—whereas diversity, particularly in the cited direction, better captures knowledge integration. Similarly, van den Besselaar (2025) argues that betweenness centrality captures structural bridging rather than cognitive integration per se, cautioning against equating structural brokerage with knowledge integration and advocating complementary measures. Recent bibliometric reviews further confirm that betweenness centrality remains one of the most widely used structural indicators in citation network analysis, although its interpretation varies depending on network construction and analytical context (Mukhtar, et al., 2024).

While betweenness centrality has been applied at the publication level—for example, to identify structurally pivotal or potentially transformative papers Chen (2006)—its interpretation as an indicator of interdisciplinarity has primarily been articulated in journal-level studies. In publication-level citation networks, betweenness centrality captures structural mediation: the extent to which a paper lies along citation paths connecting otherwise weakly connected regions of the network. Unlike diversity-based

indicators, which describe the heterogeneity of referenced knowledge domains, betweenness centrality reflects relational positioning within evolving citation flows.

Although both diversity-based measures and structural centrality indicators have been discussed in the literature on interdisciplinarity, they are grounded in different analytical logics. Diversity concerns the cognitive composition of knowledge inputs, whereas betweenness centrality concerns positional brokerage within network structure. Whether structural mediation at the paper level coincides with reference-based knowledge integration—and whether the two exhibit similar or divergent associations with research outcomes—remains insufficiently examined.

2.3 Knowledge Integration, Structural Mediation, and Research Outcomes

A substantial body of research has examined the relationship between interdisciplinarity and scientific impact, typically using citation counts as a proxy for research influence. Empirical findings remain mixed. Some studies report that interdisciplinary research, as measured by disciplinary diversity, tends to receive more citations on average (Chen, et al., 2015; Chen, Qiu, et al., 2021; Larivière, et al., 2015; Wang, et al., 2015), whereas others suggest weak or even negative effects (Larivière & Gingras, 2010; Levitt & Thelwall, 2008; Rafols, et al., 2012). Rather than a simple linear association, several investigations identify nonlinear patterns: moderate levels of interdisciplinarity are often associated with higher citation impact, while very low or very high diversity may correspond to lower performance (Chakraborty, 2018; Larivière & Gingras, 2010; Yegros-Yegros, et al., 2015). These findings suggest that knowledge integration, as captured by disciplinary diversity, does not uniformly translate into citation advantage.

From a network perspective, structural mediation within citation networks has also been linked to research outcomes. Papers occupying mediating positions along citation pathways may benefit from greater visibility and connectivity, as they serve as structural intermediaries between otherwise weakly connected knowledge clusters (Burt, 1992; Freeman, 1977). Bibliometric studies suggest that betweenness centrality—an operational indicator of structural mediation—can correlate with citation performance, although its interpretation depends on network configuration, normalization procedures, and analytical scale (Leydesdorff, 2007a; Leydesdorff & Rafols, 2011a). These findings indicate that structural mediation represents a relational dimension of scholarly influence distinct from compositional knowledge integration.

Beyond citation counts, interdisciplinarity has been associated with patterns of knowledge diffusion. Reference diversity reflects the heterogeneity of intellectual inputs, whereas forward diversity, as captured by citing documents, reflects the cross-field spread of research outputs (Carley & Porter, 2012). Empirical investigations indicate that greater disciplinary diversity may correspond to broader diffusion across fields (Carley, et al., 2019), although diffusion stability and longevity may vary (Hou, et al., 2024). These strands of research suggest that both intellectual composition and structural positioning can shape how research circulates across disciplinary boundaries.

Despite these advances, relatively few studies have jointly examined how knowledge integration and structural mediation relate to citation impact and diffusion at the paper level. In particular, it remains unclear whether papers that integrate diverse knowledge sources also occupy mediating positions within citation networks, and how these dimensions interact in shaping research outcomes. This gap motivates the empirical analyses conducted in the present study.

3. Data and Method

3.1 Data

To examine the structural relationship between cognitive diversity and betweenness centrality at the paper level, we constructed a field-bounded citation network within the domain of scientometrics.

The initial dataset consists of papers published between 2008 and 2017 in two leading journals in the field, Scientometrics and Journal of Informetrics. This publication window ensures a minimum citation observation period of five years for each paper, enabling a more reliable assessment of citation impact. The underlying bibliographic data were accessed and processed through the ICSR Lab platform in May 2024, initially comprising 3,318 publications. ICSR Lab provided access to the underlying bibliographic records, citation relations, and a programmable computing environment. All subsequent data-processing procedures, including citation-network construction, extraction of the largest connected component, cascading citation expansion, and calculation of the bibliometric and network indicators, were implemented by the authors using self-written code within the platform. Although these two journals represent central publication venues for scientometrics research, they do not fully capture the entire scope of the field, and some papers published in these journals may fall outside the domain of scientometrics. To mitigate this issue, we constructed a citation network linking papers within Scientometrics and Journal of Informetrics and extracted the largest connected component using self-written Python code and the NetworkX library installed in the ICSR Lab computing environment. This procedure helps remove papers that are weakly connected to the core citation structure of the field.

The largest connected component of the initial citation network contains 2,485 papers. To further expand the coverage of relevant scientometrics research, we applied the Cascading Citation Expansion method proposed by Chen (2018). This method iteratively extends a seed set of publications by incorporating papers that cite the seed set (forward citation expansion) and papers cited by the seed set (backward citation expansion), thereby broadening the representation of the research domain.

When expanding the foundational dataset, we required supplemented papers identified through both forward and backward citation expansion to have at least five citation links with the seed set. This threshold was used to ensure that the added papers were more closely aligned with scientometrics research while also reducing the computational complexity of betweenness centrality calculation. In addition, all expanded papers were restricted to the same publication window (2008–2017) to maintain temporal consistency with the core dataset.

Using the 2,485 papers in the largest connected component as the seed set, cascading citation expansion identified an additional 1,540 papers that satisfied the citation-link criterion. The final dataset therefore consists of 4,025 scientometrics-related publications, which serve as the basis for all subsequent analyses.

The resulting network represents a local citation environment of scientometrics research. Structural indicators such as betweenness centrality are therefore calculated within this field-bounded citation network, allowing structural mediation to be interpreted relative to the internal citation structure of the field rather than the global citation system. Prior research has shown that the interpretive validity of betweenness centrality in citation-based analyses depends strongly on the delineation of the citation environment and the local context in which the network is defined (Leydesdorff, 2007a; Leydesdorff & Rafols, 2011a).

3.2 Measures

3.2.1 Diversity (Knowledge Integration)

To operationalize knowledge integration, this study employs two reference-based disciplinary diversity indicators: the Rao–Stirling (RS) index and the Leinster–Cobbold diversity index (LCDiv). Both measures quantify the heterogeneity of disciplinary inputs in a paper's reference portfolio, thereby serving as structural proxies for the compositional dimension of knowledge integration.

The RS can be expressed as follows:

$$RS = \sum_{ij}^{n} p_i p_j (1 - s_{ij})$$

where s_{ij} denotes the similarity between discipline *i* and *j*, *n* is the number of disciplines and p_i represents the proportion of references belonging to discipline *i*.

