Positioning knowledge: schools of thought and new knowledge creation

S. Phineas Upham · Lori Rosenkopf · Lyle H. Ungar

Abstract Cohesive intellectual communities called "schools of thought" can provide powerful benefits to those developing new knowledge, but can also constrain them. We examine how developers of new knowledge position themselves within and between schools of thought, and how this affects their impact. Looking at the micro and macro fields of management publications from 1956 to 2002 with an extensive dataset of 113,000+ articles from 41 top journals, we explore the dynamics of knowledge positioning for management scholars. We find that it is significantly beneficial for new knowledge to be a part of a school of thought, and that within a school of thought new knowledge has more impact if it is in the intellectual semi-periphery of the school.

Keywords Innovation · Management · Schools of thought · Clustering

Introduction

New knowledge developers work in an intellectual and scientific landscape with social structures that shape their actions. In doing so they navigate within, between and among intellectual "schools of thought" that deeply affect their contributions (Kuhn 1962; Small 2003). Previous research has explored that new knowledge has more impact when it is well situated in an existing school of thought and/or when it incorporates outside knowledge (Trajtenberg 1990; Fleming 2001). Drawing from previous research, we develop a strategic understanding of the positioning incentives for researchers creating new knowledge in the social science field of management. We contribute to the existing

S. P. Upham (✉) · L. Rosenkopf
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6370, USA
e-mail: uphams@wharton.upenn.edu; phin.upham.wg05@wharton.upenn.edu

L. Rosenkopf
e-mail: rosenkopf@wharton.upenn.edu

L. H. Ungar
CIS Department, University of Pennsylvania, Philadelphia, PA 19104-6309, USA
e-mail: ungar@cis.upenn.edu

literature in this field by quantitatively testing how the position of researchers within or outside schools of thought relates to the impact of their contributions. We then test how positioning within schools of thought and a researcher's experience in multiple schools of thought affects performance. Our tests widen the applicability of the dynamics of knowledge positioning and put forth a clearer understanding of their potential for future management research.

First, we wish to understand the systematic mechanisms by which schools of thought engage and incentivise individual developers of new knowledge. Second, we wish to use citation data to explore how schools of thought encourage both local and distant search, integrating the findings of recent innovation literature with our theory. Third, we wish to quantitatively explore the consequences on intellectual impact of new knowledge positioning both within and between schools of thought by examining the intellectual organization of the field of management.

An epistemic community or mini-paradigm, often called a school of thought, is a socially constructed and informal community of researchers who build on each other's ideas and share similar interests and who consequentially share patterns of citation in their work (Crane 1972, 1980). Research on schools of thought in citation analysis has been limited but very suggestive (Small and Crane 1979; Ennis 1992). While researchers have identified and delineated schools of thought in various fields ranging from management of information science to theoretical high energy physics, they have rarely looked at how these schools function or how they affect new knowledge performance (Crane 1980; Culnan 1987).

Schools of thought can powerfully influence the process of individual knowledge creation in at least three ways. First, the socially agreed upon boundaries of schools of thought influence how developers of new knowledge explicitly think about and position themselves within their field; thus there is an explicit strategic dimension to knowledge positioning (Castro et al. 2001). Second, schools of thought are labels for dense social networks that distribute information through personal ties, conferences, conversations, etc. They deeply influence the knowledge developer's searches, access to and ease of finding information and, in aggregate, an individual researcher's knowledge stocks and resulting new knowledge contributions (Doreian 1988; Moody 2001). Third, schools of thought represent mental paradigms that unconsciously influence authors' view of the boundaries of their intellectual world (Crane 1980; Pfeffer 1993; Small 2003).

Seminal work by Thomas Kuhn in *The Structure of Scientific Revolutions* argues that groups of researchers with a coherent scientific and intellectual world view and a shared set of questions and methodologies are a fundamental part of intellectual thought and rigor (Kuhn 1962). In her description of the field of management of information science, Culnan (1986) describes invisible colleges (which we call schools of thought) as inherent to innovative research:

> Researchers in any discipline tend to cluster into informal networks, or "invisible colleges," which focus on common problems in common ways…. The history of the exchanges between members of these subgroups in a discipline describes the intellectual history of the field. (p. 156)

It is perhaps to be expected that loosely defined groups of like-minded researchers within academic fields will tend to study similar questions with overlapping methodologies. This holds particularly true in the social sciences, where methodology and goals are fragmented and schools of thought are prevalent.

Pfeffer (1993), discussing the field of organizational theory, argues that researchers need a strong paradigm to direct and organize the advancement of knowledge through agreed upon goals and vocabulary so that their work can incrementally build on each other's. He argues that a level of external intellectual "borrowing" from outside one's paradigm causes a lack of coherence in a field. As he puts it, "consensus is a critical precondition for scientific advancement (p. 600)." At the opposite extreme these same forces can socially embed new knowledge builders so that they are structurally disinclined to try to communicate or learn valuable ideas from those outside of their circle—which can lead to intellectual isolation and stagnation.

To analyze strategies for maximizing effectiveness in an ideological landscape, we build on the ideas of positioning theorists (Hotelling 1929; Downs 1957; McGann 2002) and search theorists (March and Simon 1958; Levinthal 1997; Cohen et al. 2000). Positioning theorists seek to explain the behavior of agents who try to position themselves to appeal to a maximum number of consumers. Search theorists analyze the consequences of local and distant search strategies on outcomes such as innovation. Using these approaches, we show how the implicit incentive structures created by schools of thought affect the impact of new knowledge.

The field of study examined in this paper, micro and macro management strategy, has been characterized throughout its history by competing schools of thought offering different and sometimes mutually exclusive causal explanations for business phenomena and identifying the underlying drivers of firm behavior (Barney 1986; Mintzberg 1994). We use a database of 113,000+ papers from 41 top management journals from 1956 to 2002, which covers the modern life of management studies, to explore what affects the impact of the publications of management scholars.

Theoretical development

Positioning

Hotelling (1929) described a competitive game that was later adapted to explore strategic positioning in ideological space. In the game, two hypothetical newspaper sellers, competing for readers who are distributed evenly along "Main Street," can set up their stand anywhere in town. Assuming that, for the same price, customers will buy the closest newspaper if one newspaper seller were to position himself anywhere but the center of Main Street, the other would position himself a little closer to the center point and gain the majority of the customers. Thus both sellers end up converging at the midpoint of Main Street. In this model each player explicitly takes into consideration the moves of other players when acting. A political variant of this principle was applied to the ideological landscape of voters to explain the middle-of-the-road views generally espoused by candidates of major political parties (Downs 1957). Given an even distribution of voters, politicians in a two-party race, in order to appeal to the greatest number of voters, will converge to mainstream positions where they maximize their access to voters in a Nash equilibrium. More generally, Downs' work finds that in an intellectual landscape, given Hotelling's assumptions, a central position closest to the greatest number of consumers is optimal.

Further extending Hotelling's theory, Downs (1957) also challenged his assumption of a "normal" distribution of consumers, and other researchers have further extended Hotelling's game by including multiple players or additional consumer (voter) or competitor

(candidate) entry (Krishna 2001; McGann 2002). The Hotelling–Downs framework has been usefully applied broadly to such areas as marketing and brand positioning (Choi and Coughlan 2004), news coverage (Gasper 2005) and simulations (Marks and Albers 2001).

The principles of acting to maximize intellectual proximity to the greatest number of consumers and of employing a dynamic strategy that takes others' moves into consideration provide a powerful framework for analyzing new knowledge development. The strategic positioning of new knowledge developers explored in this paper resembles a very complex multi-player version of the Hotelling problem—one, however, that differs along two key dimensions. First, we attempt to include the dimension of schools of thought, which makes the landscape "clustered" and has profound consequences for the application of Hotelling's strategic principles. Second, in our framework the candidates and the voters are flip sides of the same coin (each producer of new knowledge is also a consumer of new knowledge in our knowledge landscape), simultaneously competing with and supporting each other. In the context of new knowledge development, the positioning theory model significantly under-estimates the importance of schools of thought in the knowledge landscape.

Search—near and distant

The tension between local and distant search has been explored by juxtaposing the strategies of exploration and exploitation (March 1991). Exploration (involving distant search) is an attempt to add value by finding a new opportunity, while exploitation (involving local search) involves building on existing resources or knowledge in an attempt to extract value. Search strategies have significant effects on the development and structure of their landscape (Levinthal 1991, 1997, 1998). In the long term, exploration does produce benefits, but it must be "paid for" by exploitation (Barnett and Sorenson 2002). In the shorter term and from a research perspective, interdisciplinary research—explorative by its very nature—increases the difficulty of publishing papers, training graduate students, or receiving funding for a subject (Birnbaum 1981b).

