The Meeting of Minds: Forging Social and Intellectual Networks within Universities

Tobias H. Stark, J. Ashwin Rambaran, Daniel A. McFarland

a) Utrecht University; b) University of Michigan; c) Stanford University

Abstract: How are social and intellectual relations structured and given shape within research universities? To answer these questions, we test to what extent various theoretically predicted processes explain the dynamics of academics’ networks of collaboration and shared language use in a unique longitudinal data set (1994 to 2005) of 2,631 faculty at a large private American university. Using the latest advances in stochastic actor-oriented models (in RSiena) and text analysis, we found that social and intellectual relations are clustered and centralized on bridging faculty who form a broader interdisciplinary hub of research in the university, and that, over time, this hub disseminated its style of (interdisciplinary) research to other faculty. These networks are shaped by selection based on age, gender, race, and academic rank as well as the coevolution of social and intellectual relations over time. Clear differences emerge in science, technology, engineering, and mathematics (STEM) fields and are strongly driven by structural mechanisms of clustering and centralization, whereas non-STEM fields (social sciences and humanities) are strongly driven by personal preferences of faculty members.

Keywords: knowledge organizations; academic collaboration; sociology of science; multiplex network analysis; longitudinal social network analysis

A primary goal of research universities is to forge internal social and intellectual communities in which faculty come together to generate novel forms of research. What structures do university intellectual communities have and how do faculty within them structure themselves socially and intellectually? Prior research offers some understanding of how and why a research university might structure intellectual relations between faculty, but this understanding is hampered by fragmented views and very divergent characterizations. Some view faculty as balkanized factions (Blau 1973; Hearn 2007; Kadushin 1976), others as spanning ridge structures (Biancani, McFarland, and Dahlander 2014; Friedkin 1998; Moody 2004), or as centralized hubs (Biancani et al. 2018; Rawlings et al. 2015). Yet other work suggests science, technology, engineering, and mathematics (STEM) and non-STEM (social sciences and humanities) segments are organized in completely different ways (Kagan 2009; Snow 1959). Which of these views hold and which do not? Prior work failed to employ state-of-the-art social network methods to model the development of collaborations as a stochastic process, and few fully take into account selection and preferences, or how social relations may be driven by preferences and rational choices of finding certain collaborators (see for an exception Dahlander and McFarland [2013]). As such, it is unclear what form intellectual networks take within universities and whether these social structures are forged by network mechanisms or selection pressures and preferences of individual faculty.
Even more challenging is the fact that most studies of faculty networks fail to consider intellectual associations. A great deal of the sociology of science concerns the formation of thought styles (Fleck 1979; Knorr Cetina 1999; Kuhn 1970), and this extends into recent quantitative studies of culture (Goldberg et al. 2016; Goldberg and Stein 2018). Some of this work argues that shared cognition can shape the way persons relate to each other (see McLean 2017 for review). For research universities, it is thus far unclear if all our prior studies missed a key component of association: via shared ideas and research interests—or the meeting of minds. It may be that collaborations lead scholars to develop such shared interests as some network research contends (Borgatti and Cross 2003; Rawlings et al. 2015), or even as work on the importance of proximity contends (Festinger, Schachter, and Back 1950; Kabo et al. 2014). Or it may be that, over time, scholars are more guided by their interests as Internet search becomes more prevalent and one can find the best-fitting thought partner (Evans 2008).

But independent of whether shared intellectual interests lead or lag collaboration, there is the question of whether the structure of shared interests and ideas is differently shaped and formed than that of collaboration. Do interests and ideas have distinct structural forms and mechanisms? Are intellectual interests structured around departments or broader fields? Are they centered on and stratified around certain scholars’ ideas and works? And do different mechanisms generate intellectual structures, such as status? We lack answers to these questions because the ideas of scholars are seldom brought to bear and are never brought into relation with social collaborations.

Such a concern has broader implications than just research universities and their structuring. It concerns research and development organizations and their core function in today’s knowledge economy (Powell and Snellman 2004). A far wider class of organizations than universities is intent on developing thriving intellectual communities, and they too wonder how these more elective associations and interests form and what drives them in certain directions. Knowing this may help organizations in the knowledge economy better engineer environments that break from balkanized forms or hone research interest in certain directions (Mäkinen, Evans, and McFarland 2020).

This focus also has broader theoretical relevance as it speaks to the structure–culture and network–meaning discussions in broader sociology (McLean 2017). For more than a century, sociologists have given different accounts of how networks and ideas interrelate: whether relations of production form cultural beliefs (Marx), and social positions shape ideology (Mannheim), or beliefs shape social relations (Weber), or whether they are co-constituted (Simmel). By examining rich longitudinal cases with state-of-the-art social network analysis, we can investigate and address this interrelation.

This article studies the forging of intellectual and social relations within universities by analyzing a single large U.S. elite university. In particular, we focus on the social collaborations and intellectual pursuits of 2,631 faculty for 12 years (54,067 collaboration ties and 131,130 shared idea ties) and employ new social network modeling strategies to better identify the assortment of mechanisms that drive collaboration and intellectual affinity, to ascertain their distinct and evolving
structure, and to develop a richer understanding of how intellectual communities develop. By focusing on a single university, we obtained rich longitudinal information of unprecedented detail that allows us—for the first time—to better specify which mechanisms drive network formation. However, our focus also comes at the cost of ignoring interinstitutional mechanisms and not being able to generalize to all universities, which we reflect on in the conclusion. Nonetheless, given the paucity of rich data on faculty networks within single universities, a focal case seems warranted here as a means of learning how universities forge internal social and intellectual communities.

Predictors of Collaboration and Shared Academic Interests

A great deal of prior research speaks to the structure and change of social and intellectual networks within universities and organizations more generally (Friedkin 1998; Kabo et al. 2014). The research on social networks internal to universities has diminished in recent years as large-scale corpora have emerged and studies have shifted to entire fields or entire countries (Ferligoj et al. 2015; Trujillo and Long 2018; Varga 2019). Given our focus on the development of networks of scientists within universities, not all of our expectations translate directly to the structure of the broader academic community that encompass scholars from different universities. Collaborations and intellectual correspondence at one university are restricted by the local and often rigid opportunity structure of available colleagues and shared ideas, whereas they are much freer across institutions. Moreover, faculty may structure themselves socially and intellectually at their own university in response to local incentive structures such as financial stimulation of interdisciplinary collaboration. These incentives may take very different form for collaborations with colleagues from other universities (e.g., grants for consortia).

Nevertheless, understanding how social and intellectual collaboration is structured among faculty within one university is the focus here, and to speak to it we draw on literature from multiple subfields of somewhat indirect relevance and bring them to bear here in synthesized form: social networks, sociology of science, organization studies, and social theory. Each of them offers salient ideas related to the formation of social and intellectual communities.

Academic Collaboration

Face-to-face interactions in shared local social environments are important for conducting academic research (Kabo et al. 2014; Lee et al. 2010). These interactions have consequences for academics’ collaborative relationships and their research interests. Research has established that the more opportunity people have to interact with others, the more likely they are to affiliate with them (Blau 1977; Festinger et al. 1950; Kabo et al. 2015) and adopt their ideas (Friedkin 1998; Rawlings et al. 2015).

In large research organizations such as universities, scientists are organized in concentric social circles, such as departments, that are embedded in schools.
This setup into foci of activity (Feld 1981) structures academics’ opportunities to interact with colleagues and subsequently start collaborations and learn from each other (Kabo et al. 2015). In fact, departments have been considered constraint foci (Merton 1976) because faculty members are by definition in close proximity, share similar research interests, teach the same students in similar topics, and perform joint administrative tasks. It is thus more likely that collaborative ties develop and persist within schools and departments (Dahlander and McFarland 2013).

Cross-departmental, interdisciplinary collaborative ties are much harder to establish and sustain because ties that bridge groups require frequent negotiations between different interests (Burt 2002). Research has shown that it is difficult to evaluate interdisciplinary work (Klein 2008). Interdisciplinary research centers, therefore, tend to select productive scientists with a proven track record of past successes and facilitate their productivity further (Biancani et al. 2018). These scientists might centralize research efforts through Matthew Effect–type cumulative advantages (Merton 1968) and become academic “stars” that attract many other collaborators (Azoulay, Zivin, and Wang 2010; Rawlings et al. 2015; Zhang et al. 2018). In turn, these scholars benefit from increased productivity and access to new ideas and novel techniques brought in by their collaborators.

