

Shared Leadership in Scientific Teams:

Heterogeneity vs. Homogeneity

Huimin Xu

School of Information, University of Texas at Austin, Austin, TX 78701, USA

Meijun Liu

Institute for Global Public Policy, Fudan University, Shanghai 200433, China

Yi Bu

Department of Information Management, Peking University, Beijing 100871, China

Shujing Sun

School of Management, University of Texas at Dallas, Dallas, TX 75080, USA

Yi Zhang

Faculty of Engineering and IT, University of Technology Sydney, NSW 2007, Australia

Chenwei Zhang

Faculty of Education, University of Hong Kong, Hong Kong 999077, China

Daniel E. Acuna

School of Information Studies, Syracuse University, Syracuse, NY 13244, USA

Steven Gray

School of Business, University of Texas at Austin, Austin, TX 78712, USA

Eric Meyer

School of Information, University of Texas at Austin, Austin, TX 78701, USA

Ying Ding

School of Information, University of Texas at Austin, Austin, TX 78701, USA

Corresponding author: Ying Ding (ying.ding@ischool.utexas.edu).

Shared Leadership in Scientific Teams:

Heterogeneity vs. Homogeneity

Abstract: Creative, complex, knowledge discovery-based scientific tasks require intensive and interdependent teamwork, thus having more than one leader is crucial. This paper aims to advance our understanding of shared leadership in scientific teams. We define three kinds of leaders, junior (10-15), mid (15-20), senior (20+) based on career age. By considering the combinations of any two leaders, we distinguish shared leadership to heterogenous ones in different age cohorts and homogenous ones in the same age cohorts. Drawing on 521,693 Computer Science (CS) teams with two leaders in the DBLP dataset, we identify that heterogenous shared leadership brings higher citation impact for teams than homogenous shared leadership. Specifically, when junior leaders are paired with senior leaders, it significantly increases team citation. The result is significant when we control team size, year, mean team career age, gender dominance, and industry/academia collaboration. We explore the patterns between homogenous leaders and heterogenous leaders from team scale, expertise composition, and knowledge recency perspectives. Compared with homogenous leaders, heterogenous leaders are more adaptive in large teams, have more diverse expertise, and trace both newest and oldest references.

Keywords: Shared Leadership; Homogenous Leaders; Heterogenous Leaders; Scientific Collaboration; Team Impact

Introduction

The complex challenges of scientific discovery, novel syntheses of knowledge, and the large scale data and computing demanded for scientific breakthroughs, render us to rethink whether the traditional sole leadership is still suitable to scientific teams, or shared leadership where multiple leaders engage in leadership activities is more desirable to creative tasks. Intrinsically, innovation fuels scientific discovery that advances industry invention, government policy and public good (Ahmadpoor & Jones, 2017; Yin et al., 2022). Extensive research focuses on the

novelty of scientific publications and strives to maximize the team innovation (Fortunato et al., 2018). In scientific publications, creative ideas arise from recombining existing knowledge in an unconventional way (Uzzi et al., 2013), building new connections to bridge different concepts (Hofstra et al., 2020), and overshadowing prior work by attracting subsequent attentions (Wu et al., 2019). Prior literature suggests that gender diversity in team composition (Nielsen et al., 2018), flat and egalitarian team structure (Xu et al., 2022a; Xu et al., 2022b), agile small teams (Wu et al., 2019), and fresher teams with weak collaboration ties (Zeng et al., 2021) could enhance team innovation. Internally, the demands of the scientific workforce are different. Researchers are highly educated workforce who contribute intellectual capitals. These knowledge workers are not merely satisfied with material wealth (e.g., money), they also have the emotional and professional pursuit, such as leadership impact (Pearce et al., 2004). The absolute authority of sole leader will stifle the creativity of scientists, especially for those coming from different domains that norms/practices and thought processes are dramatically varied. Drawing on social identity theory, Hogg (2008) emphasized the identity dimension of leadership. Social identity defines an individual's identification through social categorization process (Tajfel et al., 1979). In innovative work contexts, group belongingness can foster individuals carrying different knowledge to transform personalized identity to serve the collective creative identity as a means to maintain positive self-image in groups (Hirst et al., 2009). Leaders have the apparent paradoxical identity that they not only conform but also oppose to group norms (Hollander, 1958). Leaders who fit the group norms best can gain disproportionate popularity, trust and legitimacy from followers (Hogg, 2008), and thus accrue enough innovation credits to be transformational and steer the group in new direction (Abrams et al., 2008). The innovation of leaders would directly motivate the creative efforts of followers in the whole team (Hirst et al., 2009). Externally, scientific teams need to keep pace of the rising trend of interdisciplinarity and internationalization. Interdisciplinary projects involve experts from different fields and require greater labor of division (Haeussler & Sauermann, 2020). For example, a successful AI health project needs the knowledge fusion from an AI expert and a medical expert. International cooperation brings higher transaction costs, language barrier and cultural difference, which would hinder the creativity of research (Wagner et al., 2019). A sole leader is hard to own management skills, professional knowledge especially in interdisciplinary areas, and digital literacy to handle all problems and achieve transformative innovation. Multiple leaders can