However, since the dichotomy between exploration and exploitation is always operationally dependent on the choice of boundary, the way in which we delineate boundaries will determine whether we define a search as near or distant. A number of recent studies examine the tradeoffs of different search strategies, but they define their relevant boundaries of analysis in different ways. Katila and Ahuja (2002) argue that firms can differentiate themselves by creatively and meaningfully reusing old technology to create new knowledge as well as by finding new technologies to achieve breakthroughs. Nerkar (2003) sees firms successfully choosing between recent, cutting-edge knowledge and knowledge that integrates understandings developed across time spans. They thus differentiate between middle-level and radical exploration. Herbert Simon famously looked at satisfying to explain the tension between usefulness and truthfulness—that at a certain point one stops looking for a better answer if the one has an adequate one (March and Simon 1958; March and Shapira 1987; March 1991). Over all, these researchers find that search patterns have profound effects on knowledge creation. This previous research is tied together by the core idea that when new knowledge is developed, some boundary, internal or external, is extended or challenged.

Although theories of local search often take for granted that there is local and distant knowledge, what makes knowledge accessible and close or inaccessible and far is left unexplored from an intellectual, psychological, and resource point of view. Delineating these boundaries is indeed a complex process. We argue that a key part of this intellectual boundary-shaping in knowledge development can be found in socially constructed schools

of thought. Schools of thought are a key factor for new knowledge developers in perceiving information as near or far, and we believe that an appropriate intermediate spanning of boundaries between schools of thought is an important potential driver of knowledge creation. By providing an explanation for why new knowledge is close or far, we believe we move towards generalizing and integrating previous research on boundary spanning in knowledge creation.
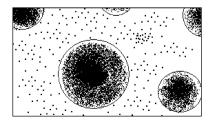
Hypotheses

In this section we extend three hypotheses about the relationship between the position of new knowledge within a school of thought and the subsequent impact of that knowledge. First, we argue that being part of a school of thought increases intellectual impact. Indeed, most new knowledge is not in any school of thought at all. Second, we argue that a position near the semi-periphery of a school of thought (but still within the school) leads to greater overall impact of new knowledge. We also find that two strong and competing hypotheses are supported by existing theory. In our third hypothesis, we argue that over time an author's tendencies to explore between schools of thought, i.e., whether a knowledge developer tends to explore many different paradigms or specialize within a paradigm, will affect the impact of his or her ideas (Fig. 1).

Hypothesis 1   New knowledge has more impact if it is within a school of thought than if it is not.

Given that new knowledge is part of a school of thought, within that school of thought it can also be at the "core" or "periphery." New knowledge that is core to its school is consistent with the rest of the school in its sources of ideas; new knowledge that is peripheral draws on knowledge that differs in some significant way, usually uncommon knowledge or knowledge from outside the school. The level of compliance of new knowledge with a particular school of thought depends on the number of criteria it satisfies in order to identify with that school. Its association with the school weakens as it draws more and more from notions that may not be firmly help by members of the school or our the central focus of other schools. In this way, for the purpose of research and exploration of knowledge creation, a theoretical distinct line can be created to separate schools from one another.

We believe that while new knowledge creators receive benefits from being in a school, membership can also be constraining if they blind themselves to good ideas outside that school. Specifically, we believe successful research usually draws from its own school and also a few core ideas from one or perhaps two other schools, synthesizing knowledge that is near and distant. In the area of technological innovation, studies have shown that patents

Fig. 1 Knowledge can be positioned within or outside of a school of thought

which combine technologies from different patent classes tend to have more diverse, and potentially greater, impact (Trajtenberg 1990).

Previous research in local and distant search suggests that new knowledge which is core to its school of thought is likely to be intellectually embedded within that school and have less impact outside of that school, and new knowledge located closer to the periphery of its school tends to more explicitly engage ideas meaningful both to its own field and to audiences beyond its field (McCain 1986, 1987). Even within a school, new knowledge that remains too close to the core ideas of a school and does not search for and use new ideas is less likely to have innovative impact (Meyer and Zucker 1989; Fleming 2001; Fleming and Sorenson 2001) and is thus less likely to influence others and be more highly cited by those within its school (Rosenkopf and Nerkar 2001). At the same time, knowledge too distant from the core might not reap as many of the benefits of membership.

This would imply that the relationship between a position at the center and periphery of a school is curvilinear—that a position at the semi-periphery of a school, straddling more than one school or reaching beyond one's school, would tend to draw the largest audience for new knowledge and receive the most overall citations. Such boundary-spanning research is more likely to draw fresh, interesting outside work into a school, which would potentially result in more impact (Fig. 2).

Hypothesis 2 A position toward the intellectual semi-periphery of a school of thought results in greater impact than a position at the center or periphery of a school of thought.

We now wish to take into consideration not only the characteristics of the new knowledge but the exploratory tendencies of creators of new knowledge as well. Over time a researcher has a tendency to either explore a diversity of knowledge domains or to focus on one. For example, in the field of economics, Oliver Williamson has published most of his works in one school of thought—transaction-cost economics (Williamson 1975, 1979)—and he is intellectually central to that school. At the same time, some great new knowledge producers systematically publish works in different fields and are enormously and broadly influential. James March, a peripatetic management scholar who studies organizations, for example, also publishes widely in many schools of thought including decision theory, organizational learning, and adaptation (March and Simon 1958; March and Shapira 1987; March 1991).

Ron Burt has argued that "theory developers" focus on deepening and refining theory within one field, while "theory synthesizers" span organizational and knowledge boundaries to combine knowledge in fresh and innovative ways (Burt, unpublished).

The advantages of remaining within a community are well-established in network theory—trust, reputation, and learning, among other social and intellectual benefits, were
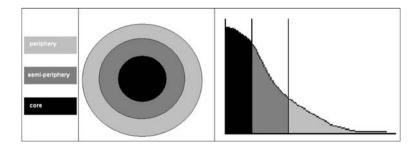

Fig. 2 Knowledge can be positioned in the periphery, semi-periphery, or core of a school of thought

discussed previously. Further, learning the norms and knowledge structure of a school is an investment that may have to be paid again if one changes schools. Simultaneously, from the perspective of influence of ideas, one could argue that remaining in a school of thought may cause lesser impact after an initial introduction of that idea, whereas a specific idea can be made new many times if transported to different fields (Amir 1985; Adner and Levinthal 2000).
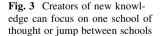
New knowledge developed without fresh inputs and without looking further than itself would be in danger of missing important insights, or as Rosenkopf and Nerkar (2001) put it, might "lead firms to develop 'core rigidities' or fall into 'competence traps'" (p. 288). An innovator may try to avoid falling into such traps by exploring numerous schools of thought seriously, interacting with a broad array of knowledge in multiple fields. Francis Crick, known for his work on the structure of DNA, seems to believe such theory-hopping is essential for creative insight when he argues that "professional [scientists] know that they have to produce theory after theory before they are likely to hit the jackpot. The very process of abandoning one theory for another gives them a degree of critical detachment that is almost essential if they are to succeed."
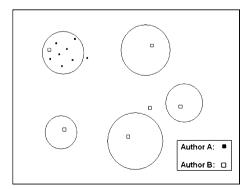
We believe that the interests and history of an author, whether eclectic or focused, make a difference in his or her impact and readership. We find that two strong and competing hypotheses are supported by existing theory. On one hand, authors who publish widely and are involved in the intellectual pursuit of multiple areas lend themselves in their eclectic pursuits to the benefits of knowledge combination and thus gain the advantages of synthesizing between schools of thought (Fleming 2001). On the other hand, depth of expertise lends itself to more incremental new knowledge production, gaining the advantages of a closed cluster, so perhaps focusing on one area can be more beneficial to a new knowledge creator (Birnbaum 1981a, b; Fig. 3).

Hypothesis 3a   New knowledge created by those who actively engage in multiple schools of thought over time has greater impact.

Hypothesis 3b   New knowledge created by those who concentrate on a very few schools of thought over time has greater impact.