Scientific collaboration has also been found to follow ridge structures, or “sequentially overlapping regions of structural cohesion in social space” (Friedkin 1998:159). Research collaborations show topologies with areas of lower density and clusters of higher density. For instance, research studying entire scientific fields, not just collaboration at one university, has found such cohesive clusters of associations between sociologists (Moody 2004) and in other disciplines (Newman 2004). Such connectivity may grow over time as access to digital materials grows and Internet search increasingly becomes the norm (Evans 2008). Forming clusters of collaborations has advantages for the individual researcher. Simmel (1950) already observed that a dyad will be stabilized by a third party that connects both members of the dyad (transitivity). Accordingly, collaborations between two scientists have been found to be more likely and more stable if they involve a third member (Dahlander and McFarland 2013; Kossinets and Watts 2009).

Research on scientific collaboration across universities has also found evidence of assortative mixing with regard to collaborations (Newman 2004). That is, scholars with many collaborators tend to collaborate with similarly successful scholars, perhaps forming “elite clubs” of dissemination, whereas scientists with few collaborators tend to work mostly alone or with others that also have few collaborators. This tendency to associate with other scholars who are equivalently positioned in collaborations leads to stratification in the academic network (Rawlings et al. 2015).

Assortative mixing can also be driven by preferences to collaborate with colleagues who are similar with regard to relevant background characteristics and research focus because people with similar origins and habitus have similar research interests and tastes (Bourdieu 1984, 1988). Bourdieu’s work in Homo Academicus shows that French social scientists and humanists in the 1960s came from lower-income, liberal backgrounds, whereas medical scholars came from wealthier, conservative backgrounds. From this, he inferred that upbringing and habits of
mind led scholars to assume certain intellectual interests, collaborations, and career paths over others.

Scholars associate with other scholars, in part, because they are drawn to those with the same tastes and personal characteristics as themselves (Blau 1977; McPherson, Smith-Lovin, and Cook 2001). People with similar traits and characteristics tend to understand each other better, they often communicate more easily, and find each other more likable and predictable (Byrne 1971). The growing use of public scholar profiles and Internet searches probably amplify the effects of similarity over time as faculty now can select more widely and on multiple dimensions their collaborators and research topics of interest (Evans 2008). Collaboration has thus been found to be partially a consequence of people’s preference to affiliate with those who are similar with regard to gender, age, and status (Kossinets and Watts 2009; Zhang et al. 2018).

Such preferences may be different for members of minority groups on campus. Research found that the preference for same-ethnic friendship is stronger among adolescents who belong to ethnic minority groups, particularly if they have more same-ethnic peers to affiliate with (Smith et al. 2016). Similar processes may extend to academic collaborations. Perhaps some of the balkanization and clustering associated with departments and fields may be due to the self-selection of people with certain background characteristics into particular fields (McPherson et al. 2001).

But what of preferences for colleagues with certain resources and capacities? Collaborations may also be a result of scholars’ strategic decisions based on their own and other scholars’ resources, knowledge, and skills (Dahlander and McFarland 2013; Melin 2000; van Rijnsoever, Hessels, and Vandeberg 2008). Junior faculty might seek out higher-ranked colleagues to benefit from mentoring in terms of obtaining funding, tenure support, or general visibility but also by joining an academic thought community. Senior faculty can benefit from the new skills and updated ideas junior colleagues bring to the table. Research in another professional domain found that advice relationships can turn into collaboration and vice versa (Lazega and Pattison 1999). This exchange of knowledge and ideas across academic ranks is distinct from centralization and may reflect localized exchange relations. Although both benefit, junior faculty might have less capacity for engaging in such cross-rank collaboration because they can build on less experience and less reputation that would make them visible to others (van Rijnsoever et al. 2008).

**Intellectual Affinities**

Seminal work in social network science discussed preferences for shared attitudes or values as a key predictor of the formation of social relationships (McPherson et al. 2001). Recent contributions have shown that such preferences extend to shared language, experiences, and basic cognitive affinity (Goldberg and Stein 2018). But this work has yet to ask how the cognitive structure that underlies the formation of social relationships is shaped and evolves. To derive predictions, we turn to work in the sociology of culture and science that provides insight into how cognitive structures may be shaped in universities and knowledge organizations.
Departments and schools form the building blocks of universities and facilitate the development of separate epistemological (Knorr Cetina 1999) and academic cultures (Kagan 2009; Snow 1959). Durkheim ([1893] 1933) has already suggested that intellectual specialization and fragmentation of disciplines would not create mutual interdependencies and instead be anomic, preventing the knowledge flow across disciplines (Blau 1973). Accordingly, work in the sociology of science remark on the lack of intellectual integration and increased balkanization within universities (Blau 1973; Hearn 2007; Kadushin 1976). These works note how scholars tend to form clusters of academic interests, thought processes, and homogenous research styles. In fact, the ridge structure found in academic collaboration (Friedkin 1998; Moody 2004), has also been found in the topology of shared ideas among scholars from different universities in terms of citations of the same articles (Moody and Light 2006) and may extend to shared ideas within universities. These insights are complemented by research showing that local clustering can enhance knowledge creation (Whittington, Owen-Smith, and Powell 2009).

However, intellectual affinities may not perfectly fall within departments because of internal fractionalization (Abbott 2001) and may reflect broader paradigms that extend beyond departments (Kuhn 1970), such as big data or technoscience. Disciplinary divisions in thought styles may become particularly blurry at times when university administration and funding agencies aim to promote scientific progress through focusing on interdisciplinary collaborations (Jacobs and Frickel 2009). Moreover, with the rise of the Internet, scholars’ searches have become more focused and market-like than relationally driven (Evans 2008). Scientists increasingly search for the right complement or exact shared expertise in new scholarly writings rather than draw on expertise from existing social ties. Evans (2008) finds that academic citation grew far more interdisciplinary, which suggests greater intellectual connectivity. That is, even without collaboration, people are easily alerted to related work, which increases the likelihood of knowledge transmission (Friedkin 2009). These processes may promote the clustering of shared ideas independently of departments or fields.

In addition, the structuring of ideas might be driven by highly productive and visible academics who promote intellectual cohesion (Collins 1994; Crane 1972). These “academic stars” draw attention to their own work and might serve as facilitators of ideas and knowledge (Azoulay et al. 2010; Moody 2004). The Internet has likely increased the level of centralization. Evans (2008) shows that with the rise of the Internet, scholarly citations focused on fewer and newer texts. Research studying millions of scientific abstracts also found that the rare real innovations tend to form ideational bridges between knowledge clusters, subsequently centralizing scholars’ intellectual work (Foster, Rzhetsky, and Evans 2015).

Insofar as the scientific endeavor is considered a “regulated competition” (Bourdieu 2000:126) or a “regulated struggle” (Bourdieu 2004:62), scientists have to make strategic choices to achieve peer recognition (Bourdieu 1975). This line of thought suggests that continuation of research traditions is instilled through scientists’ education and eventually rewarded, whereas deviance is punished by neglect (Foster et al. 2015). As a consequence, knowledge regimes may emerge in which shared ideas and research styles define membership. This may stratify scientific knowledge in a
core–periphery structure of ideas. One can also imagine “schools of thought” such as the Chicago school or Weberian tradition in sociology (Abbott 2009) with high internal cohesion and few bridging ties between them.

Just as preferences may drive collaborations, they may also drive intellectual matters. Bourdieu’s (1988) analysis of French academics in the 1960s showed that people with similar background characteristics (similar socialization) tend to share intellectual interests. Other research found, for instance, that highbrow cultural tastes tend to forge exclusionary networks, whereas popular culture tastes forge weak, bridging ties (Lizardo 2006). As such, persons of similar backgrounds may also share intellectual focus. Moreover, the aforementioned localized exchange relations between junior and senior faculty extend to intellectual interests. Academics collaborate not just to increase productivity but also to achieve new skills and gain new insights (Azoulay et al. 2010; Borgatti and Cross 2003). Senior scholars have greater experience with the research community and the dominant research paradigm, whereas junior scholars might bring in fresh ideas and new methods.

Lastly, the way in which the organization of academic work is driven by these mechanisms might differ between academic cultures (Fleck 1979; Kagan 2009). Early work by Snow (1959) argued that scholarly activity is divided into two cultures (oversimplified for brevity): on the one hand, there is a positivist culture that continually advances ideas (STEM fields), whereas on the other, there is a culture that is humanistic and continually elaborates on itself (non-STEM fields).1

In many regards, this notion of two distinct scholarly cultures echoes the work of Fleck (1979) on thought styles and thought communities and Kuhn (1970) on scientific paradigms. These scholars made special reference to the differences in language, methods, clusters of association, and norms of research activity, all forming a gestalt, or social and cognitive configurations. Kuhn (1970) noted that non-STEM fields were preparadigmatic and lacked a cogent thought style. Today, STEM fields have become more engaged in rapid discovery team science with large teams and fast-paced work, whereas the non-STEM fields are slower and more monastic. Such variation by field is clearly important, and we attempt to explore differences in the various mechanisms we propose.