exhibit their respective areas of skills, expertise and abilities to evolving tasks (Pearce & Conger, 2002). Also, the pursuit of innovation is inherently a risky endeavor (Van Knippenberg, 2011). Relying on one person to make decision will increase the risk in face with fierce global competition. Thus, multiple leaders in shared leadership can provide a more innovative climate for knowledge-intensive research than unitary leadership (Pearce et al., 2004). But few literature contributes to the understanding of shared leadership in scientific team innovation and impact.

Labor of division in knowledge production determines the role of leading and supporting. By analyzing self-report contributions in publications, Xu et al. (2022a) measured the number of leaders with keywords (e.g., design, lead, supervise) in proportion to team size. They found that flat team structure with more leaders in teams is conducive to disruptive innovation and long-term impact. Although they studied the role of shared leadership, they treated all leaders the same and ignored the difference and complementarity between leaders. Yet, leaders in different seniority levels are equipped with distinct leadership skills (Mumford et al., 2007). The order of authorship is another common manifestation of leadership (Chinchilla-Rodríguez et al., 2019; Liu et al., 2014). Previous literature usually takes both first authors and last authors as leading authors, which is a form of shared leadership. First authors and last authors typically contribute more than middle authors in publications in all disciplines (Larivière et al., 2016). Thus, authors who occupy such important positions win more recognition, credits and publication bonus (Fuyuno & Cyranoski, 2006). Geographical proximity between first-last authors from the same institution or country is positively related to paper citation since the last author could provide direct supervision to the first author (Hsiehchen et al., 2015; Lee et al., 2010). We can sense the first author and the last author exhibit different levels of seniority and contributions, but prior research seldom distinguishes them. In experimental research, the convention is first authors are usually the students supervised by last authors (Costas & Bordons, 2011). Young scholars who perform most of the technical parts tend to be the first authors, whereas senior scholars who conceive the conceptual design are usually the last authors (Larivière et al., 2016). In this sense, first authors do not fit the definition of scientific leadership that leaders have the ability to obtain resources, have the expertise to initiate and develop projects, and have higher impact and productivity (Chinchilla-Rodríguez et al., 2019). Career age is a direct indicator of contribution. Senior researchers tend to shoulder the leadership responsibility. Drawing on Merton's principle

of cumulative advantages (1973), the gap between senior researchers and junior ones is exacerbated since senior authors are more productive and are more cited in publications, accumulate enough experience and knowledge in research, have the ability to provide rich funds and advanced equipment to support research. Milojević et al. (2018) defined different researchers based on career age, from junior (1-10), early-career (11-15), midcareer (16-20) to full-career scientists (>20). As authors shift from junior to full career, they shoulder more responsibilities in leading the team than supporting the team. Specifically, Larivière et al. (2016) calculated the mean age of scholars who performed the experiments (6.8 years), analyzed the data (10.5 years), and designed, wrote, and contributed materials (12.6 years). At different career stages, researchers have corresponding status or positions. Bu et al. (2018a) defined milestones of CS researchers that they usually go through PhD studies in the first five years, postdocs or assistant professors in the following five academic years, then assistant/associate professors (Year 10-15), next associate/ full professor (Year 15-20), finally finish their career. Hence, we can see different time points mark different levels of maturity in leadership. As scholars age, their leading positions in the scientific hierarchy rise, from young leaders, midcareer leaders, to senior leaders. However, current research seldom distinguishes scientific leaders with different career ages and their roles in shared leadership.