Central to our analysis is the idea that new knowledge development is often a result of the constant dynamic tension within the area of a school of thought between the search for synthesis—i.e., recombination, introducing new ideas—and specialization—i.e., developing the world view of the field, deepening the core methods or proposition of a paradigm—and moreover that scholars face strategic choices in positioning themselves within and

Fig. 3 Creators of new knowledge can focus on one school of thought or jump between schools

among schools of thought. The content of the new knowledge is thereby enriched by the dialectic between the diversity of its sources as well as the efforts of the researcher to ground the knowledge by defining its scope and developing its focus. In this constant negotiation of tensions, while multiplicity of schools contributes to the knowledge, the number of schools is crucial to maintain the specialization and focus of the research.

Methods

Clustering

Co-citation analysis has been used to systematically map and examine the network structures of research papers or patents since Small and Griffith introduced the first computerized method to accomplish this in the 1970s. Early on, Small and Crane (1979) used large-scale clustering techniques to isolate and identify the structure of scientific disciplines. They thus matched a potent methodology with a coherent intellectual explanation for the "clusters" that Small and Griffith observed during earlier research. Increasingly sophisticated methods have been developed in technical and scientific fields to cluster and analyze high-dimensional similarities in enormous databases. These methods take advantage of greater computational power and the potential for multiple passes over the data during clustering (Popescul et al. 2000; Pantel and Lin 2002; Kandylas 2005; Kandylas et al. 2005). To find similar clusters of papers in our data we have drawn from advances in Internet search and genetic data mining to construct a clustering methodology called StrEMer that is substantially faster than and of comparable quality to past search and mining methods for estimating high-dimensional similarity clusters and takes into consideration our theoretical model for schools of thought.

Most current popular clustering programs in computer science search and genetic data mining assume that clusters over time are static (Popescul et al. 2000). This is a tolerable simplifying assumption for short periods in stable environments but more troubling for data over time in a dynamic environment where we must consider emerging clusters, merging clusters, and dying clusters. To address this we build an iterated "overlapping" clustering methodology into our algorithm that re-clusters in overlapping 10-year blocks, stepping forward by 1 year at a time. Therefore the elements in year 1990 would be clustered based only on the elements in 1980–1990, papers in year 1991 would be clustered based only on 1981–1991, and so on. Thus, the cluster structure of the data in 1975 can be very different from that of 1995. This allows for new clusters to be created and for existing clusters to merge or wither away, while the overlapping 10-year time span enforces some continuity on the clusters over time. Further, unlike less efficient iterated programs such as k-means and expected maximization algorithms, our StrEMer approach allows us to set criteria that give bounds on the error of single-pass algorithms. The clustering by committee (CBC) approach allows for background clusters but does not have a clear criterion being optimized and overparameterizes the algorithm for our purposes (Pantel and Lin 2002; Kandylas 2005).

The StrEMer clustering algorithm maximizes the coherence of the clusters as specified by an objective function measure. Objects to be clustered are initially assigned to clusters by a random procedure, and maximization is achieved by reassigning each object to the existing cluster with the most similar centroid (which is the sum of the similarity of the objects in a cluster). The whole process is repeated until the objective function cannot be substantially improved, considerably reducing path dependency based on initial clusters. Essentially our method plots citation (network) structure in high-dimensional space and

uses the objective function to minimize the distance from nearest centroid for all papers simultaneously. Distance represents the angle between the vector representing the paper and the centroid of the cluster it is placed in.

Our StrEMer program accomplishes clustering in three steps (repeated when doing iterated "overlapping" clustering). In Step I, we make a single pass over the data and construct several rough clusters. In Step II, we get a collection of high-quality clusters, called committees, based on the clusters obtained in Step I. These committees are tight and not similar to each other, which mean high inter-group similarity and low intra-group similarity. In Step III, we either assign each element to its most similar clusters, or add it to the residue list if it is not similar enough to any cluster.

To find our knowledge communities we use a clustering algorithm to identify clusters of like papers. Our methodology utilizes the structure of co-citations in paper bibliographies to group papers that are "similar" in the papers they cite. Essentially, therefore, we are comparing the citations of all papers to all other papers to find papers that use similar citations. Papers with similar citations are clustered together and a "distance" from the average feature characteristics (called "centroid") of that cluster is computer for each paper.

Technically, our algorithm constructs a similarity measure using a "feature vector" for each element to represent the citation patterns between this element and other elements. We then compute the similarity between two elements A and B (hypothetically illustrated below) using the cosine coefficient of their feature vectors

$$\cos(v) = \mathrm{sim}(A, B) = \frac{\sum\limits_{k} A_k \cdot B_k}{\sqrt{\sum\limits_{k} A_k^2 \cdot \sum\limits_{k} B_k^2}}.$$

We further define the similarity between an element e and a cluster c as the similarity between e and the cluster feature centroid of c: $\mathrm{sim}(e, c) = \mathrm{sim}(e, \mathrm{cen})$, where cen is the cluster feature centroid vector of c. Finally, the similarity between two clusters c_i and c_j is the similarity between their cluster feature centroid vectors: $\mathrm{sim}(c_i, c_j) = \mathrm{sim}(\mathrm{cen}_i, \mathrm{cen}_j)$, where cen_i and cen_j are centroid vectors of c_i and c_j respectively, and the centroid intuitively represents the "average" of the papers in the cluster. The distance from this centroid (how much each paper varied from the average of the cluster) comes to represent how distinct the paper is from the norms of its paper. Operationally, this means a given paper's citation structure is similar enough to be in the cluster but still has variation.

In the illustration below, W and X represent groups of papers that will become clusters, while Y and Z represent lone papers that will not be assigned to clusters. In this simplified example, the papers appear as points on the circle, extending from a common starting point—the origin, where similar papers are very near each other on the circumference of the circle. Another way of measuring this "closeness" is to look at the angle between each pair of papers. Clusters are identified by our method by searching for groups of papers with small angles between them (in a high dimensional space).

Data

In this paper we examine micro and macro management scholarship—a fast-moving, relatively young, and highly fragmented academic field, taught in business schools, which is known for its diversity of ideas and its vigorous schools of thought (Abrahamson 1996; Meyer 1999). To select our data we began with previous rankings of journals in

management. Taking the top lists that rank the most impactful journals including Coe and Weinstock (1969), Coe and Weinstock (1984), Sharplin and Mabry (1985), Podsakoff et al. (2005) and employing Johnson and Podsakoff's strategy (1994) to generate a master list from these disparate (and highly overlapping) lists by including any journal that is counted by more than one of these lists [excluding some of the ad hoc specialty journals added by Podsakoff et al. (2005)] we found 41 core journals in micro and macro management (see Table 2). Accessing the Thomson ISI database, we have collected complete sets of all articles and their citations for these 41 journals since 1956. This list of 41 journals includes within the field of management both the macro (which is heavily influenced by economics and sociology) and the micro (which is heavily influenced by psychology) specialties (see Table 1). The journals in the ISI database published since 1956 yield a total of 113,014 papers for analysis, the most complete database in this area compiled, to our knowledge. The papers include 2+ million citations in the bibliographies of these articles and 1.5 + million citations made to these articles.

Scientists and academics have formal shared discourse that in a very subtle way, and often without actual physical communities, becomes their primary source of self-identification and shapes their intellectual community (McCain 1986). The printed page or computer screen is where scientists interact. Before the Internet and long-distance telephones, scholars around the world rarely if ever met, interacting via the printed page. Scientists are not, in other words, intellectually defined—qua scientist—so much by with whom they meet but rather by with whom they agree intellectually and on whose innovations and work they build (Merton 1965). Of course, the intellectual/social structure in intellectual communities can be complemented by personal interactions at conferences and at universities. Nevertheless, scientists are often more "like," more related to, someone with whom they agree intellectually and have never met than they are to the researcher across the hall.

Citations are an unusually rich trail of formally codified relationships in scientific discourse, where the exchange of ideas (the content of the exchanges) is written down. These acknowledgments explicitly represent the exchange of ideas and knowledge, and represent for us intellectual impact and can be a proxy for whether the paper presented any new insights that were thought to be useful (Small 2003).