The Interrelation of Collaboration and Shared Interests

Once we consider the salience of shared ideas and interests in social relationships, we then have to confront the question of how they relate to each other. Classical sociological theories hold that shared ideas and social structure might coevolve, but these theories disagree on the causal order in which this happens (Mannheim 1936; Weber 2002) or whether causality can be assumed at all (Bourdieu 1984, 1988; Simmel 1971).

Recent work by relational sociologists has tried to address this chicken and egg problem (McLean 2017). An emerging theme is that in some organizations, collaboration will not facilitate shared interests and intellectual focus but rather the opposite must hold first. One has to speak the same language to have successful collaborations (Goldberg et al. 2016; Srivastava et al. 2018). Shared cognition or language then is a key component to the formation of academic collaboration. This
echoes Max Weber’s (2002) argument that belief systems can shape the mode and means of production. Accordingly, many scholars have argued that shared interests and ideas shape academic social structure (Knorr Cetina 1999; Kuhn 1970), and findings from various domains show that shared interests can indeed guide the formation of social ties (Lewis and Kaufman 2018; Stark and Flache 2012; Vaisey and Lizardo 2010).

Other classical sociological work has long argued for the reverse causal relation. Karl Marx (2000) contended that the relations of production drive ideology, and Karl Mannheim (1936) argued that social structural positioning influences how a person perceives and interprets the world. Based on these theorists, one would expect that shared ideas result from the persistent patterns of interaction within intellectual communities or, in short, that structure shapes scholarly culture (Merton 1973). Such social influence processes leading to shared interests within social relationships have been found in various domains (Centola and Macy 2007; Festinger et al. 1950; Rambaran, Dijkstra, and Stark 2013). Collaboration has been shown to create knowledge flows among academics (Rawlings et al. 2015), and it may, in a similar way, drive scholars’ research interests.

A third body of work argues that shared ideas and social structure are co-constitutive without favoring one causal direction over another. From a microangle, Simmel (1971) foresaw form and content, or relations and sentiments, as a dualism in which content was always the flip side of form, and form was always the flip side of content. Similarly, but from a broader societal perspective, Bourdieu (1984, 1988) argued for a homology of position and tastes. Many network scholars reference this idea when they consider relationships like friendship, which they characterize as both a particular content of positive affect and a particular form of a reciprocal tie (Martin 2011). Recent social network research has come to adopt this correspondence view (see McLean [2017] and Mische [2011] for a review). Academic collaboration and shared ideas at a university might thus form a homology, and it is even feasible that the causal arrow reverses from time to time as discussions of thought communities and thought styles interrelate differently in different epochs. For instance, Fleck (1979) argued that the social (community) and cognitive (style) could decouple temporarily in academic cultures when debates arose.

All of this points to the need for a rich longitudinal case that explores the interrelation of social and intellectual relations over time. From this, we can begin to untangle these possible formulations and their processes of creation. Previous work has focused only on academic relationships (e.g., co-author networks; see Moody [2004] and Newman [2004]), only on intellectual ties (e.g., co-citation networks; see Trujillo and Long [2018] and Varga [2019]), or only on the flow in one direction (e.g., from collaboration ties to citation flow; see Rawlings et al. [2015]). The bidirectional flow within one knowledge organization has thus far not been explored. This article does just that and uses the rich case of a single large elite university over the course of 12 years as its empirical focus.
The University under Study

The university we focus on is an elite American research university and its associations and activities spanning 12 years (1994 to 2005). This university and time span are well suited as an organizational case of a research and development firm aimed at forging social and intellectual associations. The university is first and foremost a research institution whose primary mission is to forge and discover new ideas. Moreover, it has a long history of working toward guiding and facilitating such endeavors. During the period of data collection, this university exhibited a particular organizational culture and set of large-scale resource allocations to improve the research community in certain ways.

Prior to the period observed, the university had been trying to forge greater interdisciplinary collaboration and better integrate the university community. In general, the administration (and faculty senate) sought to overcome balkanization and replace what was perceived as tired and parochial interaction within intellectual enclaves with exciting and boundary-spanning interactions across campus. The university even created an "eighth school" to facilitate this—the dean of research office. This office housed multiple interdisciplinary centers and prevented territorial disputes over ownership that occurred previously between schools and departments every time a new interdisciplinary project, grant, or center formed.

By the time our data collection began, the university had become a premier university and had facilitated a variety of interdisciplinary projects and efforts, all at a time when interdisciplinarity had yet to emerge as a national priority in research funding agencies. Within the university, a cadre of elite interdisciplinary scholars had succeeded in obtaining Nobel prizes and many grants and in energizing their related departments. This hub of activity was centered on biology/physics/chemistry and engineering parts of campus, and especially microlevel aspects of biology (bioengineering, biophysics, biochemistry). With the entry of a new administration, the university began a massive funding drive (in the billions) to create interdisciplinary institutes that could replicate this innovative form of interdisciplinary team, technoscience in other segments of campus. A variety of supercenters emerged and catalyzed these efforts, spreading the model of science to other younger faculty and even new departments on campus (for summaries, see Biancani et al. [2018, 2014] and Mäkinen et al. [2020]).

These efforts and narrative partly describe the social structural and intellectual form the faculty community took and evolved toward over the period observed. As such, the university was characterized as having clustered endeavors and interests but that it moved toward more interdisciplinary bridging. Moreover, although a hub of interdisciplinary stars formed a core catalyst, this shifted over time as bridging efforts spread in the university via interdisciplinary initiatives. Unclear in this account, however, is first whether this actual structure was observed in both collaborations and shared research interests; and second, whether and what mechanisms were at work—such as selection, homophily, and proximity, or social and intellectual closure, centralization, and assortativity. Through an empirical case study, we explore and uncover this form and its process of creation.
Materials and Methods

Data
To test our hypotheses, we used complete network data and individual background information of all faculty members at a major private U.S. research university between 1994 and 2005. Data were systematically gathered from university central offices on numerous faculty affiliations and background characteristics as well as their collaborations. These administratively obtained data were linked with data on all faculty publications obtained from the most widely used academic database (Clarivate’s Web of Science). The initial sample consisted of 1,396 scientists who were faculty members in 1994. New members could join the university from 1995 onward, whereas existing members could leave at any time (see Table A1 of the online supplement for the transition changes between two consecutive years).

Measures

Dependent variables. The coevolution of scientific collaborations and shared ideas was modeled with three different types of social networks.

Collaboration networks. Coauthoring grants and articles as well as working on awarded grants indicate the strongest possible forms of collaboration among faculty members in the present study (Dahlander and McFarland 2013). In determining collaboration ties, we took into account that a joined research project can last several years and there is usually an additional delay until a journal article is published or a grant is submitted (Biancani et al. 2014). We considered a 3-year period to be reasonable for coauthored publications, whereas a 2-year period deemed fit for coauthored grants, and the immediate year for awarded grants because they are ongoing. For instance, a collaboration tie for the year 2000 was indicated when two faculty members had at least one joint publication in the years 2000, 2001, 2002, or when they coauthored (a) grant(s) in 2000 or 2001, or when they had a grant that was running in the year 2000. A collaboration tie was considered decoupled when no new collaboration between the two faculty members took place in the 3 years following the year of joint publication(s) or in the 2 years following the grant-awarded (running) year (Dahlander and McFarland 2013).

Shared language networks. To determine shared ideas of academics, we examined the similarity of faculty members’ language usage in their journal articles by comparing the text of their published abstracts. The collected words in their article abstracts were weighted according to the “term frequency–inverse document frequency” or tf–idf algorithm, the most widely used technique for information retrieval in texts (see Manning, Raghavan, and Schütze 2008). In each abstract, each word is given a weight that depends on its frequency in that specific abstract (tf), and on the inverse of its frequency in the corpus as a whole (idf). Thus, a word that is rare overall, appearing in few abstracts, but frequent in one specific abstract, will be given a higher weight for that specific abstract. A word that is mentioned in many abstracts has a lower weight in all abstracts.

Each author was represented by a weighted vector for all words in their total text vocabulary. Pairs of authors were compared by measuring the distance in
their respective vectors using the cosine similarity \([1 – \text{cosine(vector a, vector b)}]\), which varies from 0, signifying no overlap in actor-pairs language similarity, to 1, indicating identical language. Cosine similarity has the advantage of normalizing for the length of vectors, to allow more appropriate comparisons between author-pairs who have published differing numbers of articles in our 3-year time window for journal publications.

For pairs of authors (person \(i\) and person \(j\) who copublished, each person’s tf–idf vector was recalculated to exclude the abstracts of articles coauthored among that pair. In this way, articles coauthored by person \(i\) and person \(j\) would not contribute to \(i\) and \(j\)’s language similarity measure but would contribute to the similarity observed between \(i\) and a third person \(k\), and that between \(j\) and \(k\).