Shared leadership also have shortcomings, namely low efficiency (Pearce, 2004), diffusion of responsibility (Zhu et al., 2018), competition and conflicts (Weiss et al., 1992). O'Toole et al. summarized the success of co-leader model lies in “how complementary the skills and emotional orientations and roles of the leaders are” (2002, p. 71). Distinguishing heterogenous leaders and homogenous leaders is a key to understanding the advantages and disadvantages of shared leadership in science. Diversity and complementarity from heterogenous leaders strengthen shared leadership, including handling complex tasks (Carpenter & Fredrickson, 2001). The role overlap from homogenous leaders can weaken shared leadership, as noted above. Demographic-based characteristics (e.g., age), is a crucial aspect of heterogeneity and homogeneity in team leaders (Sperber & Linder, 2018). In scientific teams, when leaders in the same career age cohorts are paired, they are homogenous leaders. Otherwise, they are defined as heterogenous leaders. Few research analyzes the role of shared leadership through the lens of homogenous/heterogenous leaders in different team scales. Previously, team science researchers

compared the difference of division of labor (Meyer & Schroeder, 2015; Whitley, 2000), cognitive resources (Bantel & Jackson, 1989), novelty and disruption (Wu et al., 2019), communication costs (Staats et al., 2012) between large teams and small teams. From expertise perspectives, researchers have a homophily tendency to work with individuals who have similar expertise (Ding, 2011; Kraut et al., 2014). Still, expertise diversity in teams can help boost labor of division and solve complex problems (Giuri et al., 2010; Miles & Kivlighan, 2010). At the knowledge recency level, extensive literature discusses the citing practice for researchers with different career age. Senior researchers tend to cite older literature whereas junior researchers follow recent ideas (Merton, 1973; Gingras et al., 2008; Cui et al., 2022). When senior researchers are paired with junior ones, the average citing age decreases and the innovation level increases (Cui et al., 2022). This body of research studied the homogeneity and heterogeneity for all members in the whole team, but did not emphasize the similarities and differences for leaders. How heterogeneous leaders and homogeneous leaders influence expertise diversity, knowledge recency and the ultimate performance is seldom explored in scientific teams.

This paper studies the relationships between shared leadership and team performance in creative scientific discoveries, with the special focus to highlight heterogeneous and homogeneous shared leadership to identify unique patterns. This study addresses above research gaps by taking the Computer Science (CS) as a test field to compare heterogeneous and homogeneous shared leadership in team performance. This paper is structured as the followings. Section 2 highlights the related literature. Section 3 outlines the research methods. Section 4 details the findings and provides comparison and discussions. Section 5 concludes the study and points out future research directions.

Literature Review

Shared Leadership

Leadership is evolving dynamically from individual endeavor to shared efforts (Voss, 1999). Traditionally, leadership emphasizes the individualism in creative discovery and research tasks, such as irreplaceable roles and leadership traits. For example, Zucker & Darby (1996)

investigated scientists in the genetic sequence discoveries who were productive in publications and influential in citations. These star scientists not only lead the technological innovation but also bring commercial success when they collaborate with companies, such as product development and marketing, and employment growth. Leaders who engaged in the transformational leadership have distinct characteristics of idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Bass et al., 2003). By surveying the R&D teams in software development companies, Gumusluoglu and Ilsev (2009) found that these characteristics could encourage employees to generate and implement creative ideas, as well as develop new products and services at the organizational level. However, faced with knowledge-intensive work which requires high interdependency, creativity and complexity, it is hard for a person to handle all tasks (Pearce, 2004). Extant shared leadership research in different contexts stresses that team creativity is driven by the influence of multiple members and their interaction, instead of an individual's efforts. In information technology companies, when team members have high tendency to speak up and meanwhile accept others' opinions, and have high confidence in their expertise, they would be more likely to participate in leadership activities and come up with creative solutions and novel ideas (Ali et al., 2020).

In Wikipedia community, Zhu et al. (2011) analyzed 4 million messages from editors, and distinguished the roles of different leaders in the shared leadership process. Core leaders emphasize on the emotional support, whereas peripheral leaders are task-driven and provide positive or negative feedbacks. In general, creative tasks (e.g., visual puzzles, brainstorming, negotiating), the collective intelligence of a group to perform tasks is highly correlated with the equal speaking-turn distribution among team members instead of the most intelligent individual (Woolley et al., 2010). In scientific publications, the trend that multiple leaders collectively manage the scientific teams and share the leading power is rising. A signal is that the percentage of equal first authors and corresponding authors increases linearly in scientific publications (Hu et al., 2009). First authors usually contribute most to the projects and last authors shoulder the corresponding and supervision roles in experimental science, but it varies vastly among fields and regions (Liu & Fang, 2014). For example, the mathematics field list authors alphabetically, whereas leaders in Egypt appear in the first positions and meanwhile serve as corresponding authors (Liu & Fang, 2014). In addition to identifying leaders through authorship, Xu et al.

(2022a) defined leaders as those contributed to conceiving, designing, supervising, and writing in publications, and suggested that flat structure with multiple leaders was better for team novelty and long-term impact than sole leadership. Nevertheless, how multiple leaders assume shared leading roles interactively in scientific teams is seldom discussed.