The LCDiv metric incorporates an adjustable parameter q, which regulates the relative weight assigned to frequent versus infrequent components. In interdisciplinary research, q is conventionally set to 2. Under this specification, LCDiv is defined as:

$$\frac{1}{\sum_{ij}^{n} s_{ij} p_i p_j}$$

Here, s_{ij}, *n* and p_i retain the same definitions as provided above.

Both indices integrate three components commonly associated with disciplinary diversity—variety, balance, and disparity—through different mathematical formulations.

To construct these measures, each paper's references were first classified into disciplinary categories using the Science-Metrix classification system, which assigns subject categories at the publication level.

The classification is generated using a machine-learning model trained on bibliographic metadata, including titles, abstracts, keywords, affiliations, and citation patterns (Archambault, et al., 2011). The taxonomy comprises 6 major domains, 22 broad fields, and 174 subfields. In this study, diversity is computed at the most granular level (subfield).

Next, a subfield-level co-citation matrix was constructed to estimate intellectual proximity among disciplines. Using Scopus-indexed publications from 2017–2018, co-citation frequencies were calculated and normalized via cosine similarity (Ahlgren, et al., 2003)

$$s_{i,j} = \frac{\sum_{k=1}^{N} cc_{ik} cc_{jk}}{\sqrt{\left(\sum_{k=1}^{N} cc_{ik}^2\right)\left(\sum_{k=1}^{N} cc_{jk}^2\right)}},$$

where c_{ik} indicates how often subfields *i* and *k* were cocited.

Finally, RS and LCDiv were computed by combining the proportional distribution of referenced subfields with their pairwise similarity values. These diversity indicators capture the heterogeneity of intellectual inputs and thus operationalize the compositional dimension of knowledge integration at the paper level.

3.2.2 Betweenness Centrality (Structural Mediation)

To operationalize structural mediation, this study employs paper-level betweenness centrality. Betweenness centrality quantifies the proportion of shortest paths (geodesics) between pairs of nodes that pass through a focal node. For a given node k, betweenness centrality is defined as:

$$BC_k = \sum_i \sum_j \frac{g_{ijk}}{g_{ij}}$$

where g_{ij} denotes the total number of geodesic paths connecting node *i* to node *j* and g_{ijk} represents how many of these paths pass through k.

In citation networks, betweenness centrality reflects the extent to which a paper occupies intermediary positions within the overall citation topology. It captures how frequently a paper lies along shortest citation paths connecting other publications.

A citation network may be represented as either directed or undirected, resulting in two related but analytically distinct measures: directed betweenness centrality (BCd) and undirected betweenness centrality (BC).

(1) Directed Betweenness Centrality (BCd)

In a directed citation network, edges follow the direction of citation from referenced to citing documents. BCd is computed by strictly preserving this directionality. A paper attains a high BCd value when it lies on a substantial proportion of directed shortest citation paths. In this representation, BCd reflects

positioning within directed citation chains, capturing how a paper is situated between earlier and later works in the flow of citations (Fig. 2a).

(2) Undirected Betweenness Centrality (BC)

When the citation network is treated as undirected, directionality is disregarded and connections are assumed to be symmetric. Under this representation, betweenness centrality reflects intermediary positions in the overall citation structure, independent of citation direction.

As illustrated in Fig. 2, high undirected BC values may arise from several structural configurations:

- **Bridging between clusters (**Fig. 2a**)**, where a paper connects otherwise separated groups;
- **Connections across multiple input communities (**Fig. 2b**)**, where references span distinct areas;
- **Being cited across multiple output communities** (Fig. 2c), where a paper functions as a shared reference point.

These configurations correspond to different patterns of citation structure. Importantly, they arise from network topology rather than from the internal semantic composition of references.

Distinguishing between BC and BCd therefore allows us to examine different aspects of citation positioning. While BCd captures intermediary positions along directed citation paths, undirected BC reflects broader structural embedding within the network. In this study, both measures are used to characterize how papers are situated within the scientometrics citation network.

3.2.3 Diffusion Indicators

The diffusion of a paper can be assessed by examining the disciplinary distribution of its citing publications. To ensure methodological consistency, we apply the same diversity metrics used for reference-based analysis (RS and LCDiv) to the disciplinary composition of citing papers. The key distinction lies in the direction of measurement. Whereas reference-based diversity captures the heterogeneity of disciplinary inputs cited by a paper, the diffusion indicators capture the heterogeneity of disciplinary outputs that cite the paper.

Formally, the diffusion index of paper k is defined as the disciplinary diversity of its citing publications. The two indicators are computed as:

$$Citing\ RS = \sum_{ij}^{n} p_i p_j (1 - s_{ij})$$

$$Citing\ LCDiv = \frac{1}{\sum_{ij}^{n} s_{ij} p_i p_j}$$

Where p_i denote the proportion of citing papers belonging to discipline i, and s_{ij} represent the similarity between disciplines i and j.

These formulations parallel those used for reference-based diversity, but are applied to the distribution of citing disciplines. They therefore capture the extent to which a paper's citations are distributed across disciplinary boundaries.

3.3 Statistics method

This study employs regression analysis to examine the effects of reference-based diversity and betweenness centrality on citation impact and knowledge diffusion. Because citation counts are non-negative integers and typically exhibit highly skewed distributions (Seglen, 1992), count regression models are appropriate (Fleming, 2001; Wang, et al., 2017). While Poisson regression assumes equality of the mean and variance, citation data often display overdispersion. We therefore apply negative binomial regression, which relaxes this assumption by allowing the variance to exceed the mean. The models are estimated using the glm.nb() function in R.

Following prior research, several control variables were included to account for factors associated with citation performance, including publication age, team size, and international collaboration (Bornmann, et al., 2014; Katz & Hicks, 1997; Peters & Vanraan, 1994; Tang, 2013; Vieira & Gomes, 2010). Publication age (ywin) was measured as time since publication. The number of authors was log-transformed to address skewness. International collaboration was captured using a binary variable (icol), coded as 1 if the paper involved institutions from two or more countries. Because the dataset is restricted to a single research domain, disciplinary variation in reference counts is assumed to be limited; therefore, the number of references was not included as a control variable.

To examine knowledge diffusion, measured using continuous diversity indices of citing papers, ordinary least squares (OLS) regression was applied. The same control variables were included to ensure comparability across models.

4. Results

4.1 Relationship Between Integration and Mediation

4.1.1 Correlation analysis

Table 1 reports the correlation matrix among the examined variables. A significant positive correlation is observed between directed betweenness centrality (BCd) and undirected betweenness centrality (BC) (0.52***), indicating that the two centrality measures capture related aspects of structural positioning within the citation network.

In contrast, the reference-based diversity indices (RS and LCDiv), which capture knowledge integration, exhibit no significant correlation with undirected betweenness (r = -0.01). Their association with directed betweenness is statistically significant but extremely weak (r = − 0.03, p < .001). The small magnitude of

these coefficients suggests that knowledge integration and structural mediation are largely independent at the paper level.

These patterns indicate that papers integrating heterogeneous disciplinary knowledge are not necessarily positioned on central mediation paths within the citation network. Conversely, papers occupying structurally central positions do not systematically exhibit high levels of reference diversity.

Table 1
Correlation matrix of integration and structural mediation indicators

BCd	**BC**	**BCd**	**RS**
	0.52* **		
RS	-0.01	-0.03***	
LCDiv	-0.01	-0.03***	0.61***
Notes: ***p < 0.001. Correlation coefficients are Pearson correlations.			