Bibliometrics, or the quantitative study of bibliography, uses as its unit of analysis the citations made from a published piece of work to other published pieces of work—and usually code these citations as a proxy for "impact" or influence (using the appropriate controls). Similarly, almost all studies using patents use patent citation as a proxy for patent success (Trajtenberg 1990). In paper-citation work it is commonly assumed that a

Table 1 Fields in which papers in our study fall (as specified by ISI classification of journal)

Citations received	Papers	Average cites/paper	Field
921,507	40,392	22.81	Psychology
357,996	42,193	8.48	Management
350,254	5,737	61.05	Behavior
181,531	27,063	6.71	Sociology and anthropology
51,692	9,081	5.69	No category
38,089	12,176	3.13	Economics
14,284	2,059	6.94	Political science and public administration

All citation counts as of 2002

Table 2 Journals and summary statistics

Citations received	Papers	Average cites/paper	Journals
297,182	7,775	38.22	*Journal of Personality and Social Psychology*
195,697	3,530	55.44	*Psychological Bulletin*
154,557	2,207	70.03	*Psychological Review*
131,085	14,509	9.03	*American Journal of Psychology*
93,836	8,039	11.67	*American Sociology Review*
91,083	4,636	19.65	*Journal of Applied Psychology*
64,297	4,470	14.38	*Management Science*
60,031	10,948	5.48	*American Journal of Sociology*
55,953	2,991	18.71	*Administrative Science Quarterly*
49,606	2,360	21.02	*Academy of Management Journal*
32,756	1,501	21.82	*Academy of Management Review*
31,554	1,283	24.59	*Strategic Management Journal*
27,664	8,076	3.43	*Social Forces*
23,448	877	26.74	*Organizational Behavior and Human Performance*
23,374	9,393	2.49	*Harvard Business Review*
22,498	1,717	13.1	*Journal of Vocational Behavior*
22,384	5,150	4.35	*Perspectives in Psychology*
20,594	2,502	8.23	*Human Relations*
17,836	1,116	15.98	*Organizational Behavior and Human Decisions*
14,596	5,132	2.84	*Industrial Labor Relations Review*
14,256	1,566	9.1	*Journal Human Resources*
12,022	1,321	9.1	*Journal of Conflict Resolution*
11,658	875	13.32	*Journal of Management*
9,082	1,185	7.66	*Journal of International Business Studies*
8,144	2,257	3.61	*Journal of Management Studies*
8,028	1,822	4.41	*California Management Review*
8,012	930	8.62	*Journal of Occupational Psychology*
7,846	904	8.68	*Decision Sciences*
7,221	1,868	3.87	*Sloan Management Review*
7,050	1,534	4.6	*Industrial Relations*
7,021	868	8.09	*Journal of Organizational Behavior*
5,633	1,598	3.53	*Journal of Business Research*
5,398	7,600	0.71	*Monthly Labor Review*
5,259	1,155	4.55	*Journal of Applied Behavior Science*
5,015	140	35.82	*Research in Organizational Behavior*
4,755	3,491	1.36	*Long Range Planning*
3,669	789	4.65	*Organizational Dynamics*
2,262	738	3.07	*Administration Society*
1,390	2,304	0.6	*Labor Law Journal*
757	693	1.09	*Journal of Collective Negotiations in the Public Sector*
590	1,114	0.53	*Arbitration Journal*

citation in a specific piece of work indicates intellectual influence and (which we code as "impact") on the published work and value to the citing author (Small 1978). This may not always be the case, though, as a well-known citation may be used simply to "represent" a point of view, or citations may be made for social reasons. For example, citations can be influenced by a social exchange phenomenon based on the notion of reciprocity. The rigors of the review process and the well-documented correlation between citations and other measures of influence, however, lead us to believe that citation metrics are at least a useful proxy for influence (Bayer and Folger 1966; Cole and Cole 1967; Osareh 1996). This methodology is generally attractive for its impersonality, objectivity, replicability, and scalability (Culnan 1986). Tables 1 and 2 summarize our data.

Variables

Our unit of analysis is individual papers. Our variables therefore describe characteristics of papers, both descriptions of them and also how they relate to other papers published at that time. As discussed, we characterize all research papers published in a top journals as a unit of analysis (containing new knowledge) since journals, using peer review, explicitly demand that a paper contribute to the aggregate knowledge of a field to be accepted. Further, well cited papers tend to have contributed unique ideas to a researchers work—so by only including cited papers in the regressions we are screening for papers published which had no differentiated impact.

Total citations ("impact")

To measure the impact of a paper we draw on a deep literature which argues that total citations can, in many cases, be a useful measure of the influence that paper has on other research. So for our dependent variable we measure the total impact of a paper as the number of citations it has received subsequent to its publication through 2002, controlling for exogenous characteristics below (Fig. 4b). See Clustering for details and methodological discussion.

Bibliography size

To control for the increasing number of citations made in papers, we control for the number of entries in a paper's bibliography. As seen in Fig. 4a, the average bibliography size increased steadily over the time spanned by our data.
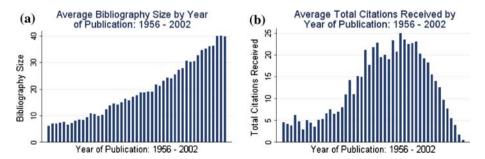

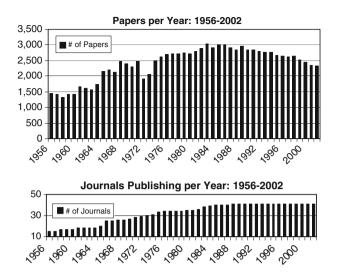**Fig. 4** Average bibliography size by year (*left*) and average total citations received by year (*right*)

Fig. 5 Papers per year (*top*) and journals published or covered by data-source in that year (*bottom*)

Year of publication

We created dummy variables for each year from 1956 to 2002.

Journal of publication

Because different journals tend to receive different citation rates as a result of varying size of readership and journal prestige, we included dummy variables for each of the 41 journals in our data (Table 2; Fig. 5).

Coauthorship (binary)

We include a binary variable for whether a paper has more than one author (1 if coauthored, 0 if single author). This helps control for the differences in the process of joint and individual knowledge production. Almost 34% of the papers in our data are coauthored.

Cluster (binary)

This binary variable is coded 1 if the paper is in a cluster of other similar papers (representing a "school of thought"—see Clustering for a technical discussion of how we characterize 'similar'—essentially it represents how different the citation structure of a papers is in high dimensional space, controlling for the independent variables) when it was published and 0 if the paper is not in a cluster when it was published. In summary, 'similar' papers are alike in that they have similar bibliographies as measured in high dimensional space, controlling for key paper characteristics. A little over half of our papers did not belong to any cluster (Table 5). See Clustering for technical details.
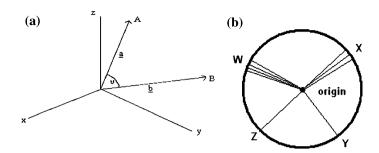
Fig. 6 **a** Distance between a paper and its centroid. **b** Representation of high dimensional clustering pattern

Distance

A paper is central to its cluster when its citation structure is very similar to the mean of the citation structure of all papers in its cluster (the centroid). This distance is measured by how similar or different a paper's bibliography is to the 'average' bibliography of the papers in its cluster, measured in high dimensional space and controlling for key paper characteristics. Papers that are typical of their clusters will have small distances. Papers that differ from the rest of the group by citing outside sources or by citing uncommonly cited papers will have larger distances. We measure distance as the angle υ of the vector representations between the citation structure of a paper and the centroid of its cluster (see Fig. 6a). Papers not in a cluster are not assigned a distance. See Clustering for technical details of how we measure and define distance.

Because clusters are generated with data from the year of the paper and the 9 previous years, $t^{(-10)}$ to $t^{(0)}$, distance represents the centrality of the paper to its cluster historically at the time of publication, not the centrality of the paper after publication.[1] We also include a second-order effect for distance to capture curvilinearity.

Diversity of publications

In order to find a proxy for an author's more general tendency to seek a diversity of knowledge and viewpoints we count the number of clusters in which an author has published throughout our dataset. We believe this will give an estimate of an author's tendency to stay in a school of thought or move between schools of thought. We use an entropy or diversity measure, defined as:

$$H = -\sum_{i=1}^{20} [p_i \cdot \ln(p_i)]$$

where p is the probability of being in state i. For papers with multiple authors we average their diversity measures.

Total number of papers published

We count an author's total number of publications, which is potentially correlated with the *Diversity of Publications* measure. For papers with multiple authors we sum their publication counts (Table 3).