Language similarity scores were heavily skewed with mean values around 0.01. Normalization using log-log transformation (after adding one to the scores) was necessary to obtain a meaningful cutoff value for shared language ties. Two faculty members’ language use was considered shared when their transformed score for language similarity was greater than two standard deviations above the mean score across a 3-year time window (the period in which language similarity in journal publications was assessed). Shared language use determined by this cutoff value was represented by a one in the binary shared language network. It was not possible to weight the strength of the similarity because of the limitations in the statistical model (see below).

**Co-citation networks.** Because there is no obvious right choice for determining a cutoff value indicating shared language use, we validated our results by comparing them with the most typical measure of thought communities, namely co-citation: the extent to which faculty members cited the same articles in their articles. This was measured as the overlap in number of references faculty member \(i\) and \(j\) had in articles that were not jointly coauthored by \(i\) and \(j\). In maintaining consistency, the same 3-year time window of published journal articles was used to determine this form of shared ideas between faculty members. We considered author \(i\) and author \(j\) as having similar ideas when they shared more co-citations than the mean number of co-citations among all faculty in the entire observation range (1994 to 2005), which was \(M = 2.75\). A co-citation tie existed between two faculty members in the binary co-citation networks when they cited three or more of the same articles within a 3-year time window.

**Explanatory variables.** Individual and dyadic variables were constructed to examine their unique effects on the formation of ties in networks of collaboration, sharing language, and co-citation.

**Individual-level variables.** We included faculty members’ gender (0 = female, 1 = male), race (0 = non-Hispanic white, 1 = minority), and age (in years). Two dummy variables tested for differences between tenured faculty members (as reference) and those who were untenured or clinical faculty (clinical faculty are only in the STEM fields). We distinguished STEM field faculty comprising the natural sciences, physical sciences, medicine, and engineering (initial \(n = 888\) in 1994) from those in the humanities and social sciences, which included law, business, and education (initial \(n = 508\) in 1994), with a dummy variable (0 = non-STEM field, 1 = STEM field).
Dyadic-level variables. Four binary dyadic variables for pairs of faculty were constructed: same gender (cross-gender as reference), same race (dichotomous as minority group versus non-Hispanic white; cross-race as reference), similar age (with, at most, a 2-years age difference), and same rank or tenure (cross-rank as reference). A dyadic network covariate measure of same affiliations, indicating contact opportunities, was included as well (department and school, where 0 = different field and different department, 1 = same field and different department, 2 = same field and same department).

Analytic Strategy

Social and intellectual associations between faculty can be seen as a function of strategic decisions of faculty (actors), which can be adequately represented with the stochastic actor-oriented model (SAOM) using simulation investigation for empirical network analysis (SIENA; Snijders 2001; Snijders, Van de Bunt, and Steglich 2010; Steglich, Snijders, and Pearson 2010). This model investigates whether change in association depends on theoretically specified assumptions about the underlying preferences of the actors. The SAOM simulates network development based on the idea that tie formation, tie maintenance, and tie dissolution follow from each person’s preference to optimize their position in the network (Ripley et al. 2020). The exponential random graph model, the other prominent class of social network model, instead predicts ties in the network based on the state of the entire graph and not based on preferences of focal actors (Block, Stadtfeld, and Snijders 2019). Given our argument that faculty consider and decide upon their relations, the SAOM seems the most suited methodology.

Building on recent advances in the coevolution of multiplex network analysis in the stochastic actor-oriented modeling framework (Snijders, Lomi, and Torlo 2013), we examined how networks of collaboration and shared language (or co-citation in the robustness check) codevelop over time. To assess effect sizes, we calculated the relative importance of each effect on the probability of a tie change (Indlekofer and Brandes 2013). This allowed us to determine the extent to which collaboration and shared ideas (mutually) form one another, or are generated by structural processes, strategic decisions, opportunity, and preferences.

Networks were modeled using a time window of four consecutive academic years, which yielded three nonoverlapping observation periods (1994 to 1997, 1998 to 2001, and 2002 to 2005). The last observation was 2005 because the networks were based on collaborations and shared language use in articles published in the current and following 2 years and our last data source was 2007. Narrowing the observation window to 4 years kept the dynamic network analyses via the SAOM empirically tractable, as our models ceased to converge well with larger and more network panels because of their exponentially increasing complexity. As an additional advantage, modeling different periods separately allowed for the comparison of changes in the dynamics over time. Moreover, this approach partially prevented the assumption of stable network dynamics across the entire time span of 12 years (Lospinoso et al. 2011).
*Model specification.* The same set of effects was included in both network types (collaboration and shared language, or co-citation in the robustness check) to draw conclusions about community building and idea convergence.

*Structural processes.* We included degree, clustering, centralization, and stratification as four important selection effects to examine endogenous structural processes in networks of collaborations and of shared language. Degree reflects the general tendency of faculty to start forming new collaborations or to start using the same language in their journal articles over time, comparable to the intercept or grand mean in linear models.

Transitive triads assessed whether the collaborator of a collaborator was likely to subsequently become a collaborator in the future. This reflected clustering and network closure because all collaborations are reciprocal ties. The same measure in the shared language network tested whether two authors with similar research language to a third person adopted similar language to one another in the future.

Degree of alter assessed centralization, which in the collaboration network tested whether faculty members with many collaborations were more likely to increase in collaborations over time, whereas the same measure in the shared language network tested whether faculty members that shared the research language of many other colleagues had even greater language similarity with others over time. In a sense, this is akin to the notion of the Matthew Effect, where the wealthy become wealthier, or the popular attract further attention (Merton 1968). The end result is increased centralization.

Last, degree assortativity assessed the tendency toward network stratification and indicated whether persons with many collaboration (or shared language) ties were more likely to start new collaborations with persons with similar levels of association; or whether persons with few collaborations were more likely to form future collaborations with each other. As such, the term captured the extent to which individuals stratified associations into caste structures, where the have’s and have not’s occupied distinct social and intellectual spaces.

*Opportunity structure.* The effects of sharing the same academic environment were assessed with a dyadic effect for same affiliations that represented overlapping organizational foci. A positive estimate would indicate that faculty members were more likely to form collaborations or share language with people who are in their own school and department than with people who are located elsewhere.

*Preference.* We examined individual preferential differences based on faculty members’ gender, race, and age. A dyadic similarity effect was included for each of these three individual characteristics to represent faculty’s preferences to form and maintain collaborations and to share language with other faculty members who were of the same demographic category (same gender or same race) or who were of similar age as themselves. When assessing the dyadic similarity effects, we controlled for potential degree differences between persons based on their own individual dispositions (male, minority group, and age). A positive estimate for these effects would indicate that males, racial minorities, and older faculty were more likely to collaborate and to share language with others compared with females, racial majorities (constituting non-Hispanic whites), and younger faculty
(all mean-centered). These additional individual effects were needed to determine the strength and direction of the similarity effects.

Effects of academic rank were tested by including two dummy variables that indicated whether untenured (junior) or clinical faculty were more or less likely to collaborate and share language with others compared with tenured (senior) faculty. In addition, a dyadic effect of same rank measured whether faculty were more likely to form collaborations or intellectual associations with others of similar occupational status than with others of dissimilar occupational status.

*Academic cultures.* Differences in the development of collaborations and shared ideas between STEM fields and non-STEM fields were assessed with a covariate for STEM (non-STEM as reference), representing faculty members’ broader scientific field affiliation. We compared effect sizes between STEM and non-STEM fields using the sienaRI method (for an explanation, see Indlekofer and Brandes [2013]; for an empirical example, see Rambaran et al. [2020]).

*The link between social structure and ideas.* To assess whether people shared language first and then collaborated, or vice versa, or whether there was no clear causal direction between the two, we included two dyadic between-network effects (one in each direction). These multiplex network effects examined whether existing ties in one network type (existing collaboration tie [X]) affected the development and continuation of ties in another network type (new language tie [Y]) and vice versa (existing language tie [X] predicts new collaboration tie [Y]). A positive effect of the collaboration network on the shared language network indicated that collaborating faculty members were more likely to adopt or maintain similar research language over time. The opposite effect of the shared language network on the collaboration network indicated that faculty members who used similar research language were more likely to (continue to) collaborate with each other in the future.

**Results**

*Descriptive Findings*

Descriptive statistics of the individual characteristics and network information on faculty members from 1994 to 2005 can be found in Table A1 of the online supplement. Several trends in the data show a gradual and consistent change in the demographic makeup of the research university. Over a period of more than 10 years, faculty members were increasingly female (though 77 percent of faculty members were male in 2005), racially diverse, and of younger age. The percentage of clinical faculty members increased from 16 percent to 29 percent during this time. This indicates the research university’s effort to expand their hospital and medical school.