4.1.2 Typological Patterns of Integration and Mediation

To further investigate the relationship between disciplinary diversity and betweenness centrality, we apply cluster analysis to classify scientometrics papers using these two indicators as clustering dimensions. K-means clustering was employed, and the optimal number of clusters was evaluated using the Elbow Method and the Silhouette Score. The Elbow Method examines changes in the sum of squared errors (SSE) across different cluster numbers, while the Silhouette Score assesses clustering quality based on intra- and inter-cluster distances, with higher values indicating better separation. As shown in Fig. 3, the elbow curve did not exhibit a uniquely pronounced inflection point, whereas the silhouette score reached its maximum at k = 3, supporting the three-cluster solution as the most parsimonious and well-separated configuration.

We compared the clustering solutions with three and four clusters and observed that three well-defined clusters consistently emerged. Under the four-cluster specification, one cluster contained only a very small number of observations; in some combinations (e.g., RS with BC/BCd), it reduced to a single node. Given its instability and limited substantive meaning, we therefore adopt three clusters as the final specification.

Figures 4a–d present the clustering results derived from combining the two disciplinary diversity indicators (LCDiv and RS) with the two centrality measures (BC and BCd). Across these combinations, three distinct clusters are identified: (1) low disciplinary diversity – low betweenness; (2) high disciplinary diversity – low betweenness; and (3) high betweenness – low disciplinary diversity. Notably, no cluster emerges that simultaneously exhibits both high disciplinary diversity and high betweenness.

The results indicate a clear separation between integrative diversity and structural mediation: papers with high disciplinary diversity tend to exhibit low betweenness centrality, whereas papers with high

betweenness centrality generally display low disciplinary diversity. This pattern is particularly visible in the combinations involving LCDiv with both BC and BCd. In addition, a small number of observations show unusually high betweenness centrality despite low disciplinary diversity, suggesting that certain cognitively narrow papers may nonetheless function as structural intermediaries within the citation network.

4.2 Interpreting high betweenness centrality papers: illustrative examples

The following examples are illustrative rather than inferential and are used solely to clarify the substantive meaning of BC and BCd as indicators of structural mediation within the citation network. As noted above, betweenness centrality reflects the extent to which a paper lies on shortest citation paths connecting other papers. It therefore captures structural mediation within citation pathways rather than reference-based knowledge integration.

Tables 2 and 3 list the top 10 scientometrics papers ranked by BC and BCd, respectively. Our analysis shows that these highly central papers predominantly focus on methodological developments, analytical tools, theoretical frameworks, or comprehensive review studies that synthesize and systematize existing approaches. Their elevated BC and BCd values are consistent with their positioning along widely traversed citation paths, making them frequently referenced nodes in the structure of the field.

However, when examining reference-based disciplinary diversity, most of the papers listed in Tables 2 and 3 do not exhibit particularly high levels of diversity. This suggests that papers occupying structurally central positions in citation pathways are not necessarily characterized by heterogeneous disciplinary integration. This pattern reinforces the distinction between structural mediation and knowledge integration: high centrality appears to reflect structural embedding within widely used methodological or conceptual trajectories rather than extensive disciplinary recombination.

A closer inspection of Tables 2 and 3 further illustrates the distinctions between BC and BCd. All the papers in Table 3 exhibit high BCd values, indicating that they frequently lie on directed citation chains. In Table 2, five cases (the 3rd, 6th, 7th, 8th, and 9th papers) display relatively high BC values but comparatively lower BCd values. These cases reflect different structural patterns within the network:

- Papers 3 and 7 show higher citing diversity values, indicating that they are referenced across a relatively wide range of research areas.
- Paper 6 exhibits comparatively higher reference diversity, suggesting a stronger degree of knowledge integration.
- Papers 8 and 9 present a more balanced profile, combining moderate diversity with noticeable structural centrality.

At the same time, several papers exhibit both high BC and high BCd values. These papers occupy structurally prominent positions in both directed and undirected network representations, indicating their

importance within the citation topology of the field. Importantly, such structural prominence does not uniformly correspond to high levels of disciplinary diversity.

Finally, a clarification regarding BC and BCd is necessary. In undirected networks, all intermediary nodes along any path are counted in the betweenness calculation, regardless of citation direction. In contrast, directed betweenness centrality considers only intermediary nodes along directed citation paths, which reduces the number of admissible paths. Consequently, BC values are typically higher than BCd values. However, in cases where clear directed citation chains exist and alternative undirected paths bypass the focal node, BCd may exceed BC. Table 3 includes several examples in which BCd values are relatively higher than BC values, consistent with such structural configurations.

Table 2
Top 10 publications ranked by their BC values

ID	Title	BC	BCd	RS	LCDiv	citingRS	citingLCDiv
1	van Eck and Waltman (2010)	25.571	1.475	0.161	1.191	0.436	1.831
2	Bornmann and Daniel (2008)	10.189	2.640	0.208	1.514	0.189	1.258
3	Mongeon and Paul-Hus (2016)	8.134	0.574	0.076	1.082	0.353	1.586
4	Tahamtan, et al. (2016)	5.474	4.170	0.260	1.351	0.272	1.422
5	Aria and Cuccurullo (2017)	4.951	2.605	0.169	1.203	0.427	1.808
6	Moed (2010)	3.987	0.478	0.289	1.762	0.129	1.183
7	Blondel, et al. (2008)	3.406	0.140	0.074	1.080	0.327	1.513
8	Waltman and van Eck (2012)	3.353	0.730	0.271	1.371	0.196	1.265
9	Waltman, et al. (2011)	3.349	0.485	0.151	1.178	0.106	1.126
10	Waltman, et al. (2010)	3.302	1.460	0.286	1.400	0.383	1.669

Notes: Due to the small magnitudes of the original BC and BCd values, the BC and BCd values in the tables are the results of multiplying the original values by 1000.

Table 3
Top 10 publications ranked by their BCd values

ID	Title	BC	BCd	RS	LCDiv	citingRS	citingLCDiv	
1	Cobo, et al. (2011b)			2.892	5.283	0.181	1.221	0.355	1.589
2	Tahamtan, et al. (2016)			5.474	4.17	0.26	1.351	0.272	1.422
3	Leydesdorff and Rafols (2011b)			0.563	4.167	0.28	1.536	0.185	1.228
4	Bornmann and Daniel (2008)			10.189	2.64	0.208	1.514	0.189	1.258
5	Aria and Cuccurullo (2017)			4.951	2.605	0.169	1.203	0.427	1.808
6	Hsieh (2011)			0.388	2.488	0.174	1.211	0.152	1.179
7	Cobo, et al. (2011a)			2.557	2.376	0.217	1.277	0.354	1.604
8	Meho and Rogers (2008)			0.958	2.05	0.106	1.173	0.307	1.459
9	Vanclay (2012)			2.068	2.009	0.297	1.653	0.192	1.262
10	Perianes-Rodriguez, et al. (2016)			1.776	1.739	0.144	1.169	0.411	1.718

Notes: Due to the small magnitudes of the original BC and BCd values, the BC and BCd values in the tables are the results of multiplying the original values by 1000.

Taken together, these illustrative examples further support the empirical separation between structural mediation and knowledge integration observed in the correlation and clustering analyses, setting the stage for examining how each relates to research outcomes.