---

[1] Reclusterings using data for 5 years after publication $t^{(0)}$ to $t^{(5)}$ and for the 10 years around the publication $t^{(-5)}$ to $t^{(5)}$ yielded comparable similar distance measures, implying that cluster centrality changes gradually.

Table 3 Correlation table and descriptive statistics ($n = 113{,}014$)

	Correlations						
	1	2	3	4	5	6	7
Total citations	–						
Bibliography size	0.3013*	–					
Coauthorship (binary)	0.1443*	0.3228*	–				
Cluster (binary)	0.1460*	0.3930*	0.3590*	–			
Distance	0.1803*	0.4070*	0.3634*	0.7347*	–		
Distance squared	0.1308*	0.2889*	0.2655*	0.5433*	0.9137*	–	
Diversity	0.1232*	0.1773*	0.3455*	0.3079*	0.3217*	0.2403*	–
Mean	13.619	20.855	0.340	0.436	0.161	0.052	0.751
SD	52.292	31.077	0.474	0.496	0.161	0.093	0.550
Min	0	0	0	0	0.001	0	0
Max	4,580	801	1	1	1.480	2.190	2.453

* Denotes significance at the $\alpha = .05$ level using the Bonferroni correction for multiple pairwise tests

Analysis

Regressions

The data used for our analyses are non-negative counts of the number of forward citations for papers published in 41 journals covering both micro and macro management (sometimes called organizational behavior and strategy, respectively). As with previous measures (Small and Crane 1979; Morris and Moore 2000; Ramos-Rodriguez and Ruiz-Navarro 2004) of publication citation counts, the data exhibit a variance in the number of citations larger than would be expected from a Poisson distribution. We considered using a simple negative binomial (NB) model to account for the excess variance; however, the numerous zero counts in our data further indicate that a zero-inflated negative binomial regression (ZINBR) model would outperform the NB model (see Fig. 7). The use of two-stage ZINBR models is helpful when there may be a distinct process influencing the occurrence of a proportion of data points with the value of zero.[2]

To statistically verify our intuition-based model selection we formally tested for over-dispersion with a Likelihood Ratio test and excess zero counts with a Vuong test. The results confirmed that the ZINBR is the best model for our data. To further account for possible uncaptured heteroskedasticity in our models we report significance using Huber-White standard errors.

Models

Formally, our model is defined as:

---

[2] We anticipate that numerous papers will not be cited for structural reasons, including article type, journal, and bibliography characteristics. Other papers in our database received no citations for the time period covered simply due to chance. There are also papers that are cited frequently, leading to over-dispersion in our dataset.

$$\Pr(y=0) = p + (1-p)\left(1 + \frac{\lambda}{\alpha}\right)^{-\alpha}$$

$$\Pr(y>0) = (1-p)\frac{\Gamma(y+\alpha)}{y!\Gamma(\alpha)}\left(1 + \frac{\lambda}{\alpha}\right)^{-\alpha}\left(1 + \frac{\alpha}{\lambda}\right)^{-y}$$

where p, the probability of a structural zero count, and λ are modeled as:

$$\ln\left(\frac{p}{1-p}\right) = c_p + a_1 v_1 + a_2 v_2 + \ldots \quad \text{(the zero - inflation portion of the model)}$$

$$\ln(\lambda) = c_\lambda + b_1 w_1 + b_2 w_2 + \ldots$$

where v and w are our independent variables, a and b are the corresponding regression coefficients, and the c's are our regression constants (intercepts). Here v and w are labeled differently though they coincide in our models. The over-dispersion parameter α is determined by the iterative maximum-likelihood procedure used to fit the model.

Thus, the predicted mean number of citations for a paper, given its characteristics, is $\lambda \cdot (1-p)$ (note this is independent of α). Using our formulations for p and λ from above and taking natural logarithms, this becomes

$$\ln(\lambda \cdot (1-p)) = -\left(c_p + a_1 v_1 + a_2 v_2 + \ldots\right) + \ln(p) + \left(c_\lambda + b_1 w_1 + b_2 w_2 + \ldots\right)$$

$$\lambda \cdot (1-p) = \frac{\exp(c_\lambda + b_1 w_1 + b_2 w_2 + \ldots)}{1 + \exp\left(c_p + a_1 v_1 + a_2 v_2 + \ldots\right)}.$$

This formulation clearly displays the difficulty of interpreting the final regression coefficients. The predicted number of citations depends on the coefficients in both the zero-inflated section and the NB section. However, the coefficients do not function in a simple additive manner. To attempt to illustrate the effect sizes of the coefficients we will utilize marginal effects plots to convey the average influence that our explanatory variables exert in the ZINBR model. The zero-inflated portion of the model should be interpreted as predicting the likelihood of a zero count; thus a negative coefficient in this portion
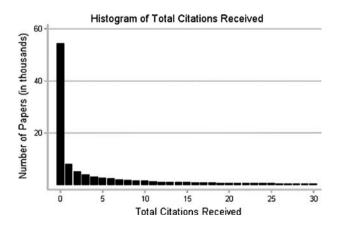

Fig. 7 Total citations. Distribution of papers by number of citations shows the extreme dispersion of the data on the right, implying a negative binomial model, and the large number of zero points on the left suggests that we ought to use a zero-inflated model. We cut off the histogram at 30 to make the trend visible. The papers excluded have total citations ranging up to 4,580

indicates a decreased probability of having zero citations. The NB portion of the model can be interpreted in the usual statistical manner. Thus we would expect strongly significant predictors to exhibit opposite signs in the two portions of the ZINB model.

To capture as much of the variance in paper citations as possible before testing our hypotheses, we create a Base Model that predicts a paper's total citations based on external paper and field characteristics without including any cluster-specific information. For the Base Model we construct a ZINB model and include bibliography size, coauthorship, journal, and year effects as our explanatory variables. The year effect is necessary because papers that are published earlier tend to have accumulated a greater number of citations. We therefore control for publication year with dummy variables. In predicting the inflated zero counts we utilize these same explanatory variables.

In Model I we build upon our Base Model to explore Hypothesis 1, which asks whether a paper benefits from membership in a cluster. To this end we augment our Base Model with the binary cluster membership variable identifying whether or not a paper belongs to a cluster.

Model II investigates Hypothesis 2, which argues that a paper at the semi-periphery of its cluster is more likely to be highly cited. We include in the analysis, in addition to the Base Model, the variable's distance from center and the squared term of distance from center. We use splines to fully characterize any nonlinear trends. Papers that are not in a cluster are excluded from consideration when fitting this model since they have no meaningful measure for distance. Consequently, the binary cluster membership variable used in Model I is not included in this model.

For Model III we adjust Model I to account for the extent to which an author is benefited or harmed by publishing in many schools of thought (represented by clusters) throughout his or her career. We aim to determine whether individual papers receive more citations if the author has a diverse experience with multiple schools or within few schools of thought in our data—a proxy for an author's more general exploratory tendencies. To do this we include in the analysis a diversity measure to capture author publication diversity in addition to the independent variables included in Model I.

Results

Our Base Model reveals that we have constructed a sound basis for modeling the number of citations received by papers. The coefficients for all included independent variables are significant ($p < 0.001$) in both the zero-inflation and NB portions of the model, as shown in Table 4. To test the effects of our combined year and combined journal dummy variables we fit reduced models, eliminating each group of dummy variables in turn, and compare these models to our full models using Likelihood Ratio tests. The results of these tests confirm our expectations that both journal of publication and year of publication are significant predictors for all of our models. The results for these tests are given in Table 4 (Fig. 8).

The addition of an indicator for cluster membership in Model I allow us to test whether belonging to a cluster has a positive or negative effect on a paper's total number of citations. Our Model I (see Table 4) includes significant coefficients for cluster membership in both portions, fully supporting Hypothesis 1. Cluster membership, on average, is associated with receiving 15.03[3] more citations, holding all other variables unchanged.

---

[3] 15.03 is the slope of the line plotted in Fig. 6. It is the difference between predicted citation counts for in-cluster papers and non-cluster papers after removing extreme outliers.