Heterogenous vs Homogenous Shared Leadership

Researchers committed to studying the heterogeneity/homogeneity of leaders in creative teams, manifested through variables such as age, expertise, character (O'Toole, 2002), tenure, educational and functional composition (Hambrick et al., 1996), leadership styles (Zhu et al., 2011), and status (Watts, 2010). In the US airline companies, when leaders are heterogenous in function, tenure and education background, they have great vision and external connections to initiate creative frequent-flyer programs and expand new markets (Hambrick et al., 1996). But the heterogeneity could produce conflicts between leaders and thus impede the response process. In the online knowledge production community, heterogenous leaders with different leadership styles could shoulder different labor of division (Zhu et al., 2011). Transformational leaders are socially oriented and person focused, focusing on inspirational motivation and intellectual stimulation, whereas transactional leadership is task driven, offering rewards for followers who are compliant but punish those who are deviant (Bass et al., 2003). The combination of both styles can not only motivate followers but also ensure the fulfill of projects in the community. In the open source software community, O'mahony & Ferraron (2007) classified leadership into hands-off leader, technical manager, visionary leader, organization builder, and organization leader based on their concerns to organizations. At different stages of organizational development, the heterogeneity of these leaders enables them to rotate leadership roles and apply their particular skills to meet evolving demands. In the scientific discovery, Watts (2010) indicated that heterogenous leaders (e.g., novice leader and experienced leader) have more potential benefits than homogenous leaders (e.g., leaders with equal status). Firstly, this is beneficial to create an environment where leaders can respect and learn from each other “characterized by intellectual generosity” (Watts, 2010, p. 336). Early-stage leaders are equipped with digital skills, master the most up-to-date methods and technology in research, and have fresh perspectives (Powell, 2021), which can be complementary to senior leaders. Junior leaders can practice their soft skills (e.g., management, communication, collaboration) and build

connections, while working with senior leaders. Secondly, the heterogeneity of shared leadership can encourage student-centered supervisory ways. Phillips & Pugh (2015) suggested that in some universities, the mode that a leader who leads the project and another leader who provides the support is better than the mode of two equals. Both emotional and rational elements are required for successful supervisory practice (Firth & Martens, 2008). Thirdly, leaders' heterogeneity can advance interdisciplinarity. By interviewing 12 respondents about what kind of information exchange occurs in interdisciplinary collaboration, Haythornthwaite (2006) summarized that leaders from science and social science fields can boost knowledge exchange, learning process, research collaboration, new idea generation, and network contacts. The harmonious relationship will be broken in some cases (Phillips & Pugh, 2015). For example, when two leaders have competing ideas or conflicts, the student might feel confused, distracted or contradictory. Students are more likely to see their leaders separately rather than together. It will increase the difficulty of reaching the consensus. However, the discussion about the advantages and disadvantages of homogenous/heterogenous leaders only limit to small scale groups, qualitative methods. How to scale up the results within millions of scientific teams is still a new direction.

Shared Leadership and Team Scale

Literature in team science always highlights the differences between large teams and small teams. Large teams enable clear division of labor and specialization in solving large problems particularly when there is a high degree of certainty and independency about the tasks to be accomplished (Meyer & Schroeder, 2015; Whitley, 2000). Specifically, by studying the collaboration practices of sociologists, Hunter & Leahey (2008) concluded quantitative methods were more suitable for large teams since they enabled easily-divisible tasks, like collecting, cleaning, coding and analyzing data. Meanwhile, large teams have extra cognitive resources because they can be more diverse in skills and knowledge than small teams (Bantel & Jackson, 1989). For example, Larivière et al. (2015) analyzed team size using three indicators, number of authors, number of institutions, and number of countries. They found that as team becomes larger and more diverse, they can receive more influence in citations. However, large teams are more likely to produce conflicts due to increased communication costs and decreased support (Staats et al., 2012). Although large teams can generate more ideas than small teams in experimental brainstorming settings, being exposed to too many ideas can lead to distraction and

information overload (Paulus et al., 2013). By analyzing millions of teams in paper publication, patent invention and software development, Wu et al. (2019) suggested that small teams can achieve disruptive innovation although they do not receive much attention, compared with the incremental innovation in large teams. Because of the inherent task complexity and coordination requirements, a senior leader with rich research experience, research resources and team management experience would be more qualified to build large teams. Senior leaders have more knowledge, funding, connections, prestige, power, such that, they can enlarge their teams by attracting or hiring collaborators (Merton, 1973, Xu et al., 2022b). In organizations, there are different requirements for junior, mid and senior managers in leadership skills (Mumford et al., 2007). Compared with junior leaders, senior leaders have higher business (e.g., personnel resources and financial resources) and strategic skills (e.g., visioning) requirements. Even though junior leaders are less experienced and lack resources, they could provide more direct supervision, such as detailed technical methods (Powell, 2021). Junior leaders tend to have a fewer students, thus they can invest more time and energy to train their students. In such small teams, cooperation is agile, which can lead to effective communication and adaptability (Katzenback & Smith, 1993). Current research focuses more on the comparison between small and large teams, how to explore the roles of shared leadership in these teams with distinct size is still unknown.