4.3 Differential Effects on Citation Impact

Table 4 presents the negative binomial regression results assessing how knowledge integration (RS and LCDiv) and structural mediation (BC and BCd) relate to citation impact. Both BC and BCd exhibit significant positive coefficients, indicating that papers occupying structurally central positions within the citation network tend to receive more citations. In other words, papers that lie on frequently traversed citation paths are more likely to accumulate citations over time. Among the control variables, team size, international collaboration, and publication age all show positive and statistically significant associations with citation counts across specifications, consistent with prior findings in bibliometric research. These effects do not alter the substantive relationships observed for structural mediation and knowledge integration.

Table 4
Negative binomial regression results for citation counts

Independent variables	Dependent variable: Citation counts			
	(1)	(2)	(3)	(4)
BC	1.632***	-	-	-
	(0.026)			
BCd	-	2.894***	-	-
		(0.069)		
RS	-	-	-1.087***	-
			(0.196)	
LCDiv	-	-	-	-0.296***
				(0.071)
lnau	0.190***	0.270***	0.219***	0.217***
	(0.028)	(0.030)	(0.034)	(0.034)
icol	0.089**	0.137***	0.129***	0.131***
	(0.032)	(0.034)	(0.039)	(0.039)
ywin	0.140***	0.162***	0.156***	0.160***
	(0.005)	(0.005)	(0.006)	(0.006)
Intercept	2.152***	1.953***	2.701***	2.863***
	(0.056)	(0.059)	(0.077)	(0.112)
Note: ***$p < 0.000$. **$p < 0.001$.				

Parenthetical values indicate standard errors.

In contrast, both measures of reference-based disciplinary diversity (RS and LCDiv) show significant negative associations with citation counts. This finding contrasts with studies in other domains where interdisciplinary integration has been associated with broader audience reach and higher citation impact. However, the relationship between interdisciplinarity and impact is widely recognized as field-dependent. Prior research suggests that the citation consequences of interdisciplinary integration depend critically on the citation intensity of the disciplines being combined. For instance, Larivière and Gingras (2010) and Chen, Song, et al. (2021) demonstrate that interdisciplinarity tends to enhance impact primarily when it incorporates highly cited disciplines, whereas integration involving lower-

citation fields may dilute visibility. In this context, the negative association observed in scientometrics may reflect the specific citation ecology of the field rather than a generalizable disadvantage of knowledge integration.

To explore whether this contextual factor may contribute to the negative association observed in Scientometrics, we compared the average citation counts of Scientometrics (2008–2017) with those of its primarily cited disciplines (Table 5). The data reveal a substantial citation density gap. Although Scientometrics is categorized under Information & Library Science, its average citation count (68.24) is markedly higher than that of the broader field (12.92). Moreover, other disciplines frequently intersected by Scientometrics, such as Education (18.38) and Economics (20.48), also exhibit considerably lower citation averages. This pattern suggests that, within the citation ecology of Scientometrics, increasing disciplinary diversity may involve integrating knowledge from fields with comparatively lower citation density. Such a configuration could, at least in part, attenuate citation performance, thereby offering a plausible contextual interpretation for the observed negative association between disciplinary diversity and scientific impact.

Table 5
The average citations of several mainly cited disciplines of Scientometrics

Mainly cited disciplines	**The number of cited papers**	**Average citations**
Information & Library Sciences	2862	12.917
Science studies	2696	31.353
Artificial intelligence & image processing	1840	19.21
Business & management	1389	32.825
Economics	1270	20.478
Education	1084	18.382
General Physics	1025	27.424

Taken together, the results indicate that structural mediation and knowledge integration relate to citation impact in different ways. Structural positioning within the citation network exhibits a consistent and positive association with citation counts, whereas reference-based diversity shows a weak or negative association in this field-specific context. This pattern suggests that occupying structurally central positions in citation pathways may contribute to scholarly visibility even when integrative diversity does not translate into citation advantages.

These findings highlight the importance of distinguishing between network position and knowledge composition when examining the relationship between interdisciplinary characteristics and citation performance. While diversity captures the heterogeneity of intellectual inputs, structural mediation reflects a paper's location within citation structures, and these dimensions do not appear to influence citation outcomes in the same manner.

4.4 Differential Effects on Cross-Disciplinary Diffusion

We further examine the relationship between knowledge integration, structural mediation, and knowledge diffusion. Tables 6 and 7 present regression results assessing how citing diversity relates to reference-based diversity and betweenness centrality.

Across all models, both betweenness centrality measures (BC and BCd) and diversity-based indicators (LCDiv and RS) exhibit significant positive associations with citing diversity. These results indicate that papers characterized by higher structural mediation or higher reference diversity tend to be cited by works originating from a broader range of disciplinary areas.

Specifically, higher betweenness centrality is associated with greater citing diversity, suggesting that structurally central papers are more likely to appear along citation paths that span multiple areas. Similarly, higher reference-based diversity is positively related to citing diversity, indicating that papers integrating heterogeneous knowledge sources are also more frequently referenced across disciplinary boundaries.

The effects of control variables are broadly consistent with those observed in the citation impact models and do not substantially modify the main patterns concerning integration and mediation.

Table 6
The regression analysis result of citing LCDiv and interdisciplinarity

Independent variables	**Dependent variable: CitingLCDiv**		
	(1)	**(2)**	**(3)**
BC	0.036***	-	-
	(0.005)		
BCd	-	0.053***	-
		(0.013)	
LCDiv	-	-	0.227***
			(0.011)
lnau	0.035***	0.036***	0.035***
	(0.005)	(0.005)	(0.005)
icol	0.003	0.003	0.004
	(0.006)	(0.006)	(0.006)
ywin	-0.002.	-0.002.	-0.003***
	(0.001)	(0.001)	(0.001)
Intercept	1.228***	1.229**	0.943***
	(0.011)	(0.011)	(0.017)
Note: ***$p < 0.000$. .$p < 0.01$.			

Parenthetical values indicate standard errors.

Table 7
The regression model about citing RS and interdisciplinarity

Independent variables	Dependent variable: CitingRS		
	(1)	(2)	(3)
BC	0.020***	-	-
	(0.003)		
BCd	-	0.036***	-
		(0.007)	
RS	-	-	0.484***
			(0.016)
lnau	0.021***	0.021***	0.019***
	(0.003)	(0.003)	(0.003)
icol	-0.001	-0.002	-0.001
	(0.003)	(0.003)	(0.003)
ywin	0.000	0.000	0.001
	(0.001)	(0.001)	(0.000)
Intercept	0.146***	0.146***	0.056***
	(0.004)	(0.006)	(0.005)
Note: ***p < 0.000.			

Parenthetical values indicate standard errors.

Overall, the findings suggest that knowledge integration and structural mediation are each positively associated with diffusion across disciplinary contexts. Importantly, while both dimensions relate to citing diversity, they represent analytically distinct characteristics of papers, consistent with the earlier results showing that integration and structural mediation do not strongly coincide at the paper level.

5. Discussion

5.1 Distinguishing Integration from Structural Positioning

The analyses indicate that knowledge integration and structural positioning represent empirically distinct structural properties within citation networks. Reference-based diversity indicators capture the heterogeneity of intellectual inputs incorporated into a paper, whereas betweenness centrality reflects a

paper's location within citation pathways. Although both relate to knowledge connections in different senses, they do not strongly coincide at the paper level.