The average number of collaborations per faculty member (referring to average degree) varied between 4.22 and 5.88 across all observation periods. There were considerably more collaboration ties in the more positivist research cultures of STEM fields—wherein the average degree varied between 5.72 and 8.32—than in the more monastic research cultures of the humanities and social sciences, for which the average degree of collaboration ties varied between 0.43 and 0.94.
The average degree of shared language ties increased from an average of 7.36 in 1994 to an average of 16.45 in 2005. This indicates that faculty members shared increasingly more ideas in scholarly articles over time. The average degree in STEM fields varied between 11.00 and 23.85 across the observational periods, suggesting large thought communities on campus. The average degree of shared language ties was considerably lower in the humanities and social sciences and varied between 0.70 and 1.64, indicating that researchers in the humanities and social sciences were similar to much fewer colleagues in their language usage.

Collaborations and shared language ties were mostly among scholars of the same gender (66 percent to 77 percent) and of the same race (69 percent to 82 percent), but this decreased over a period of 12 years, whereas the percentages of collaborations and of shared language ties among faculty in the same department or field remained stable and high during this time (84 percent at the minimum). Both collaboration and shared language networks were characterized by high volumes of tie changes, and half of the ties were stable from one time point to the next (Jaccard indexes were about 50 percent), yielding enough tie stability to employ longitudinal social network analysis using the SAOM. The networks were further characterized by moderate levels of clustering in terms of transitivity (transitivity indices were about 20 percent).

Figure 1 shows a visualization of the collaboration core network (Figure 1A) and the shared language core network (Figure 1C) in the second period (1998 to 2001; visualizations of the other two periods can be found in Figures A1 and A2 of the online supplement). These core networks include the 1,214 faculty members who had at least one collaboration tie that was connected to the largest network structure, and they include the 1,385 faculty with at least one shared language tie associate with the largest structure. The colors of the nodes (representing the scholars) indicate school affiliations in both networks. As can be clearly seen, scholars from the medical school dominated both collaborations and shared language. The simplified versions of the core networks (Figure 1B and 1D) show that medical scholars had more ties within their field (indicated by the size of the node) and most ties with other disciplines (indicated by the thicker lines to other fields). Thus, medical scholars not only collaborated widely but they also were more “fluent” in the language used in other disciplines. Faculty from engineering had the second most collaborations within their discipline and relatively more collaborations and shared language ties with the natural sciences and physics than with the other fields (apart from medicine).

The detailed versions of the core networks (Figure 1A and 1C) show that the sections of the network formed by faculty from medicine and engineering were much denser and more complex than those of the other fields. There were many central actors in these two disciplines that seem to dominate the collaborations and shared language usages at this university. The opposite emerged to be true for the social sciences and humanities. Faculty from these fields were quite dispersed with only a few central nodes in the social sciences. Just as for the other fields, the bulk of the interdisciplinary collaboration and shared language ties of social scientists were with faculty from the medical school. In sum, these descriptive findings point
Figure 1: Illustration of collaboration and shared language networks in the period 1998 to 2001. Notes: (A) The network consisting of collaboration ties \( n = 8,280 \) in the period between 1998 and 2001 among all faculty \( n = 1,214 \) who belong to the core network (referring to faculty with at least one tie to the large structure; \( n = 709 \) excluded). (B) A simplified version of complete network in panel (A). (C) A network consisting of shared language ties \( n = 20,848 \) in the period between 1998 and 2001 among all faculty \( n = 1,385 \) who belong to the core network (referring to faculty with at least one tie to the large structure; \( n = 538 \) excluded). (D) A simplified version of the complete network shown in panel (C). Larger nodes in panels (B) and (D) have more within-discipline ties, and thicker lines have more between-discipline ties.
to fundamental differences in collaboration and shared language networks between STEM fields and non-STEM fields.

**SIENA Findings**

Table 1 presents the findings of the dynamic multiplex network analysis predicting the coevolution of collaboration and shared language among faculty members for the three observation periods under investigation (1994 to 1997, 1998 to 2001, and 2002 to 2005) (SAOMs performed with RSiena version 1.2-17). Graphical representations of the hypothetical change process in network configurations underlying each effect can be found in Table A2 of the online supplement. The first three columns with effect estimates (and standard errors) present the results of SIENA analyses regarding changes in the collaboration networks and the last three columns present the results of SIENA analyses regarding the changes in the shared language networks.

The rate parameters in the first three rows of Table 1 (effects 1 to 3) indicate the average number of changes (so-called change opportunities in the SAOM) in the collaboration networks and shared language networks across each estimation period consisting of four consecutive years each. The rate of change fluctuated somewhat between the three estimation periods. The larger estimates of the rate parameters for the language ties indicate a higher average number of changes in the shared language networks than in the collaboration networks (see Table A3 of the online supplement for tests of whether the effects were significantly different in the collaboration and shared language networks), suggesting that language ties had a relatively higher turnover rate than did collaboration ties.

**Structural effects.** In accordance with the low density of the collaboration and shared language networks, the negative effects of degree (effect 4 in Table 1) indicates, considering the large network size, a low probability of faculty members collaborating with other faculty members or having similar language as colleagues. Net of degree, the positive transitive triads effects (effect 5) show that faculty were likely to form clusters and to forge third-party ties in collaboration and shared language networks. That is, they engaged in social and intellectual closure.

Degree of alter effects (effect 6) tell a different story. For collaboration, degree of alter became more positive from period 1 to period 2 (and remained positive in period 3), which suggests that as time proceeded, faculty members preferred to (continue to) collaborate with well-connected colleagues. By contrast, well-connected faculty members persisted in not collaborating with each other as shown by a negative degree assortativity effect in period 2 and small and nonsignificant coefficients in the other two periods (effect 7). Thus, collaboration networks were clustered and became increasingly centralized on faculty bridges, but these bridges were not particularly likely to collaborate with each other. Instead faculty with many collaborations tended to collaborate with those with fewer collaborations.

Shared language networks entail different structural properties. The positive effects for degree of alter in language networks indicates that faculty who shared language with many colleagues attracted even more ties over time. This suggests that faculty members preferred connecting with alters who were already intellectual
Table 1: Results from dynamic multiplex network models predicting coevolution of collaboration and shared language among faculty members between 1994 and 2005.

<table>
<thead>
<tr>
<th>Rate Parameter</th>
<th>Panel 1: Collaboration Network</th>
<th>Panel 2: Shared Language Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Network rate t1→t2</td>
<td>5.78† (0.22)</td>
<td>5.27† (0.17)</td>
</tr>
<tr>
<td>2. Network rate t2→t3</td>
<td>9.94† (0.36)</td>
<td>4.45† (0.13)</td>
</tr>
<tr>
<td>3. Network rate t3→t4</td>
<td>4.63† (0.18)</td>
<td>5.00† (0.18)</td>
</tr>
</tbody>
</table>

Structure Effects
4. Degree (density) | -4.55† (0.07) | -4.24† (0.04) | -4.47† (0.04) | -3.97† (0.04) | -3.71† (0.02) | -3.55† (0.02) |
5. Transitive triads (clustering) | 0.34† (0.01) | 0.44† (0.01) | 0.24† (0.01) | 0.13† (0.002) | 0.08† (0.001) | 0.09† (0.001) |
6. Degree of alter (centralization)† | -0.003 (0.002) | 0.003 (0.001) | 0.004† (0.001) | 0.001 (0.001) | 0.002† (0.000) | 0.002† (0.000) |
7. Degree assortativity (stratification) | 0.001 (0.001) | -0.002† (0.003) | -0.005 (0.002) | 0.01† (0.001) | 0.01† (0.001) | -0.003† (0.001) |

Individual Effects
8. Age (older) | -0.02† (0.001) | -0.02† (0.001) | -0.02† (0.001) | -0.02† (0.001) | -0.02† (0.001) | -0.01† (0.001) |
9. Male (ref = female) | 0.07 (0.04) | -0.004 (0.02) | -0.02 (0.02) | -0.07† (0.02) | -0.01 (0.02) | -0.10† (0.01) |
10. Minority group (ref = white) | -0.06 (0.04) | 0.02 (0.03) | 0.02 (0.03) | -0.04 (0.02) | 0.05 (0.02) | 0.04† (0.01) |
11. STEM (ref = non-STEM) | 0.10† (0.05) | 1.00† (0.04) | 0.82† (0.04) | 0.98† (0.03) | 0.71† (0.02) | 0.77† (0.01) |
12. Untenured (ref = tenured) | -0.42† (0.03) | -0.26† (0.03) | -0.02 (0.03) | -0.34† (0.02) | -0.25† (0.02) | -0.08† (0.01) |
13. Clinical faculty (ref = tenured) | -0.26† (0.02) | -0.16† (0.02) | 0.02 (0.02) | -0.39† (0.02) | -0.51† (0.02) | -0.24† (0.01) |