Shared Leadership and Expertise Diversity

Researchers always take team members' expertise into consideration while discussing collaboration. In the scientific teams, the common methods to mine research expertise are based on keyword extraction or topic modeling in papers. The main conclusion of this line of research is researchers' expertise homophily is much easier to boost collaboration but not necessarily lead to better performance. Kraut et al. (2014) calculated the similarity between the papers of one author and those of other authors by embedding the terms in abstracts into a semantic space, and concluded that authors sharing similar expertise are more likely to collaborate. Applying the same principle, the co-author recommendation systems are built based on authors' recent publication keywords and thus seek for similar co-authors (Sie et al., 2012). Likewise, Ding (2011) found that productive authors in the information retrieval field tend to share similar research expertise with collaborators identified by the topic modeling method. Extending the top

100 productive authors to all authors in the information retrieval fields, Zhang et al. (2016) suggested common research interests are not the dominating factor in collaboration, and authors with different research interests can work together. Bu et al. (2018b) explored the relationship between authors' impact and his collaborators' topic diversity in computer science. They concluded authors who have more impact (i.e., high h-index) tend to collaborate with authors with diverse research interests. In the open source software teams, expertise heterogeneity is positively associated to project activities, such as fixed bugs and patches, released new files (Giuri et al., 2010). Specifically, the expertise and expertise level are different for core members, who are highly experienced and committed to the projects, and for periphery members, who participate in the community and provide support for projects. Projects benefit from the heterogeneity of "specialists" and "generalists", "low-skilled" and "high-skilled" members (p. 67). High expertise overlapping does not boost communication, especially when all members share the same language and the common project goal. Meanwhile, it is not beneficial for labor of division in projects (Giuri et al., 2010). In large corporates, shared leadership has become a necessity, since co-leaders can help compensate for the weaknesses of others (O'Toole, 2002). The complementarity of their expertise corresponding to different roles and tasks can help solve complex business issues. Miles & Kivlighan (2010) asserted when coleaders share similar cognition but hold different expertise or behaviors, they can reduce conflicts, facilitate interaction and increase engagement in teams. Even though science of science researchers consider the topic diversity within collaborators, they do not discuss the expertise composition of multiple leaders. Leaders' research interests are more important in research direction and can indirectly influence students' research interests. Especially for the shared leadership, how research expertise influences the collaboration of homogenous leaders and heterogenous leaders is unexplored in the scientific context.

Shared Leadership and Knowledge Recency

A paper's references represent a team's accumulated knowledge about a specific topic. Members in the team contribute differently to the knowledge base, for example, career age influences researchers' citing practice. There exists age-stratified difference in the acceptance of new ideas (Merton, 1973). Young scholars cite more recent references than older scholars. Similarly, for around 6,000 Quebec university researchers aging between 28 to 40, their reference lists include

a higher percentage literature within 5 years. From 40 until retirement, these old researchers are not active in following new ideas (Gingras et al., 2008). Milojević (2012) studied the reference citing difference for researchers who are active in core journals in five fields astronomy, mathematics, robotics, ecology and economics. She found that senior researchers cite older references on average since they accumulate more experience in the field and thus read older and foundational literature. But, they still follow cutting-edge literature like junior researchers. The limitation of this research is that Milojević only analyzed the top researchers in top journals in five disciplines. Cui et al. (2022) generalized this finding in all fields in the MAG dataset. Through uncovering individual scientists' career, they found that aging scientists favor to cite older references and their ideas still stay in the original state when they were young. What's worse, by classifying citations into constructive or contrast, they found senior researchers even challenge against new ideas. Aging brings social and cultural resistance. Kuhn purposed that science advances in a revolutionary way rather than an incremental way (1970). These scientific revolutions are started by junior scientists who are very young and new to the field. But the glory and shine of star scientists might prevent those newcomers and outsiders from challenging their authority (Azoulay et al., 2019). Azoulay et al. (2019) found that the premature death of superstar life scientists can lead to a mark increase of outsiders who contributed to the field. The reason why the older resists the younger is because they are restricted "in his response to innovation by his substantive and methodological preconceptions and by his other cultural accumulations" (Barber, 1961, p. 601). Higher proportion of older researchers tend to lead to older citations (Barnett & Fink, 2008). But working with young scholars could help them slow the aging rate and cite more recent reference, thus boosting science advance (Cui et al., 2022). Whether these conclusions similarly applied to shared leadership that heterogenous leader combinations can help teams follow recent research ideas and meanwhile keep trace of old sources still needs to be tested.