Table 4 Coefficients and significance values for zero inflated negative binomial models

	Base	1	2	3
Negative binomial model coefficients				
Cluster (binary)		0.253***		0.192***
Distance			2.743***	
Distance squared			−2.439***	
Diversity of sources				0.181***
Diversity squared				−0.148***
Year of publication	sig.***	sig.***	sig.***	sig.***
Journal dummies	sig.***	sig.***	sig.***	sig.***
Bibliography size	0.013***	0.013***	0.010***	0.012***
Coauthorship (binary)	0.193***	0.194***	0.116***	0.035***
Cluster sum				0.009***
Constant	2.029***	1.851***	1.789***	1.515***
Alpha	1.449	1.436	1.234	1.388
Log-likelihood	−255,172	−254,775	−169,004	−253,487
Number observations	113,014	113,014	60,840	113,014
Zero inflated model coefficients (likelihood of zero count)				
Cluster (binary)		−0.756***		−0.723***
Distance			−0.541	
Distance squared			−0.415	
Diversity of sources				−0.644***
Diversity squared				−0.059
Year of publication	sig.***	sig.***	sig.***	sig.***
Journal dummies	sig.***	sig.***	sig.***	sig.***
Bibliography size	−0.590***	−0.538***	−0.335***	−0.504***
Coauthorship (binary)	−1.650***	−1.621***	−1.416***	−1.688***
Cluster sum				0.020***
Constant	3.266***	3.314***	2.369***	3.281***
Number zero observations	54,250	54,250	14,033	54,250

Note: Significance reported for year of publication and journal dummies was calculated by a full versus reduced model LR comparison test

sig. significant, *n.s.* not significant (significance implied by robust standard errors)

* $p = .05$; ** $p = .01$; *** $p = .001$

Figure 9 is a visual representation of the average effect for cluster membership, as seen by the positive slope of the linear fit.

As seen in Table 4, the independent variables from our Base Model remain significant, and the directions of these coefficients, along with their interpretations, remain the same. Furthermore, there is a strong (inverse) relationship ($p < 0.0001$) between membership in a cluster and receiving zero citations, as seen in Table 5.

When computing Model II we excluded from consideration all papers that were not assigned to a cluster, since they have no meaningful distance measure and used a normalized measure of distance for papers that were within a cluster such that they range from 0.0 (very central) to 1.0 (extreme periphery). Model II includes our variable representing distance, allowing us to test Hypothesis 2. Within the NB portion of the model, the
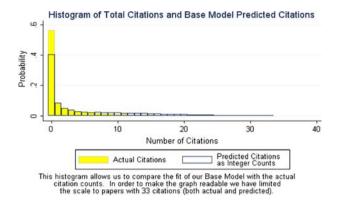
Fig. 8 Total citations versus base model predicted citations. Illustrates how well Model I predicts the actual citations. On average there is good agreement between the densities, indicating that Model I predicts an aggregate distribution of citations very similar to the true distribution of citations
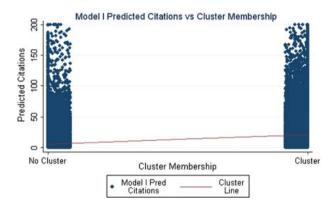

Fig. 9 Model I predicted citations versus cluster membership

Table 5 Cluster membership versus zero citations

Total cites	Cluster membership?		Totals
	No	Yes	
Zero citations	44,618	9,631	54,249
≥1 citations	19,098	39,667	58,765
Totals	63,716	49,298	113,014

significant coefficients for distance squared and for distance indicate a potentially curvi-linear relationship between distance and total citations. Within the zero-inflation portion of the model we found that both distance and distance squared were non-significant. As seen in Fig. 10, increasing distance from the core of a cluster is beneficial until distance reaches 0.54, and beyond this further increase in distance is associated with relatively fewer expected citations. Based on Model II, moving away from the optimal point in the
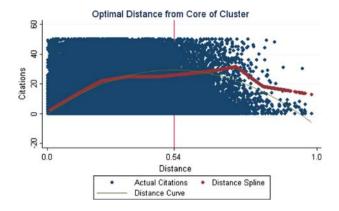
Fig. 10 Optimal distance from core of cluster

semi-periphery by $\pm\sigma$ (SD = 0.1607), the number of expected citations decreases by 2.88,[4] holding other variables constant. The variables held over from our Base Model retain significance in the same direction, leaving their interpretations unchanged.

To further characterize the relationship between distance and total citations, particularly because of the skewed nature of papers toward the core of the cluster, we use linear splines to track the effects of distance over smaller intervals. As shown Fig. 10, the quadratic curve is a reasonable parameterization of the relationship, though the spline has a plateau in the semi-periphery rather than a clear apex.[5] In Fig. 10 one can see that though there is great variation in the distribution of papers, a linear spline fits the model best (it is hard to see the density of points in the figure except through the spline) and it is curvilinear.

Model III allows us to examine diversity of publications as a predictor of total citations. The initial results support our first competing hypothesis, which argues that a diverse publication pattern does indeed lead to higher citations. Results are significant ($p < 0.01$) for both the NB and zero-inflation portions of Model III. The positive coefficient in the NB portion indicates that increased diversity in an author's publication pattern, which is

---

[4] 2.88 is the difference in Model II between predicted citations at distance = 0.54 and ±1 SD (distance $\approx$ .36 or distance $\approx$ 0.68).

[5] Given the mechanism of our clustering algorithm there could be different reasons for a paper to be on the semi-periphery of a cluster. Papers with high impact could be on the semi-periphery because they cite unusual papers, either inside or outside their own cluster, or because they cite a mix of papers both within and outside their cluster. To explore the drivers of semi-peripheral placement for high-impact papers we constructed two additional variables to capture the average distance of each paper's citations within and outside its cluster and added both these variables into our model for Hypothesis 2. We found that the coefficients for both these variables were positive and highly significant, implying that successful papers in the semi-periphery tended to cite a mix of central papers both within cluster and in central papers within other clusters.

We hypothesized that a paper which combined knowledge from its own cluster with a few, perhaps one or two, outside clusters rather than many outside clusters would be the most successful. This would allow it to act as a bridge between a few audiences or research communities. To test this we constructed for each paper a Herfindahl Index of cluster concentration for citations made to a paper within another cluster. We did this by summing, for each paper with more than four outside citations, the percentage of citations to papers in each outside cluster. A higher Herfindahl cluster concentration score would imply that the paper made a high percentage of outside citations to one cluster; a lower Herfindahl score would imply that a paper scattered its outside citations to many papers. We added this variable for concentration into the regression for Model II, and it was positive and highly significant. This confirms the intuition that papers which bring together knowledge from a few schools tend to be well cited.
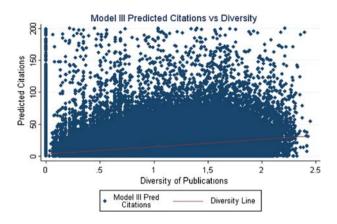
Fig. 11 Model III predicted citations versus diversity

associated with the author's interaction with very diverse knowledge and perspectives, is associated with higher citation counts. Similarly, in the zero-inflation portion, the more diverse an author's citation pattern, the less likely he or she is to receive zero citations, as indicated by the negative coefficient. As shown in Fig. 11, on average, an increase in diversity by one standard deviation (approximately 0.55) is associated with 6.31 additional citations, holding all other variables constant. In the robustness section we examine this finding more closely, finding some conflicting evidence.

Robustness

Here we examine several limitations in the above analysis and attempt to strengthen our findings. First we look more closely at our second hypothesis and explore whether the increased impact of knowledge in the semi-periphery is due to increased citations from within the school of thought or increased attention from outside the school of thought. Since we are predicting that distance has a curvilinear relationship with total citations ranging from the core to the periphery, we split our papers into two groups—*core to semi-periphery* and *semi-periphery to periphery*—based on the distance trend noted in Fig. 12 by dividing the data above and below the optimal distance point of the quadratic.

For each of these two parts of the data we replaced the dependent variable in Model II with the number of citations for that paper originating from outside its cluster and from within its cluster, alternately.[6] We proceeded to test whether changes in distance led to changes in citations originating outside the cluster, which would support our theory that drawing from knowledge outside the cluster tends to lead to more impact on knowledge outside the cluster. This new model confirmed that for the areas of the cluster from the core to the semi-periphery, increasing distance from core attracted significantly more out-of-cluster citations, and from the semi-periphery to the periphery increasing distance from the core attracted fewer out-of-cluster citations.