Dyadic Effects
14. Same affiliations (opportunity) | 0.84† (0.02) | 0.58† (0.02) | 0.71† (0.02) | 0.28† (0.01) | 0.39† (0.01) | 0.35† (0.01) |
15. Similar age | 0.66† (0.08) | 0.49† (0.08) | 0.56† (0.08) | 0.56† (0.05) | 0.37† (0.04) | 0.25† (0.04) |
16. Same gender | 0.11† (0.04) | 0.08† (0.03) | 0.10† (0.03) | 0.07† (0.03) | 0.09† (0.02) | 0.11† (0.02) |
17. Same race | 0.04 (0.04) | -0.06 (0.04) | 0.06 (0.05) | 0.04 (0.03) | -0.004 (0.03) | 0.03 (0.02) |
18. Same rank | -0.34† (0.03) | -0.05† (0.03) | -0.20† (0.03) | -0.26† (0.03) | -0.12† (0.02) | 0.04† (0.02) |

Cross-Network Effects
19. Existing tie X → new tie Y | 1.90† (0.04) | 1.01† (0.04) | 1.23† (0.04) | 1.43† (0.05) | 1.49† (0.05) | 1.26† (0.04) |

Notes: Significance tests performed by dividing the estimates by its standard error resulting in t values, which under the null hypothesis are approximately normally distributed (Ripley et al. 2020). Convergence statistics adhered to the usual criteria for convergence: all parameters had t ratios < 0.10; overall maximum convergence ratio < 0.25. ‡Degree of alter was assessed with the degPlus effect in RSiena, which operates as the degree popularity combined with the degree activity effect under the assumption that the parameters for both are the same. †Individual effects were assessed with the egoPlusAltX effect in RSiena because for undirected networks the model cannot distinguish senders from receivers (see RSiena manual for details). Standard errors are in parentheses under estimates. † p < 0.01; * p < 0.05 (two-tailed test).
bridges in the university. Moreover, the positive degree assortativity effects in the first two periods indicate that faculty who shared language with many others (intellectual bridges) tended to share one another’s language as well. Thus, the intellectual bridges were central intellectual hubs in the university. However, this tendency changes over time. A negative degree assortativity effect in the last period shows that these intellectual bridges were less likely to share language with each other in the last period. This suggests that intellectual bridges became less of a hub over time and spread out into the university.

Opportunity effects. The positive effects of same affiliations (effect 14) indicate that the more scholars shared organizational foci (same field, same department), the more likely they were to collaborate and to share language over time. Thus, net of network structure, networks of collaboration, and shared language were partly driven by shared affiliation in departments, schools, and field. This was probably due to a combination of greater contact opportunities within organizational foci and a higher likelihood of overlapping research interests within the same research discipline.

Preferences. There was evidence for age and gender homophily in all periods (effects 15 and 16 in Table 1): Collaborations and shared language use were more likely between faculty members with the same gender or of similar age. The odds of collaborations and shared language use were particularly high among younger faculty than among older faculty (see Figures A4 and A5 of the online supplement for details). The odds of female–male collaborations and shared language ties were especially low, and interestingly, gender homophily in the language networks was mostly driven by female scholars connecting to other female scholars linguistically. This shows that these two demographic factors foster clustering in networks of collaboration and of shared language.

There was no clear indication for identical processes to be found based on faculty’s race. Two faculty members had no higher tendency to collaborate or to share language if they both belonged to the same race category (either non-Hispanic white or both minority) in the first two periods. After 2002, there was some evidence for race homophily. Collaborations were more likely when faculty belonged both to a minority group. Shared language ties were much more likely among scholars from racial minority groups than among non-Hispanic white scholars or faculty from the racial majority and minority groups.

Academic rank. All estimates of being untenured or clinical faculty (compared with being tenured; effects 12 and 13 in Table 1) were negative and significant (except for collaborations in period 3). This indicates that, net of age effects, tenured faculty members were more likely to forge new collaborations and maintain ongoing ones, and they were more likely to share language with colleagues than untenured and clinical faculty members. As such, tenured faculty appear to play a leadership role in forging centralized networks and a wider intellectual hub. We note, however, that across the board there was a downward trend (see Table A4 of the online supplement for the comparison of estimates), indicating that untenured and clinical faculty were catching up with their tenured colleagues to having nearly the same volumes of collaborations and shared language ties in the last period of investigation (2002 onward).
The effects of same rank (effect 18) were negative and significant for the collaboration networks, which suggests that faculty with the same academic ranks were less likely to collaborate with each other. Same rank was also associated with a lower likelihood for a shared language tie in the first two periods. This means that strategic decisions may have led to localized exchange relations between junior and senior faculty. However, the effect of same rank was positive and significant in the last period of the shared language network. This may suggest that, at least for shared use of language, strategic decisions may have become less important in the 2000s.

Cross-network effects. The likelihood for tie creation and maintenance in the collaboration network was predicted by an existing tie in the language network, and vice versa, across all periods. These results indicated bidirectional causality (or cyclic causation) between collaboration and shared language: existing shared language ties predicted future collaboration ties, and existing collaboration ties predicted future shared language ties (see effect 19 in Table 1). This suggests there is a feedback process between shared language and collaborations reinforcing each other over time. The effects of collaboration ties on language ties were stronger than vice versa in the first two periods and were of equal strength in the last period (see Table A3 of the online supplement for the comparison of estimates). The balance thus subtly flips from collaborations that have greater influence on shared language in the 1990s to shared language guiding associations equally in the 2000s. This means that there was evidence for the perspective that social structure leads to the development of shared ideas in the first period and that shared ideas and social structure became co-constitutive in the last period.

Academic cultures. The results also reflected a distinct culture of knowledge production in STEM fields and non-STEM fields. The main effect of STEM (effect 11) was strongly positive in the collaboration network in each period. This indicates that faculty in STEM fields were far more likely to develop new and maintain collaborations with colleagues than scholars in the humanities and social sciences. STEM faculty were also more likely to share language with colleagues as indicated by the significant and positive effects in the language networks of each period. However, this difference became significantly smaller ($z = -9.13, p < 0.001$) over time, indicating the increasing use of similar language also among non-STEM faculty.

To illustrate the extent to which the two scientific cultures (STEM and non-STEM fields) differ in shared language, and in what way they constitute distinct communities of collaboration, we summarize for each explanation its relative impact within the two scientific cultures. Figure 2 shows that degree, or the general tendency of faculty to form and maintain collaborations (Figure 2A) and shared language (Figure 2B), had the strongest relative importance in explaining tie formation. For both STEM and non-STEM fields, this accounted for about two-fifths of the decision-making process in both networks and across the three periods. The sizeable contribution of degree for both STEM and non-STEM faculty means that the other factors included in our model explained fewer decisions made in both fields. For non-STEM faculty, their choices to collaborate and share language was weakly related to the other factors in our model: opportunity and homophily con-
Influences on collaborations (1994–2005) and shared language (1994–2005) among non-STEM (first column) and STEM (second column) faculty members (averages and variation shown across time). Shared language = existing tie X in shared language network → new tie Y effect in collaboration network. Collaboration = existing tie X in collaboration network → new tie Y effect in language network. Homophily = same attributes (same gender, same race, and similar age). Opportunity = same environment (same scientific field and same department). Structure = transitive triads effect (clustering), degree of alter effect (centralization), and degree assortativity effect (stratification). Degree = degree effect (density).

Robustness Check: Co-citation Networks

To validate the findings based on our new measure of shared language use, we analyzed the coevolution of collaborations and co-citation as an alternative indicator of shared ideas (see Table A5 of the online supplement). Most of our previous findings reappeared in these analyses. Clustering (transitive triads) and stratification (degree assortativity) reappeared as important structural processes. Interestingly, centralization also played an important role in the co-citation networks; however, the effect of degree of alter was negative here (as opposed to positive in the shared language networks). This shows that faculty who shared citations with many others tended to not attract further co-citations. The effects for academic rank and opportunity replicated likewise and so did the effects for preferences with regard to gender.

Homophily contributed, at most, 10 percent of the tie formation process, whereas the other factors contributed even less.

For STEM faculty, their choices to collaborate and share language were driven by multiple factors, but mainly by network structure (10 percent to 19 percent) and opportunity (6 percent to 15 percent), and to a lesser extent by homophily or preferences (at most 7 percent). Controlling for everything else in the model, the importance of cross-network effects was strikingly low, particularly for non-STEM fields (1 percent at most). For STEM, it was slightly higher but still of relatively lower significance compared with the other explanations (5 percent at most). In sum, these results show that the processes underlying tie formation in the STEM and non-STEM fields were fundamentally different. Collaboration and shared language in the STEM fields were more strongly driven by structural processes such as clustering and centralization, whereas scholars’ preferences played a greater role in the non-STEM fields.