The weak and slightly negative association between diversity measures and directed betweenness centrality, together with the clustering patterns, suggests that highly integrative papers are not necessarily positioned on central mediation paths in citation networks. Conversely, papers occupying structurally prominent positions do not systematically exhibit high levels of disciplinary diversity. The absence of a cluster combining both high integration and high structural mediation further reinforces this distinction.

These findings suggest that combining diverse knowledge sources and occupying a structurally central position are different aspects of scholarly organization. Integration concerns what kinds of knowledge are brought together within a paper, whereas structural mediation concerns where that paper is situated in the topology of citation relations. Conflating these properties may obscure important differences in how papers function within scientific communication systems.

5.2 Why Might Integrative Papers Not Occupy Brokerage Positions?

The observed weak relationship between reference diversity and structural mediation warrants further consideration. One possible interpretation relates to the different structural logics underlying these two properties. Reference diversity reflects the heterogeneity of knowledge sources incorporated into a paper, whereas betweenness centrality depends on how that paper becomes embedded within evolving citation pathways. A work may integrate heterogeneous knowledge without becoming situated along major citation routes that connect other publications.

Citation networks also tend to consolidate around established themes, widely adopted methods, or influential conceptual frameworks. Papers that recombine diverse knowledge sources may introduce new perspectives without necessarily becoming central connectors within these trajectories. As a result, integrative breadth does not automatically translate into structural mediation.

Conversely, structural centrality is often associated with contributions that function as methodological tools, conceptual anchors, or broadly applicable models. Such papers can become common reference points across multiple research streams, increasing their likelihood of appearing on shortest citation paths. However, these works do not necessarily draw on a highly diverse set of disciplinary references themselves.

Taken together, these considerations suggest that integration and mediation reflect different organizing principles within scientific communication. The former concerns the diversity of intellectual inputs, whereas the latter reflects positioning within an evolving relational structure. The limited empirical overlap between the two therefore indicates structural differentiation rather than contradiction. These

interpretations remain exploratory; the present study does not establish causal mechanisms but highlights the importance of analytically separating these properties.

5.3 Divergent Associations with Impact and Diffusion

Beyond the structural distinction identified above, the two dimensions also exhibit divergent empirical associations with scholarly outcomes. The contrasting associations of disciplinary diversity and betweenness centrality with citation impact reinforce the importance of analytically distinguishing between knowledge integration and structural mediation. Although both dimensions relate to how knowledge is connected or combined, their empirical relationships with citation performance differ within the context of scientometrics.

The regression results indicate a negative association between reference-based diversity (RS and LCDiv) and citation counts. This pattern should not be interpreted as evidence that integrating heterogeneous knowledge is inherently disadvantageous. Rather, it appears to reflect characteristics of the citation environment in which scientometrics is embedded. As shown in Table 5, several disciplines frequently referenced by scientometrics—such as Library and Information Science, Education, and Economics—exhibit lower average citation densities compared to core scientometric publications. When papers draw more extensively on these areas, their reference portfolios may become linked to communities with comparatively lower citation activity. The observed negative association therefore likely reflects field-specific citation dynamics rather than a general limitation of integrative research.

By contrast, both BC and BCd exhibit consistent positive associations with citation counts, underscoring the relevance of structural positioning within citation networks. Papers with high betweenness centrality tend to occupy positions along major citation pathways, increasing the likelihood that they are encountered and referenced by subsequent research. As illustrated in the Results section, many highly central papers introduce methodological tools, conceptual frameworks, or broadly applicable approaches that are widely adopted within the field. Their higher citation performance thus appears to be related to structural prominence rather than to the breadth of disciplinary inputs they incorporate.

In contrast to citation impact, the diffusion results reveal a different empirical pattern. Both knowledge integration and structural mediation are positively associated with citing diversity, indicating that papers characterized by either broader reference portfolios or more central structural positions tend to be cited across a wider range of disciplinary contexts. This suggests that citation volume and diffusion breadth capture distinct aspects of scholarly influence. While structural mediation appears more closely related to citation prominence within the field, knowledge integration may still contribute to extending the disciplinary reach of research outputs.

Taken together, these findings suggest that knowledge integration and structural mediation relate to scholarly outcomes in distinct and non-equivalent ways. Diversity reflects the composition of intellectual inputs, whereas betweenness reflects a paper's location within evolving citation structures. Their contrasting relationships with citation impact, alongside their shared association with diffusion breadth,

highlight the importance of distinguishing between visibility within citation networks and reach across disciplinary contexts.

5.4 Limitations and Future Directions

Several limitations should be acknowledged. The analysis is confined to the field of scientometrics, and the observed relationships may vary across disciplines with different citation cultures and structural dynamics. Future research could extend the framework to additional scientific domains to examine the robustness of these patterns.

Moreover, the study relies on citation-based network measures and reference-based diversity indicators. While widely used, these metrics cannot fully capture the cognitive, semantic, or conceptual dimensions of interdisciplinary knowledge production. Integrating content-based analyses or semantic network approaches may further refine the distinction between integration and structural mediation.

Despite these limitations, the study contributes to ongoing discussions about how cross-boundary research can be more precisely conceptualized and measured. By empirically disentangling integration and structural positioning, it provides a clearer analytical foundation for future investigations into the organization and circulation of scientific knowledge.

6. Conclusion

This study examined the relationship between knowledge integration and structural mediation at the paper level within the field of scientometrics. Using reference-based diversity indicators and citation network betweenness centrality, we investigated whether cognitively integrative papers also occupy structurally central positions in citation pathways, and how these properties relate to citation impact and disciplinary diffusion.

The findings demonstrate that knowledge integration and structural mediation constitute empirically distinct organizational properties of scientific communication. Reference diversity exhibits minimal empirical overlap with betweenness centrality, and clustering analyses show that highly integrative papers rarely coincide with structurally mediating ones. These results indicate that incorporating heterogeneous disciplinary inputs does not inherently imply occupying a central structural position within citation networks. In this study, disciplinary diversity and betweenness centrality are employed as operational proxies for compositional integration and structural mediation, respectively, rather than as direct measures of cognitive synthesis or conceptual bridging.

With respect to research outcomes, structural mediation shows a stronger association with citation impact, whereas both integration and mediation are positively related to cross-disciplinary diffusion. This divergence suggests that citation prominence and diffusion breadth reflect different dimensions of scholarly organization. Structural positioning appears particularly consequential for visibility within citation networks, while knowledge integration contributes more directly to disciplinary reach.

Overall, the study provides empirical clarification regarding the relationship between compositional and structural dimensions of interdisciplinarity at the publication level. Rather than assuming conceptual equivalence, the results indicate that these properties operate through different organizational logics and exhibit distinct outcome associations. Recognizing this differentiation can support more precise interpretation of diversity-based and network-based indicators in future scientometric research.

Declarations

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Competing Interests

The authors declare no competing interests.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Author Contribution

Shiji Chen: Conceptualization, Methodology, Supervision, Writing – original draft.Yuewei Li: Data curation, Formal analysis, Writing – review & editing.Yunlong Yu: Data curation, Formal analysis.Yang Li: Writing – review & editing, Project administration.