Utilizing the peaks to either side of the central plateau identified with our splines ($d = 0.3$ and 0.7), we then test whether changes in distance led to changes in the number of

---

[6] For two of these four models there were too few papers with zero citations to justify a zero-inflated model and the standard NB model produced a sufficiently accurate fit for this analysis. Thus for these two models we used a negative binomial regression (see Fig. 9).
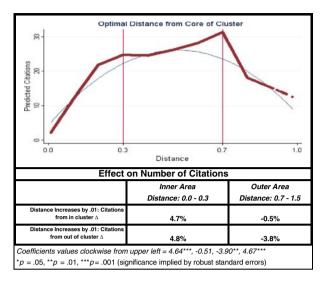
Optimal Distance from Core of Cluster

Effect on Number of Citations		
	Inner Area *Distance: 0.0 - 0.3*	*Outer Area* *Distance: 0.7 - 1.5*
Distance Increases by .01: Citations from in cluster Δ	4.7%	-0.5%
Distance Increases by .01: Citations from out of cluster Δ	4.8%	-3.8%
Coefficients values clockwise from upper left = 4.64***, -0.51, -3.90**, 4.67*** **p = .05, **p = .01, ***p = .001 (significance implied by robust standard errors)*		

Fig. 12 Optimal distance from core of cluster. For regressions for the *upper right* and the *lower right quadrant boxes* above we used a standard negative binomial regression because the zero count was not influential enough to warrant a zero-inflated model. For the *upper left* and *lower left quadrant* regressions there were substantial zero counts, so we used a zero-inflated NB model. The distance coefficient reported is only for the NB portion

citations from the same cluster—as before we predicted a positive change from the core to the apex of the curve and a negative change from the midpoint to the periphery. This would bolster the idea that a paper is combining new knowledge from outside the cluster with knowledge from that cluster to generate valuable contributions within its cluster. We used citations received from within cluster (*Same Cluster Citations*) as our dependent variable and found significant support for our predictions.

Within the core to semi-periphery, an increase in distance of 0.01 is, on average, associated with an increase in citations of 4.7 and 4.8% for in-cluster and out-of-cluster citations, respectively, holding all other variables unchanged. For the semi-periphery to periphery we found that an increase in distance of 0.01 is associated, on average, with a decrease in same-cluster citations by a factor of −0.5% and outside cluster citations of −3.8%, holding all other variables unchanged, as shown in Fig. 12.[7]

Second, we examined more closely our third hypothesis about the diversity of cluster experience. While we found an overall positive impact associated with diversity in our earlier analysis, we suspect that different explanatory mechanisms may be at work depending on how high-impact the authors of the new knowledge are. We differentiate between high-impact authors, for whom diversity of publications may be reflective of high-impact new knowledge generation, and low-impact authors, who may be motivated differently when publishing in many schools of thought without garnering citations. As an exploratory analysis, we therefore calculate an expected citation rate for each paper using the maximum of each paper's authors' average citations within our data, using this to partition our papers by "expected citations rates," and test our hypothesis on these portions separately.

---

[7] In this case the percentage change in citations is equal to $\exp(\beta * \delta)$ where δ is the incremental change in which you are interested and β represents the respective coefficient in the model (Long 1997).

Table 6 Author expected citations and cluster diversity

Model III—expected citations percentile	NB coefficients for diversity	Zero inflation coefficients for diversity
100% (full)	0.164***	−0.719***
<50% (lower half)	−0.057*	0.453***
>50% (upper half)	0.583***	−0.273***
>90% (top 10%)	−0.065 (n.s.)	−0.016 (n.s.)

n.s. not significant

* $p < .05$; *** $p < .001$

We re-ran Model III on those papers with expected citations above and below the median and in the top decile. We found similar results with respect to both direction and significance for diversity of publications among the papers with expected citations in the lower half. For those papers with expected citations in the upper half, however, we found that an author's increasing diversity of publications was associated with a small but significant *decrease* in citations. This result supports our competing Hypothesis 3b. We additionally ran Model III including only the top 10% of papers based on expected citation rate. We were surprised to find that for this highest decile of papers neither Hypothesis 3a nor 3b is supported. See Table 6 for a summary of results.[8,9]

In aggregate, this section give us added confidence in our findings for our first two hypothesis and undermine some confidence in our findings for our third hypothesis. In the section that follows we discuss the implications and consequences of our findings.

Conclusions

By applying performance measures to positioning in and around schools of thought, we reveal that where knowledge is positioned has a significant impact on its performance. Indeed, by studying a complete set of articles in the top 41 journals in the social science field of micro and macro management, we found that new knowledge which is positioned within a school of thought can expect to get, on average, 15.03 more citations. Within a school of thought, knowledge positioned in the semi-periphery of a school of thought

---

[8] We wish to confirm that our clustering algorithm is identifying meaningful schools of thought. While it compares favorably to other algorithms of its kind, we wish to separately test whether the specific tests we claim remain significant if we randomize the cluster assignment for all papers in the same proportions that exist in our dataset and monitor our results for any changes. Our first hypothesis claims that being in a cluster is advantageous in garnering citations; once we randomize citation assignments (and therefore cluster membership) this effect should disappear if the clusters are in fact meaningful. This would add confidence that our results are not an artifact of clustering or statistical methodology. We find that indeed after randomizing we do lose significance for our binary cluster variable as expected. The cluster assignments were successfully randomized ($\chi^2_{(324)} = 276.0230$, $p = 0.975$). The coefficients for our new binary cluster membership variable in our model are now insignificant (NB portion: $p = 0.403$; zero-inflation portion: $p = 0.361$). These results support the validity of our clustering methodology.

[9] Lastly, we want to further rule out the possibility that the excess zero counts in our data obscured the true trends, despite our use of the ZINBR formulation. To accomplish this we excluded all zero-citation papers and proceeded to refit our data with a standard Negative Binomial model. The direction and significance of our coefficients in the NB portion of our previous models remain unchanged. We do not believe that the zero-inflation portion of our model was incorrectly identifying and modeling the excess zero counts or obscuring the trend among those papers that received citations.

(representing knowledge that builds on a mix of knowledge common and unusual in that school) rather than at its center or periphery results in 2.88 additional citations (±1 SD).

We emphasize the complexity of new knowledge positioning because we see the act of new knowledge development to be a deeply socially structured process. Schools of thought represent more than a post hoc artifact of ideas but a dynamic and important force in the future creation and knowledge landscape. This view is in keeping with the findings of researchers who show that firms which both explore and exploit in specific ways over time tend to outperform firms that do not (Rosenkopf and Nerkar 2001; Gittelman 2003; Nerkar 2003).

The logic of knowledge positioning applies to R&D and science more generally. We reran this analysis on an extensive dataset of publications in computer science from 1992 to 2003 and found nearly identical results. Knowledge positioning will be potentially important in any field where a community of researchers exists with relatively free flows of information, interdependent research, and interconnected rewards—including the hard sciences and technology, where both papers and patents could be used as the unit of analysis.[10] Indeed, schools of thought, the theoretical prerequisite for our positioning analysis, have already been identified in virtually all fields in science and technology (McCain 1987; Braam et al. 1991; Aharonson 2004). It remains an open question, however, how sensitive the dynamics of schools of thought in these fields are to their differing knowledge-sharing norms and reward structure. More generally, our findings potentially offer a micro-theory that may be aggregated to offer macro-level insights into the development of research fields in science, social science, and technology in general. Studying these questions is the direction of our continued research.

Our goal in this paper was to better understand the consequences of implicit and explicit positioning of researchers between and among schools of thought. Such positioning problems are by their nature complex and multi-dimensional. We develop a strategic framework for analyzing how new knowledge is positioned within the knowledge landscape, considering seriously the social structure of that landscape—specifically the powerful effects of schools of thought. We find that new knowledge was positioned by its creators under the stress of two search tensions—being a part of an identifiable school of thought and simultaneously reaching beyond that school to draw on outside knowledge.

There are powerful advantages and subtle disadvantages provided by homophilous social groups, or schools of thought, in the process of knowledge creation. While such schools of thought, which represent mini-paradigms or world views, provide an audience and an intellectual structure to build on, we find that authors must also, in moderation, resist their pull and reach out beyond them to introduce fresh ideas and to appeal to outside audiences. This tension between gaining the advantages of joining a school of thought at the same time resisting homogenization from that school of thought is a significant challenge for new knowledge developers while searching for new ideas, and an important part of the dynamic evolution and development of new knowledge.

We speculate that the robust incentive structures we find in clusters are maintained through selection forces within the knowledge environment; incipient clusters that encourage too much exploration lose their integrity and fail to develop strong internal paradigms, while incipient clusters that are too internally focused may not attract sufficient attention or become too stagnant to gain momentum. Clusters that balance these two extremes in the way we describe seem to have survived to populate our dataset.