Robustness Check: Co-citation Networks

To validate the findings based on our new measure of shared language use, we analyzed the coevolution of collaborations and co-citation as an alternative indicator of shared ideas (see Table A5 of the online supplement). Most of our previous findings reappeared in these analyses. Clustering (transitive triads) and stratification (degree assortativity) reappeared as important structural processes. Interestingly, centralization also played an important role in the co-citation networks; however, the effect of degree of alter was negative here (as opposed to positive in the shared language networks). This shows that faculty who shared citations with many others tended to not attract further co-citations. The effects for academic rank and opportunity replicated likewise and so did the effects for preferences with regard to gender.

Homophily contributed, at most, 10 percent of the tie formation process, whereas the other factors contributed even less.

For STEM faculty, their choices to collaborate and share language were driven by multiple factors, but mainly by network structure (10 percent to 19 percent) and opportunity (6 percent to 15 percent), and to a lesser extent by homophily or preferences (at most 7 percent). Controlling for everything else in the model, the importance of cross-network effects was strikingly low, particularly for non-STEM fields (1 percent at most). For STEM, it was slightly higher but still of relatively lower significance compared with the other explanations (5 percent at most). In sum, these results show that the processes underlying tie formation in the STEM and non-STEM fields were fundamentally different. Collaboration and shared language in the STEM fields were more strongly driven by structural processes such as clustering and centralization, whereas scholars’ preferences played a greater role in the non-STEM fields.

Robustness Check: Co-citation Networks

To validate the findings based on our new measure of shared language use, we analyzed the coevolution of collaborations and co-citation as an alternative indicator of shared ideas (see Table A5 of the online supplement). Most of our previous findings reappeared in these analyses. Clustering (transitive triads) and stratification (degree assortativity) reappeared as important structural processes. Interestingly, centralization also played an important role in the co-citation networks; however, the effect of degree of alter was negative here (as opposed to positive in the shared language networks). This shows that faculty who shared citations with many others tended to not attract further co-citations. The effects for academic rank and opportunity replicated likewise and so did the effects for preferences with regard to gender.

Homophily contributed, at most, 10 percent of the tie formation process, whereas the other factors contributed even less.

For STEM faculty, their choices to collaborate and share language were driven by multiple factors, but mainly by network structure (10 percent to 19 percent) and opportunity (6 percent to 15 percent), and to a lesser extent by homophily or preferences (at most 7 percent). Controlling for everything else in the model, the importance of cross-network effects was strikingly low, particularly for non-STEM fields (1 percent at most). For STEM, it was slightly higher but still of relatively lower significance compared with the other explanations (5 percent at most). In sum, these results show that the processes underlying tie formation in the STEM and non-STEM fields were fundamentally different. Collaboration and shared language in the STEM fields were more strongly driven by structural processes such as clustering and centralization, whereas scholars’ preferences played a greater role in the non-STEM fields.

Robustness Check: Co-citation Networks

To validate the findings based on our new measure of shared language use, we analyzed the coevolution of collaborations and co-citation as an alternative indicator of shared ideas (see Table A5 of the online supplement). Most of our previous findings reappeared in these analyses. Clustering (transitive triads) and stratification (degree assortativity) reappeared as important structural processes. Interestingly, centralization also played an important role in the co-citation networks; however, the effect of degree of alter was negative here (as opposed to positive in the shared language networks). This shows that faculty who shared citations with many others tended to not attract further co-citations. The effects for academic rank and opportunity replicated likewise and so did the effects for preferences with regard to gender.
and race. In contrast with the shared language networks, being of similar age did not affect faculty’s likelihood to cite the same articles in most periods. Existing collaboration had a sway over shared ideas (co-citations) in the first period, but the cross-network effects became equally strong in the 2000s. Just like in the language networks, we found in the co-citation networks that STEM faculty had a higher likelihood of sharing ideas (cite the same articles) than non-STEM faculty and that this difference became smaller in the later periods. In sum, the robustness check suggests that our new measure of shared ideas among academics produced sensible results.

Discussion and Conclusion

In this article, we have asked what structure scholars’ intellectual communities have and how faculty members structure themselves socially and intellectually within a research university. Sociology of science, organization studies, and social theory have proposed various theoretical accounts for the development of collaboration and the emergence of thought communities. Explanations have focused on scholars’ opportunities to meet (Blau 1973; Kabo et al. 2015), their preferences or strategic decisions to get access to resources or knowledge (Kossinets and Watts 2009; van Rijnsoever et al. 2008), structural processes such as clustering and balkanization (Friedkin 1978; Kadushin 1976; Kuhn 1970), centralization (Azoulay et al. 2010), or stratification (Kagan 2009; Newman 2004). Moreover, modern day sociologists’ theorizing allows predictions about the causal order of social affiliation (collaboration) and culture (shared ideas) (Mannheim 1936; Marx 2000; Simmel 1971; Weber 2002). We provided the first simultaneous test of these various theoretical predictions using unique social network data of a large scientific community at one large private U.S. research university. In so doing, we analyzed ideational structures in terms of shared language use of scientists—something heretofore relatively unexplored (Foster et al. 2015). Moreover, we applied a new and powerful method for longitudinal social network analysis, an approach that is scarcely used to understand the development of shared cultures despite its ability to disentangle various competing hypotheses (Lewis, Gonzalez, and Kaufman 2012; Lewis and Kaufman 2018).

Structural Processes

Results suggest that the network structure of collaboration and of shared language at this university was organized by a continuous process of clustering and centralization on social and intellectual bridges. Clustering in cohesive groups of scholars and of shared ideas was evident in all periods, suggesting stability of research effort (Friedkin 1978; Moody 2004; Moody and Light 2006). In the 2000s, there was also strong evidence for the centralization around some highly active scholars (bridges) but no evidence for stratification. This indicates that these social bridges did not collaborate with each other, which could point toward status competition among successful academics (Lazega et al. 2016). Instead, the intellectual bridges formed an intellectual hub in the language network, which eventually diminished in the last period. These results all reflect the strong presence of interdisciplinary research
in the university and a focus on highly visible and productive scholars. These social brokers, or interdisciplinary stars (Azoulay et al. 2010), were enabled by the university to form clusters of research (labs) and associated with new faculty. And these efforts were often focused into larger research hubs called interdisciplinary centers and institutes. These centers built on an existing intellectual hub and institutionalized similar efforts elsewhere, spreading the style of research of these interdisciplinary bridges to other parts of the university to increase research output (Biancani et al. 2014; Mäkinen et al. 2020).

**Opportunity and Preferences**

The opportunity to interact as a member of the same field and department was among the most relevant predictor of collaborations and shared language. Face-to-face interactions in local contexts have been found to be important success factors in research (Kabo et al. 2014; Lee et al. 2010) and for the influence of other’s ideas (Friedkin 1998; Rawlings et al. 2015). In part, this result not only reflects an effect of the opportunity to meet but also the increased likelihood for overlapping research interest of members of the same field and department.

Net of opportunity effects, we found little evidence of selection into collaborations and shared interests based on racial background but rather saw it emerge from the pressures of gender, age, and academic rank (Kossinets and Watts 2009; Zhang et al. 2018). Men and women were equally likely to collaborate, but collaboration was more likely to take place among faculty of the same gender. Language was also more likely to be shared among scholars of the same gender and this was mainly driven by female scholars using similar language as other female scholars. Concerning age and rank, it seems younger faculty integrated themselves into collaborations and shared intellectual pursuits more than older faculty. This is in line with van Rijnsoever and colleagues’ (2008) finding of fewer collaborations among older scientists.

Higher-ranking persons were most equipped to take advantage of collaboration opportunities at the university. Net of age, tenured faculty were more likely to collaborate and to share language with colleagues than untenured or clinical faculty. In effect, our results suggest that youngish tenured faculty seemed most likely to select into the core of the research community. Moreover, faculty members of the same rank were less likely to collaborate or share language. This is in line with the notion of localized exchange relations, where collaborations are the result of strategic decisions in the pursuit of skills, knowledge, and visibility (Dahlander and McFarland 2013; Melin 2000). However, untenured scholars nearly caught up over time with their tenured colleagues in forging collaborations and sharing language. This, again, fits the efforts of interdisciplinary centers to build off stars and to spread their interdisciplinary mode of collaboration to younger faculty.

**Correspondence between Social Structure and Ideas**

There was evidence for causal effects of collaboration and sharing research language on each other. The causal path was initially stronger from collaboration to shared language in the 1990s and reached parity in the 2000s. This suggests that Marx (2000)
and Mannheim’s (1936) structural argument had greater relevance in academe of
the 1990s; a meeting of persons was more likely to precede a meeting of minds
than vice versa. The equally strong dual causal relation we observed in the 2000s
suggests that research interests and associations co-constituted one another equally
in later years. This is in line with Simmel’s (1971) notion of dualism where form
and content are the flip sides of the same coin. The increasing importance of shared
ideas for the development of scientific discovery may be due, in great part, to
technological advances—the ability to search for texts and identify colleagues of
relevance online as opposed to serendipitously falling upon them via proximity in
hallways and on bookshelves (Evans 2008).