Acknowledgement

This work was supported by the National Social Science Foundation of China (Grant No. 25ATQ007)

Data Availability

The final paper-level analytical dataset supporting the findings of this studyis provided in the supplementary materials. The dataset contains the 4,025publications included in the analyses and the paper-level indicators, dependentvariables, and control variables used in the correlation, clustering, andregression analyses. A data dictionary and the code used to reproduce theseanalyses from the retained analytical data are also provided.The underlying raw bibliographic records, citation-relation data, reference-and citing-publication disciplinary distributions, and the working files andcode retained within the ICSR Lab environment are no longer accessible. Theunderlying data were accessed through ICSR Lab in May 2024, and the dataprocessing, citation-network construction, largest-connected-

componentextraction, citation expansion, and initial indicator calculation wereimplemented by the authors using code executed within the platform. Theplatform did not permit the underlying records and platform-based working filesto be exported and has since been discontinued. Consequently, the completeworkflow cannot now be reproduced from the original bibliographic records. Thematerials provided permit reproduction of the reported correlation,clustering, and regression analyses from the retained paper-level analyticaldataset onward.

References

1. Ahlgren P, Jarneving B, Rousseau R (2003) Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. J Am Soc Inform Sci Technol 54(6):550–560. https://doi.org/10.1002/asi.10242
2. Aksnes DW, Karlstrøm H, Piro FN (2026) Self-reported and bibliometric interdisciplinarity measures rarely correspond: A survey-based comparative analysis of indicators and researcher perceptions. Scientometrics 131(1):189–208. https://doi.org/10.1007/s11192-025-05525-6
3. Archambault É, Beauchesne OH, Caruso J (2011) Towards a multilingual, comprehensive and open scientific journal ontology. In B. Noyons, P. Ngulube, & J. Leta (Eds.), Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics (Vol. 1, pp. 66–77). International Society for Scientometrics and Informetrics
4. Aria M, Cuccurullo C (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetrics 11(4):959–975. https://doi.org/10.1016/j.joi.2017.08.007
5. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech: Theory Exp., 10, Article P10008 https://doi.org/10.1088/1742-5468/2008/10/P10008
6. Bornmann L, Daniel H-D (2008) What do citation counts measure? A review of studies on citing behavior. J Doc 64(1):45–80. https://doi.org/10.1108/00220410810844150
7. Bornmann L, Leydesdorff L, Wang J (2014) How to improve the prediction based on citation impact percentiles for years shortly after the publication date? J Informetrics 8(1):175–180. https://doi.org/10.1016/j.joi.2013.11.005
8. Burt RS (1992) Structural holes: The social structure of competition. Harvard University Press
9. Carley S, Porter AL (2012) A forward diversity index. Scientometrics 90(2):407–427. https://doi.org/10.1007/s11192-011-0528-1
10. Carley SF, Kwon S, Porter AL, Youtie JL (2019) The relationship between forward and backward diversity in CORE datasets. Scientometrics 120(3):961–974. https://doi.org/10.1007/s11192-019-03163-3
11. Chakraborty T (2018) Role of interdisciplinarity in computer sciences: Quantification, impact and life trajectory. Scientometrics 114(3):1011–1029. https://doi.org/10.1007/s11192-017-2628-z

12. Chang Y-W, Huang M-H (2012) A study of the evolution of interdisciplinarity in library and information science: Using three bibliometric methods. J Am Soc Inform Sci Technol 63(1):22–33. https://doi.org/10.1002/asi.21649
13. Chen C (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inform Sci Technol 57(3):359–377. https://doi.org/10.1002/asi.20317
14. Chen C (2018) Cascading citation expansion. J Inform Sci Theory Pract 6(2):6–23. https://doi.org/10.1633/JISTaP.2018.6.2.1
15. Chen S, Arsenault C, Larivière V (2015) Are top-cited papers more interdisciplinary? J Informetrics 9(4):1034–1046. https://doi.org/10.1016/j.joi.2015.09.003
16. Chen S, Qiu J, Arsenault C, Larivière V (2021) Exploring the interdisciplinarity patterns of highly cited papers. J Informetrics 15(1) Article 101124. https://doi.org/10.1016/j.joi.2020.101124
17. Chen S, Song Y, Qiu J, Larivière V (2021) The effect of interdisciplinary components' citation intensity on scientific impact. Libr Hi Tech 39(4):1084–1096. https://doi.org/10.1108/LHT-01-2021-0038
18. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011a) An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. J Informetrics 5(1):146–166. https://doi.org/10.1016/j.joi.2010.10.002
19. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011b) Science mapping software tools: Review, analysis, and cooperative study among tools. J Am Soc Inform Sci Technol 62(7):1382–1402. https://doi.org/10.1002/asi.21525
20. Fleming L (2001) Recombinant uncertainty in technological search. Manage Sci 47(1):117–132. https://doi.org/10.1287/mnsc.47.1.117.10671
21. Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35–41. https://doi.org/10.2307/3033543
22. Gibbons M, Limoges C, Nowotny H, Schwartzman S, Scott P, Trow M (1994) The new production of knowledge: The dynamics of science and research in contemporary societies. SAGE
23. Hou JH, Li H, Zhang Y (2024) Influence of interdisciplinarity of scientific papers on the durability of citation diffusion: A perspective from citation discontinuance. J Informetrics 18(3) Article 101525. https://doi.org/10.1016/j.joi.2024.101525
24. Hsieh CM (2011) Explicitly searching for useful inventions: Dynamic relatedness and the costs of connecting versus synthesizing. Scientometrics 86(2):381–404. https://doi.org/10.1007/s11192-010-0290-9
25. Katz JS, Hicks D (1997) How much is a collaboration worth? A calibrated bibliometric model. Scientometrics 40(3):541–554. https://doi.org/10.1007/BF02459299
26. Kohus Z, Lendvai GF, Vizvári D, Demeter M, Soós S, Kerekes BB, Lukács E (2025) Exploring the collaboration networks between highly cited researchers in highly cited papers. Scientometrics 130(11):6513–6540. https://doi.org/10.1007/s11192-025-05443-7