---

[10] Patents will differ from papers in many respects. The purpose of patents is to establish a proprietary claim on a method, idea or technology while the purpose of a paper is to advance knowledge and share information. Nevertheless, we believe both will exhibit interesting and useful clusters of patterns.

References

Abrahamson, E. (1996). Management fashion. *Academy of Management Review, 21*, 254–285.

Adner, R., & Levinthal, D. (2000). Technology speciation and the path of emerging technologies. In G. Day & P. Shoemaker (Eds.), *Wharton on emerging technologies*. New York, NY: Wiley.

Aharonson, B., Baum, J., & Feldman, M. (2004). Industrial clustering and the returns to inventive activity: Canadian biotechnology firms, 1991–2000. Draft.

Amir, S. (1985). On the degree of interdisciplinarity of research programs—a quantitative assessment. *Scientometrics, 8*, 117–136.

Barnett, W. P., & Sorenson, O. (2002). The Red Queen in organizational creation and development. *Industrial and Corporate Change, 11*, 289–325.

Barney, J. (1986). Types of competition and the theory of strategy: Towards an integrative framework. *Academy of Management Review, 11*, 791–800.

Bayer, A. E., & Folger, J. (1966). Some correlates of a citation measure of productivity in science. *Sociology of Education, 39*, 381–390.

Birnbaum, P. H. (1981a). Academic interdisciplinary research—characteristics of successful projects. *SRA-Journal of the Society of Research Administrators, 13*, 5–16.

Birnbaum, P. H. (1981b). Integration and specialization in academic research. *Academy of Management Journal, 24*, 487–503.

Braam, R. R., Moed, H. F., & Vanraan, A. F. J. (1991). Mapping of science by combined co-citation and word analysis. 1. Structural aspects. *Journal of the American Society for Information Science, 42*, 233–251.

Burt, R. (unpublished). Social capital: Principles and applications.

Castro, P., & Lima, M. L. (2001). Old and new ideas about the environment and science: An exploratory study. *Environment & Behavior, 33*, 400–423.

Choi, S. C., & Coughlan, A. T. (2004). Private label positioning: Vertical vs. horizontal differentiation from the national brand. Unpublished.

Coe, R., & Weinstock, I. (1969). Evaluating journal publications: Perceptions versus reality. *AASCB Bulletin, 1*, 23–37.

Coe, R., & Weinstock, I. (1984). Evaluating the management journals—a 2nd look. *Academy of Management Journal, 27*, 660–666.

Cohen, W. M., et al. (2000). Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). National Bureau of Economic Research Working Paper 7552.

Cole, S., & Cole, J. R. (1967). Scientific output and recognition—study in operation of reward system in science. *American Sociological Review, 32*, 377–390.

Crane, D. (1972). *Invisible colleges diffusion of knowledge in scientific communities*. Chicago, IL: University of Chicago Press.

Crane, D. (1980). An exploratory-study of Kuhnian paradigms in theoretical high-energy physics. *Social Studies of Science, 10*, 23–54.

Culnan, M. J. (1986). The intellectual-development of management-information-systems, 1972–1982—a cocitation analysis. *Management Science, 32*, 156–172.

Culnan, M. J. (1987). Mapping the intellectual structure of MIS, 1980–1985—a cocitation analysis. *MIS Quarterly, 11*, 341–353.

Doreian, P. (1988). Testing structural-equivalence hypotheses in a network of geographical journals. *Journal of the American Society for Information Science, 39*, 79–85.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper.

Ennis, J. G. (1992). The social-organization of sociological knowledge—modeling the intersection of specialties. *American Sociological Review, 57*, 259–265.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science, 47*, 117–132.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy, 30*, 1019–1039.

Gasper, J. T. (2005). Political news. In *Annual meeting of the Public Choice Society*.

Gittelman, M. (2003). Does geography matter for science based firms? Epistemic communities and the geography of research and patenting in biotechnology. Draft.

Hotelling, H. (1929). Stability in competition. *The Economic Journal, 4*, 1–57.

Johnson, J. L., & Podsakoff, P. M. (1994). Journal influence in the field of management—an analysis using Salancik index in a dependency network. *Academy of Management Journal, 37*, 1392–1407.

Kandylas, B. (2005). A mixture model for document clustering by citations. Unpublished.

Kandylas, V., Ungar, L., & Forster, D. (2005). Winner-take-all EM clustering. Unpublished.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal, 45*, 1183–1194.

Krishna, R. V. (2001). Voter clustering and the theory of spacial voting with entry. Unpublished Draft.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.

Levinthal, D. (1991). Organizational adaptation and environmental selection—interrelated processes of change. *Organization Science, 2*, 140–145.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science, 43*, 934–950.

Levinthal, D. A. (1998). The slow pace of rapid technological change: Gradualism and punctuation in technological change. *Industrial & Corporate Change, 7*, 217–247.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*, 71–87.

March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science, 33*, 1404–1418.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.

Marks, U. G., & Albers, S. (2001). Experiments in competitive product positioning: Actual behavior compared to Nash solutions. *Schmalenbach Business Review, 53*, 150–174.

McCain, K. W. (1986). Cross-disciplinary citation patterns in the history of technology. *Proceedings of the American Society for Information Science, 23*, 194–198.

McCain, K. W. (1987). Citation patterns in the history of technology. *Library & Information Science Research, 9*, 41–59.

McGann, A. J. (2002). The advantages of ideological cohesion. *Journal of Theoretical Politics, 14*, 37–70.

Merton, R. K. (1965). *On the shoulders of giants: A Shandean postscript*. New York: The Free Press.

Meyer, M. W. (1994). Measuring performance in economic organizations. In N. J. Smelser & R. Swedberg (Eds.), *The handbook of economic sociology* (vol. viii, 835 pp). Princeton: Princeton University Press/ Russell Sage Foundation.

Meyer, M. W. (1999). Notes from a border discipline: Has the border become the center? *Contemporary Sociology—A Journal of Reviews, 28*, 507–510.

Meyer, M. W., & Zucker, L. G. (1989). *Permanently failing organizations*. Newbury Park, CA: Sage.

Mintzberg, H. (1994). *The rise and fall of strategic planning*. New York: The Free Press.

Moody, J. (2001). Peer influence groups: Identifying dense clusters in large networks. *Social Networks, 23*, 261–283.

Morris, M. W., & Moore, P. C. (2000). The lessons we (don't) learn: Counterfactual thinking and organizational accountability after a close call. *Administrative Science Quarterly, 45*, 737–765.

Nerkar, A. (2003). Old is gold? The value of temporal exploration in the creation of new knowledge. *Management Science, 49*, 211–229.

Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature II. *Libri, 46*, 217–225.

Pantel, P., & Lin, D. (2002). *Document clustering with committees*. Tampere, Finland: SIGIR.

Pfeffer, J. (1993). Barriers to the advance of organizational science—paradigm development as a dependent variable. *Academy of Management Review, 18*, 599–620.

Podsakoff, P. M., Mackensie, S. B., Bachrach, D. G., & Podsakoff, N. P. (2005). The influence of management journals in the 1980s and 1990s. *Strategic Management Journal, 26*, 473–488.

Popescul, A., Flake, G. W., Lawrence, S., Ungar, L. H., & Giles, L. C. (2000). Clustering and identifying temporal trends in document databases. In *IEEE advances in digital libraries, Washington, DC* (pp. 173–182).

Ramos-Rodriguez, A. R., & Ruiz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980–2000. *Strategic Management Journal, 25*, 981–1004.

Rosenkopf, L., & Nerkar, A. (2001). Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal, 22*, 287–306.

Sharplin, A. D., & Mabry, R. H. (1985). The relative importance of journals used in management research—an alternative ranking. *Human Relations, 38*, 139–149.

Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science, 8*, 327–340.

Small, H. (2003). Paradigms, citations, and maps of science: A personal history. *Journal of the American Society for Information Science and Technology, 54*, 394–399.

Small, H. G., & Crane, D. (1979). Specialties and disciplines in science and social-science—examination of their structure using citation indexes. *Scientometrics, 1*, 445–461.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *RAND Journal of Economics, 21*, 172–187.

Williamson, O. E. (1975). *Markets and hierarchies, analysis and antitrust implications, a study in the economics of internal organization*. New York: Free Press.

Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations. *Journal of Law and Economics, 22*, 233–261.