Methods

Dataset

We choose 4,894,081 papers (until 2020 April) from DBLP¹, which provides major journals and conference proceedings in the Computer Science field. The bibliographic information include unique paper id, publishing year, and authors. Tang et al. (2012) did the author name disambiguation based on the DBLP and updated the newest version online². Given we want to analyze shared leadership in teams, we exclude papers with sole author and keep papers with at least 2 leaders whose career age are above 10 (see Measures for leader definition). There are 1,092,035 papers with multiple leaders left. Among these shared leadership papers, papers with two leaders occupies 69.7% (761,161). Thus, we take the most typical two leader combinations as the representative of shared leadership in this paper. To define the leaders, we not only control career age above 10, but also restrict the positions in the authorship. By tracing the publications of ACM (The Association for Computing Machinery) fellows, Fernandes et al. (2022) concluded that these CS leaders typically place their names in the right side of the bylines, say, last positions. Besides first and last positions, the second-to-last position is the most important position of the paper, which suggests seniority and leadership (Helgesson & Eriksson, 2019). Hence, we select shared senior authors (above 10) in the first and last position or in the last two positions, given leaders usually occupy these positions in the CS field. Finally, there are 521,693 (68.5%) shared leadership teams left.

Measures

After dealing with the author name disambiguation in papers, we get the unique id of each author, thus we can trace the career age of each author after he/she published the first paper. For example, an author published his/her first paper in 2000, then his/her career age is 5 in 2005. Based on the definition of different leaders (Bu et al., 2018a; Larivière et al., 2016; Milojević et al., 2018), we firstly define leaders as those researchers whose career age are above 10, which suggests that they have accumulated 10-year research experience. Then, we distinguish three

¹ <https://dblp.org/>

² <https://www.aminer.cn/citation>

kinds of leaders, junior leader (≥ 10 & < 15), mid leader (≥ 15 & < 20), and senior leader (≥ 20). To differentiate shared leadership with two leaders, there are six different combinations among these three kinds of leaders, two junior leaders (2,0,0), two mid leaders (0,2,0), two senior leaders (0,0,2), one junior leader and one mid leader (1,1,0), one mid leader and one senior leader (0,1,1), and one junior leader and one senior leader (1,0,1). Leaders within the same age cohort are defined as homogenous leaders, whereas leaders belonging to different age cohorts are heterogenous leaders. Meanwhile, we limit these two leaders' positions as the first-last or the last two positions in publications. We describe the number and median value of career age difference between two leaders in each combination.

Table 1. The descriptive variables of leader categories

Leader Categories	Number	Median Value for Difference of Leaders' Career Age
(2, 0, 0)	92,369	1.0
(0, 2, 0)	41,221	1.0
(0, 0, 2)	54,573	4.0
(1, 1, 0)	115,114	5.0
(0, 1, 1)	85,179	8.0
(1, 0, 1)	133,237	12.0

To control confounding variables in the relationship between leaders and team impact, we use the multivariable regression below (Equation 1 & 2):

$$Team\ Impact_i = \alpha + \beta_1(Leader\ categories_i) + \beta_2(Controls) + e_i \quad (1)$$

$$Team\ Impact_i = \alpha + \beta_2(Difference\ of\ Leaders'\ career\ age_i) + \beta_4(Controls) + e_i \quad (2)$$

The dependent variable $Team\ Impact_i$ is the value for 2-year or 5-year citation percentile of a team i. The continuous independent variables are $Leader\ Categories_i$ in equation 1 and

Difference of Leaders' career age_i in equation 2. There are six different leader categories shown in Table 1. *Difference of Leaders' career age_i* means the career age difference of two leaders within a team. For example, the senior leader's career age is 20, whereas the junior leader's career age is 11, then the difference is 9. Team size is considered as control variable. To control the time-invariant factor, publication year is treated as fixed effect. β represents coefficient for different variables. Other controlling variables include male/female/equal and education/company/combined. Teams are divided to male-dominated, female-dominated, and equal categories. For instance, male-dominated means the identified male number is larger than the identified female number in a team. We apply the Bert-based model trained by Acuna & Liang (2021) to predict authors' gender information. We classify teams into pure education, pure company, and combined. Pure education (company) means teams are made up of all researchers from academia (industry) institutions, whereas combined teams have both academia and industry researchers. Using methods from Manjunath et al. (2021) to match authors' institutions with eight categories in the Global Research Identifier Database (GRID), including government, education, company, facility, healthcare, nonprofit, archive, and others. We merge the health care (mainly universities) and education into education and keep the company category. e_i represents the residual.