27. Kong M, Zhang Y, Sheng L, Hong K (2025) Citation structural diversity: A novel metric combining structure and semantics for literature evaluation. Scientometrics 130(7):4027–4060. https://doi.org/10.1007/s11192-025-05356-5
28. Larivière V, Gingras Y (2010) On the relationship between interdisciplinarity and scientific impact. J Am Soc Inform Sci Technol 61(1):126–131. https://doi.org/10.1002/asi.21226
29. Larivière V, Gingras Y, Sugimoto CR, Tsou A (2015) Team size matters: Collaboration and scientific impact since 1900. J Association Inform Sci Technol 66(7):1323–1332. https://doi.org/10.1002/asi.23266
30. Levitt JM, Thelwall M (2008) Is multidisciplinary research more highly cited? A macrolevel study. J Am Soc Inform Sci Technol 59(12):1973–1984. https://doi.org/10.1002/asi.20914
31. Leydesdorff L (2007a) Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. J Am Soc Inform Sci Technol 58(9):1303–1319. https://doi.org/10.1002/asi.20614
32. Leydesdorff L (2007b) Mapping interdisciplinarity at the interfaces between the Science Citation Index and the Social Sciences Citation Index. Scientometrics 71(3):391–405. https://doi.org/10.1007/s11192-007-1694-z
33. Leydesdorff L, Scientometrics (2018) 116(3), 2113–2121. https://doi.org/10.1007/s11192-018-2810-y
34. Leydesdorff L, Rafols I (2011a) Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. J Informetrics 5(1):87–100. https://doi.org/10.1016/j.joi.2010.09.002
35. Leydesdorff L, Rafols I (2011b) Local emergence and global diffusion of research technologies: An exploration of patterns of network formation. J Am Soc Inform Sci Technol 62(5):846–860. https://doi.org/10.1002/asi.21509
36. Leydesdorff L, Wagner CS, Bornmann L (2018) Betweenness and diversity in journal citation networks as measures of interdisciplinarity: A tribute to Eugene Garfield. Scientometrics 114(2):567–592. https://doi.org/10.1007/s11192-017-2528-2
37. Leydesdorff L, Wagner CS, Bornmann L (2019a) Diversity measurement: Steps towards the measurement of interdisciplinarity? J Informetrics 13(3):904–905. https://doi.org/10.1016/j.joi.2019.03.016
38. Leydesdorff L, Wagner CS, Bornmann L (2019b) Interdisciplinarity as diversity in citation patterns among journals: Rao–Stirling diversity, relative variety, and the Gini coefficient. J Informetrics 13(1):255–269. https://doi.org/10.1016/j.joi.2018.12.006
39. Meho LI, Rogers Y (2008) Citation counting, citation ranking, and h-index of human–computer interaction researchers: A comparison of Scopus and Web of Science. J Am Soc Inform Sci Technol 59(11):1711–1726. https://doi.org/10.1002/asi.20874
40. Moed HF (2010) Measuring contextual citation impact of scientific journals. J Informetrics 4(3):265–277. https://doi.org/10.1016/j.joi.2010.01.002
41. Mongeon P, Paul-Hus A (2016) The journal coverage of Web of Science and Scopus: A comparative analysis. Scientometrics 106(1):213–228. https://doi.org/10.1007/s11192-015-1765-5

42. Mugabushaka A-M, Kyriakou A, Papazoglou T (2016) Bibliometric indicators of interdisciplinarity: The potential of the Leinster–Cobbold diversity indices to study disciplinary diversity. Scientometrics 107(2):593–607. https://doi.org/10.1007/s11192-016-1865-x
43. Mukhtar MF, Anuar SHH, Abas ZA, Hussin MSF, Sulaiman SA (2024) Bibliometric analysis of the global trend in centrality measures. Int J Acad Res Bus Social Sci 14(9):1078–1098. https://doi.org/10.6007/IJARBSS/v14-i9/22925
44. Mutz R (2022) Diversity and interdisciplinarity: Should variety, balance and disparity be combined as a product or better as a sum? An information-theoretical and statistical estimation approach. Scientometrics 127(12):7397–7414. https://doi.org/10.1007/s11192-022-04336-3
45. Perianes-Rodriguez A, Waltman L, van Eck NJ (2016) Constructing bibliometric networks: A comparison between full and fractional counting. J Informetrics 10(4):1178–1195. https://doi.org/10.1016/j.joi.2016.10.006
46. Peters HPF, van Raan AFJ (1994) On determinants of citation scores: A case study in chemical engineering. J Am Soc Inform Sci 45(1):39–49
47. Porter AL, Chubin DE (1985) An indicator of cross-disciplinary research. Scientometrics 8(3–4):161–176. https://doi.org/10.1007/BF02016934
48. Porter AL, Rafols I (2009) Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics 81(3):719–745. https://doi.org/10.1007/s11192-008-2197-2
49. Rafols I, Leydesdorff L, O'Hare A, Nightingale P, Stirling A (2012) How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business and management. Res Policy 41(7):1262–1282. https://doi.org/10.1016/j.respol.2012.03.015
50. Rafols I, Meyer M (2010) Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. Scientometrics 82(2):263–287. https://doi.org/10.1007/s11192-009-0041-y
51. Ricotta C, Szeidl L (2006) Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. Theor Popul Biol 70(3):237–243. https://doi.org/10.1016/j.tpb.2006.06.003
52. Rousseau R (2018) The repeat rate: From Hirschman to Stirling. Scientometrics 116(1):645–653. https://doi.org/10.1007/s11192-018-2724-8
53. Seglen PO (1992) The skewness of science. J Am Soc Inform Sci 43(9):628–638. https://doi.org/10.1002/(SICI)1097-4571)43:9%3C628::AID-ASI5%3E3.0.CO;2-0
54. Tahamtan I, Afshar AS, Ahamdzadeh K (2016) Factors affecting number of citations: A comprehensive review of the literature. Scientometrics 107(3):1195–1225. https://doi.org/10.1007/s11192-016-1889-2
55. Tang L (2013) Does birds of a feather flock together matter? Evidence from a longitudinal study on US–China scientific collaboration. J Informetrics 7(2):330–344. https://doi.org/10.1016/j.joi.2012.11.010

56. van den Besselaar P (2025) Circling around interdisciplinarity. Scientometrics 130(6):3293–3311. https://doi.org/10.1007/s11192-025-05346-7
57. van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84(2):523–538. https://doi.org/10.1007/s11192-009-0146-3
58. Vanclay JK (2012) Impact factor: Outdated artefact or stepping-stone to journal certification? Scientometrics 92(2):211–238. https://doi.org/10.1007/s11192-011-0561-0
59. Vieira ES, Gomes J (2010) Citations to scientific articles: Its distribution and dependence on the article features. J Informetrics 4(1):1–13. https://doi.org/10.1016/j.joi.2009.06.002
60. Wagner CS, Roessner JD, Bobb K, Klein JT, Boyack KW, Keyton J, Rafols I, Börner K (2011) Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. J Informetrics 5(1):14–26. https://doi.org/10.1016/j.joi.2010.06.004
61. Waltman L, van Eck NJ, Noyons ECM (2010) A unified approach to mapping and clustering of bibliometric networks. J Informetrics 4(4):629–635. https://doi.org/10.1016/j.joi.2010.07.002
62. Waltman L, van Eck NJ, van Leeuwen TN, Visser MS, van Raan AFJ (2011) Towards a new crown indicator: Some theoretical considerations. J Informetrics 5(1):37–47. https://doi.org/10.1016/j.joi.2010.08.001
63. Wang J, Thijs B, Glänzel W (2015) Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. PLoS ONE 10(5) Article e0127298. https://doi.org/10.1371/journal.pone.0127298
64. Wang J, Veugelers R, Stephan P (2017) Bias against novelty in science: A cautionary tale for users of bibliometric indicators. Res Policy 46(8):1416–1436. https://doi.org/10.1016/j.respol.2017.06.006
65. Yegros-Yegros A, Rafols I, D'Este P (2015) Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. PLoS ONE 10(8) Article e0135095. https://doi.org/10.1371/journal.pone.0135095
66. Zhang L, Rousseau R, Glänzel W (2016) Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. J Association Inform Sci Technol 67(5):1257–1265. https://doi.org/10.1002/asi.23487