Analysis of relative importance (Indlekofer and Brandes 2013) revealed that exist-
ing collaborations (structure) or shared language (culture) had only very limited
impact on each other. This finding is rather surprising given the prominent interest
of cultural sociologists on the effect of culture on the development of social and
vice versa (Edelmann and Vaisey 2014; Lewis and Kaufman 2018; Lizardo 2006).
Other predictors, such as the opportunity to meet, scientists’ preferences, and par-
ticularly the structural embeddedness turned out to be much more predictive of the
academic discovery process. This shows that failing to account for such alternative
processes might lead to drawing premature conclusions about the interrelation
of collaboration and shared interests. Interestingly, structural processes—such as
centralization, clustering, and stratification—were among the strongest predictors
for the emergence of shared research ideas and became particularly important in
the 2000s. This suggests that a stronger focus on the structure underlying cultural
communities may advance our understanding of their development. Our approach
of considering not only collaboration but also culture as a social structure (Foster
et al. 2015) and modeling them simultaneously may prove particularly fruitful in
future research.

Academic Cultures

We found clear distinctions between the two scientific cultures described by Snow
(1959). STEM faculty were far more likely to engage in collaborations and shared
language than non-STEM faculty. This fits with Kuhn’s (1970) description of the
non-STEM fields as preparadigmatic and lacking a cogent thought style. There
was an increase in collaborations and shared language ties over time, as the entire
university ratcheted up production and non-STEM faculty adopted team science
endeavors and published more in journals and not just books. The steep increase of
shared language ties in the STEM fields from 11 on average in 1994 to nearly 24 in
2005 reflects, to some extent, that the university’s interdisciplinary initiatives began
in STEM fields and that the culture of interdisciplinary team science diffused more
broadly in the university. However, also at the end of our observation period, STEM
faculty continued to have considerably more collaborations and shared language
ties than non-STEM faculty.

The opportunity to interact and, particularly, structural processes appeared to be the more important predictors of collaborations and shared language use
among STEM faculty. Both factors slowly crept upward in salience over time, which
suggests that STEM faculty expanded their associations via research language with proximate colleagues and through clustering and centralization. Structural processes played a much smaller role in the non-STEM fields. Shared language and particularly collaborations were more strongly driven by the opportunity to interact and scholars’ preferences. However, the relative importance of structural processes increased considerably in the 2000s. This is in line with the interpretation of non-STEM faculty starting to adopt the team science approach prevalent in STEM fields.

Limitations and Future Research

Our analysis faces some potential limitations. First, using the same language in academic articles may indicate shared knowledge and influence between two faculty members, but it remains unclear at which point the body of work of two scholars can be considered similar. We used a cutoff value for shared language that was data driven, producing networks that allowed us to apply SAOMs using SIENA. The robustness checks with co-citation networks as an entirely independent and more commonly used measure of academics’ similarity gives confidence in our results. However, we are looking forward to replications of our findings with a different operationalization of shared ideas or research interests among academics.

Second, our data were restricted to faculty members of only one university. Although this data set gave us detailed insights into a large community of scholars and their background characteristics, allowing a unique view on the development of collaboration and thought communities, it remains unclear how well the findings generalize to other academic settings. Our investigated research university has heavily promoted multidisciplinary research in the period from which our data stem (Biancani et al. 2018; Mäkinen et al. 2020). This may have intensified the development of hierarchical structures in which research clustered around academic stars and junior faculty sought collaboration and intellectual integration with established and tenured faculty. Should similarly detailed data of other campuses become available, recent developments in multilevel modeling of SAOMs using SIENA (Ripley et al. 2020) offer a means to compare processes across different contexts.

Replication with data from another campus may also help improve understanding of how much our results were influenced by the dominant position of the medical school at this university. Our visual inspection of the data showed that medical scholars were most active in terms of collaborations and sharing language, both within their field and with other fields. In part, this is because medical faculty publish articles in greater volumes (with many coauthors) and at a greater pace than scholars from some other fields who write books (slow science of the social sciences and humanities) or publish in conference proceedings (fast science of technology, engineering, and mathematics). Moreover, medical faculty at this particular university rely on grants for research (and thus collaborations), whereas many faculty members from other schools do not.

Academics’ collaborations and adoption of shared ideas may take different forms in the larger scientific community than they do within a single institution. Scholars can be academic stars within their discipline and attract many collaborators
from other universities and inspire many colleagues from around the world to adopt their ideas. These scholars might, however, not be well embedded in the academic network of their own institutions and instead focus all attention on outside collaborators. Similarly, faculty members that form social and intellectual hubs at their own university may not be well embedded in the larger scientific community. Thus, the structure of collaboration and shared idea networks across institutions and across disciplines may differ systematically from what we found within the one university that we studied here.

In sum, our approach of modeling the coevolution of collaboration and shared ideas in the form of social structures provides many novel insights in the academic discovery process. We are looking forward to future applications of this approach in other settings to test the various theoretical accounts for the development of collaboration and the emergence of thought communities.

Notes

1 More recently, Kagan (2009) has expanded the dichotomy and argues there are three academic cultures, in which STEM, social sciences, and humanities have distinctive logics of research. Clearly there is a plenitude of theories concerning styles of scholarship and research dating back to Kant’s Metaphysics of Morals. Even depictions of sociological theory characterize different national styles (British, German, and American strands with different styles—see Levine 1995). The point here is not which theory of thought styles most pertains but rather whether actors variably adhere to different thought styles and if that intellectual adherence influences interpersonal relations and vice versa.

2 We assessed goodness of fit by comparing the observed networks to simulated networks based on the models in Table 1 with regard to auxiliary statistics that were not explicitly modelled (Lospinoso 2012). This approach indicates how well the models captured the dynamics of the collaboration and shared language networks. We focused on one particularly important auxiliary statistic, the degree distribution, that represented here the distribution of the number of collaborators or colleagues with similar language use in the entire network. The p values from the Mahalanobis distance tests were all less than 0.05, which indicated overall poor fit of the models for degree. Visual inspections of the goodness-of-fit plots (see Figure A3 of the online supplement) showed that the models underestimated the number of people without collaborators and the number of people who did not share language with anyone (isolates), whereas the amount of people with only few collaborators or who shared language with only few were overestimated. Unfortunately, poor model fit is typical for applications of SAOMs to large networks because “larger networks may contain greater internal heterogeneity that current models are unable to satisfactorily represent” (Lewis and Kaufman 2018:1736).

3 Only the effects that were significantly different from each other across periods are discussed. Table A4 of the online supplement presents significance tests of the same effect in the same network type between periods. We tested for differences between parameter estimates using z score tests, assuming equal variance across groups (with estimates $\hat{E}_{3,a}$ and $\hat{E}_{3,b}$ and standard errors $s.e_{a}$ and $s.e_{b}$, respectively), with the following formula: 

$$
\frac{(\hat{E}_{3,a} - \hat{E}_{3,b})}{\sqrt{(s.e_{a}^2 + s.e_{b}^2)}}
$$

which under the null-hypothesis of equal parameters has an approximately standard normal distribution (Rambaran et al. 2013; Ripley et al. 2020).

4 Interpretation of the homophily parameters for age, gender, and race requires taking the main effects into account as well. This is done with ego-alter selection figures in
Appendix E of the online supplement. Note that for undirected networks the model cannot distinguish senders from receivers, meaning that we can only calculate the odds of tie formation between, for instance, males, females, and between males and females (without direction).

References


Kabo, Felichism W., Natalie Cotton-Nessler, Yongha Hwang, Margaret C. Levenstein, and Jason Owen-Smith. 2014. “Proximity Effects on the Dynamics and Outcomes of Scientific


**Acknowledgments:** The data used in this project were collected by the Mimir Project conducted at Stanford University. This work was supported by the Office of the President at Stanford University and by NSF Award #0835614. This project was also funded by the European Commission (FP7-PEOPLE-2011-IOF, Grant Agreement Number 299939) and the Netherlands Organisation for Scientific Research (NWO PROO grant 411-12-027). We thank Gert Stulp for his useful suggestions on how to create the figures presented in this article and Tom Snijders for his useful suggestions on how to improve the model specification. Please direct all correspondence to Tobias H. Stark.

**Tobias H. Stark:** Department of Interdisciplinary Social Science/ICS, Utrecht University. E-mail: t.h.stark@uu.nl.

**J. Ashwin Rambaran:** Department of Psychology and Combined Program in Education and Psychology (CPEP), University of Michigan. E-mail: jrambara@umich.edu.

**Daniel A. McFarland:** Graduate School of Education, Stanford University. E-mail: mcfarland@stanford.edu.