Results and Discussion

Heterogenous vs homogenous shared leadership

Fig 1a is an illustrated case of an independent team where two leaders and two students collaborate together. Fig 1b shows six different leader combinations where two leaders in shared leadership belong to different junior, mid and senior categories. Leaders belonging to the same age cohort are homogenous teams, (2,0,0), (0,2,0) and (0,0,2), whereas leaders belonging to the different age cohort are heterogenous teams, (1,1,0), (1,0,1) and (1,0,1). We want to observe the citation patterns in heterogenous and homogenous shared leadership. To reduce citation inflation effect, we calculate the 5-year citation percentile for all papers published in the publication year. It is a relatively fair way to compare citations for papers published in different years. High (low) citation percentile means high (low) citations in the year, scaling from 0 to 100. For instance, the

5-year citation percentile of a paper published in 2013 is 99%, which means this paper belongs to the top 1% most frequently cited publications in 2013. Meanwhile, the distribution of leader categories is not even (Table 1), the number of homogenous leader combinations, such as (0,2,0), (0,0,2), are less than heterogenous leader combinations. To reduce the effect of uneven frequency distribution, we randomly choose 10,000 papers from each group for 10 times, and then calculate the average citation percentile. The results in Fig 1c show that heterogenous shared leadership has higher citation percentile than homogenous shared leadership within 5 years on average. Besides observing the difference in categories, we can also extend the categorical variables to continuous variables. Career age serves as a proxy of power since it represents experience, resources, and wisdom (Xu et al., 2022b). The career age difference of two most senior leaders corresponds to the measure of power distance between the most powerful person and the next powerful person (Eisenhardt & Bourgeois, 1988). When we calculate the specific career age difference between two leaders, we find the larger age difference in leader combinations is, the higher citation percentile is (Fig 1d). It also suggests that heterogeneity of leaders bring more citations. We choose 518,262 papers where the first author is a junior researcher (career age <10) among those with two leaders. To clearly observe the difference between homogenous leaders and heterogenous leaders, we normalize the mean 5-year citation percentile with z-score methods, $z = (x - \mu) / \sigma$, where x is the raw value, μ is the mean value of the population, and σ is the standard deviation value of the population. We conclude that heterogenous leaders are better than homogenous leaders in paper citations when we control the career age of students, from 1 to 10 (Fig 1e). This finding is consistent with the conclusions from Watts (2010) that the assemble of junior leaders and senior leaders outperforms leaders of similar age in thesis supervision. Heterogenous leaders in scientific teams can bring complementarity and diversity to handle ambiguous, creative, complex and interdependent tasks (Carpenter & Fredrickson, 2001; Hambrick et al., 1996).