Figures

Image not available with this version

Figure 1

Legend not included with this version

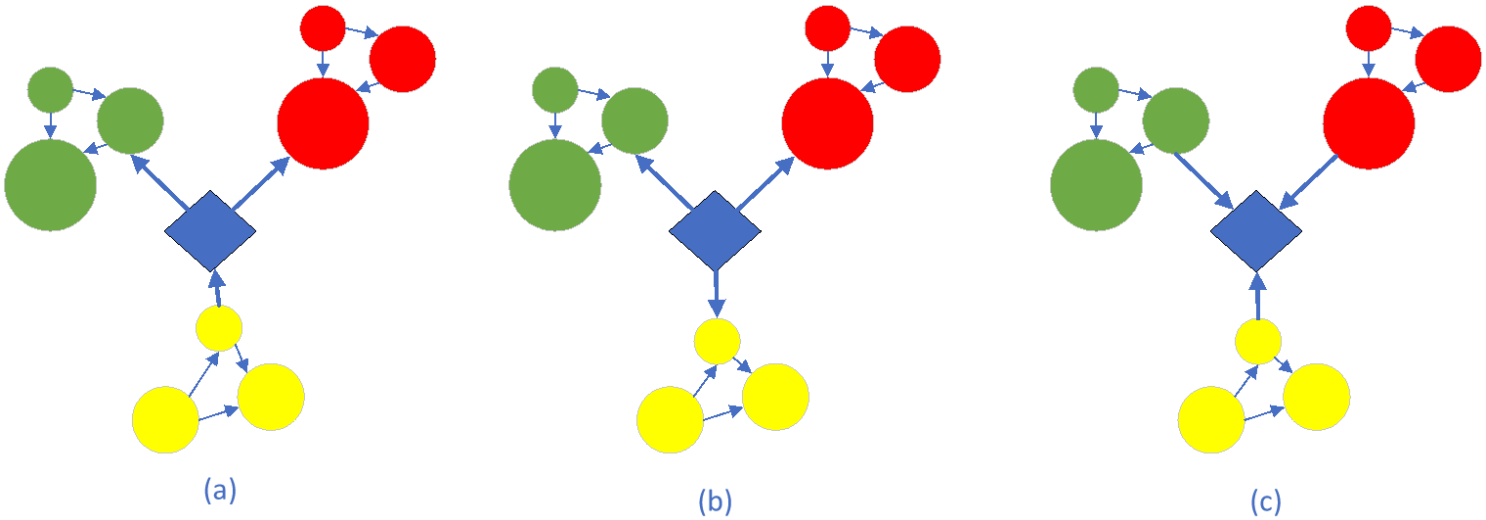


Figure 2

Structural configurations associated with directed and undirected betweenness centrality in citation networks.

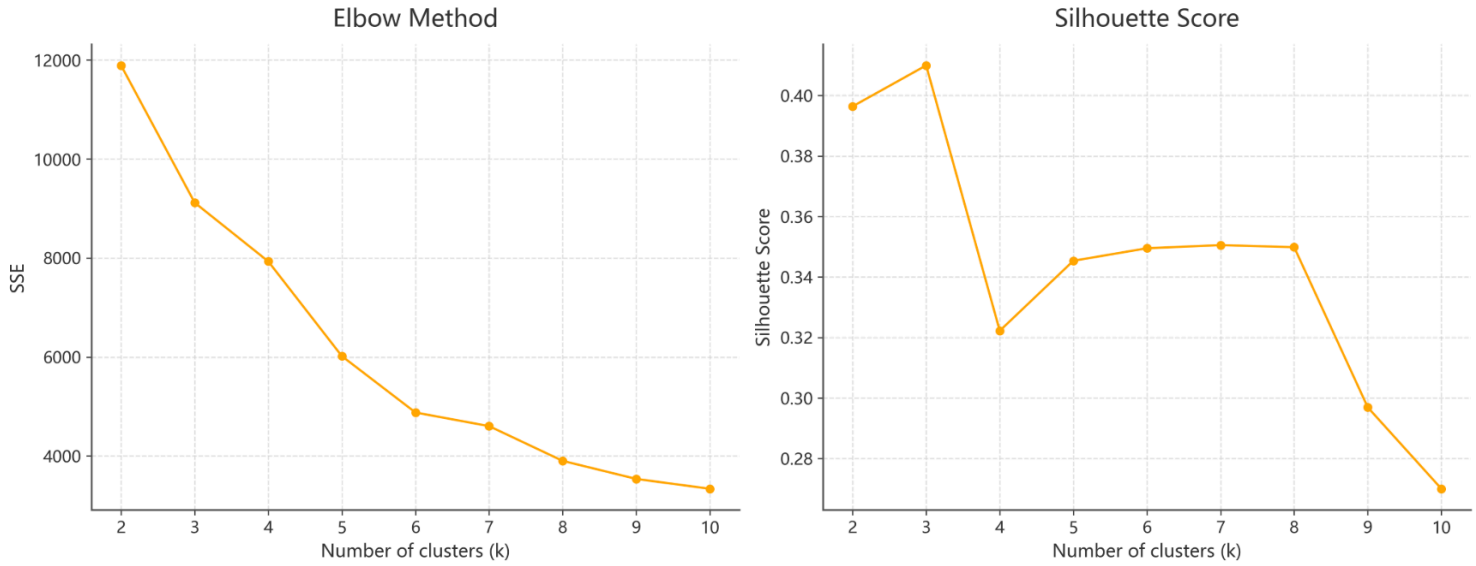


Figure 3

Results of elbow method for cluster number selection and silhouette coefficient analysis.

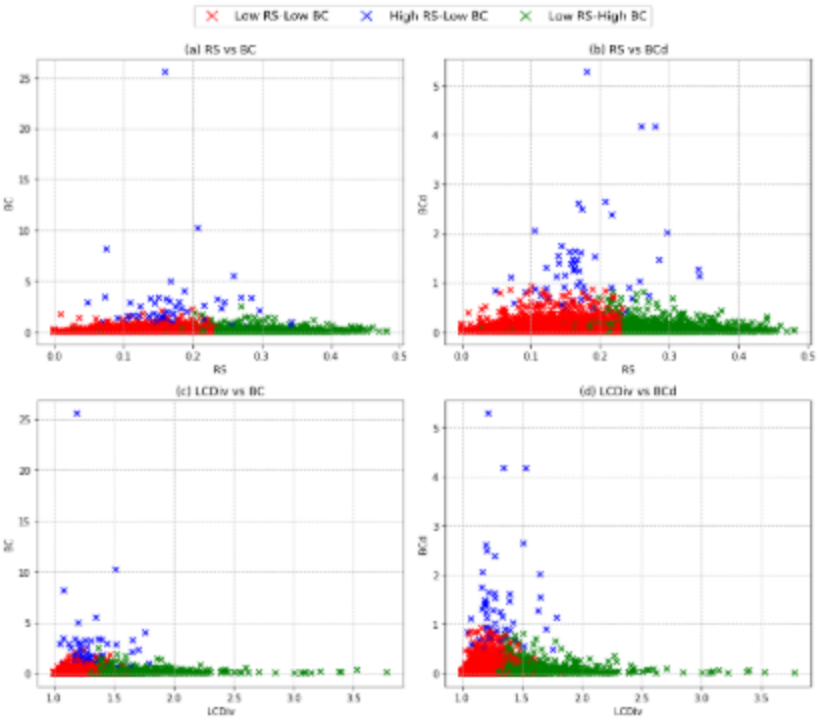


Figure 4

The cluster results of K-Means for diversity and betweenness

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Sl.zip