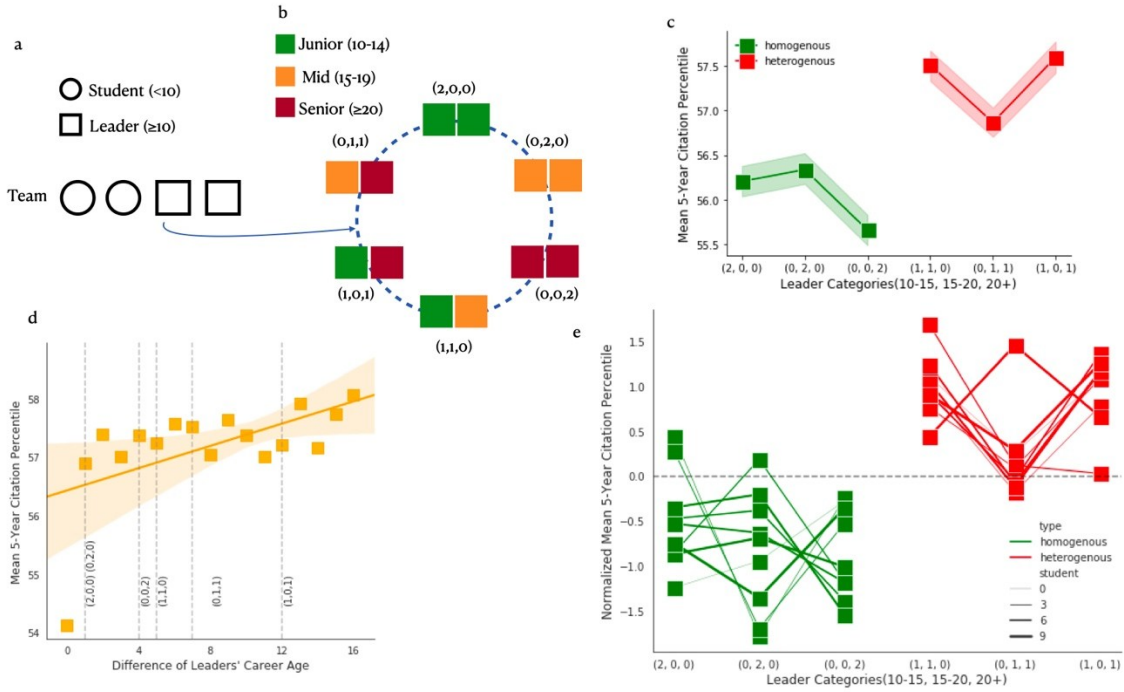


Fig 1. **a**. An example of a team consisting of two students (career age less than 10) in circle shape and two leaders (career age greater than or equal to 10) in square shape. **b**. Distinguish leaders with junior (10-14, green), mid (15-19, orange) and senior (20+, red), and categorize two leaders into six leader combinations. **c**. Heterogenous leaders (red) in career age have more citations than homogenous leaders (green). We calculate papers' citation percentile and then aggregate the mean value in the six different leader categories within 5 years. Bootstrapped 95% confidence intervals are shown as translucent bands. **d**. Larger difference in career age within leaders brings higher citations. We calculate papers' citation percentile in the publication year and then aggregate the mean value based on the difference of leaders' career age (from 0 to 16 in terms of integers) within 5 year. We fit these mean values with a linear regression model. Bootstrapped 95% confidence intervals are shown as translucent bands around the regression line. We label the median value for six leader categories of leaders' career age difference with grey dashed line. **e**. Leaders' influence for different young scholars aging from 1 to 9. We calculate papers' citation percentile in the publication year and then aggregate the mean value in six different leader categories based on papers' first author age within 5 year. The gradient color of lines represents the scale of students' career age, darker lines means higher career age. We normalize the mean 5-year citation percentile with Z-score methods. The grey dashed line clearly

differentiates the citation difference of homogenous leaders (green) and heterogenous leaders (red).

In Table 2, when we control team size and year fix effect, the coefficient values increase in five leader categories within 2 year citation percentile (1) and 5 year citation percentile (4) , compared with the (2,0,0) category. These categories are ranked based on the median value of leaders' career age difference (Table 1). This is in line with the results in (2) and (5), which shows a significant linear correlation between the difference of leaders' career age and team impact ($P < 0.001$). We also test the correlation between any two pair leader categories in (1) and (4). Through F test, all leader categories have significant difference, except (1,1,0) and (1,0,1). In models (3) and (6), we control more variables, mean team's career age, male/female/equal dominated teams, education/industry/combined teams, the result shows that leaders' career age different is still positively related to team 2-year citation and 5-year citation. Team's mean career age is positively associated with 2-year team citation. We also have other interesting findings, male-dominated teams have more citations than female-dominated and equal teams. Pure education teams have less citations than pure industry teams, while the collaboration between education and industry receive the most citations.

Table 2. Multivariable regression for leaders and team impact

	2-year citation percentile			5-year citation percentile		
	(1)	(2)	(3)	(4)	(5)	(6)
Leader Categories						
(0, 2, 0)	0.56***			0.49**		
	(0.16)			(0.16)		
(0, 0, 2)	1.01***			0.76***		
	(0.15)			(0.15)		
(1, 1, 0)	1.63***			1.54***		
	(0.12)			(0.12)		
(0, 1, 1)	1.60***			1.47***		
	(0.13)			(0.13)		
(1, 0, 1)	1.99***			1.93***		

	(0.12)			(0.12)		
Difference of Leaders'		0.10***	0.08***		0.10***	0.09***
Career Age		(0.01)	(0.01)		(0.01)	(0.01)
Team size	1.68***	1.70***	1.66***	1.69***	1.71***	1.61***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Mean Team's Career			0.03**			0.01
Age			(0.01)			(0.01)
Male/Female/Equal						
Female			-1.51***			-1.41***
			(0.15)			(0.15)
Equal			-1.32***			-1.30***
			(0.14)			(0.14)
Education/Industry/						
Combined						
Industry			3.72***			3.37***
			(0.31)			(0.31)
Combined			5.77***			5.97***
			(0.20)			(0.20)
Fixed effect: Year	Yes	Yes	Yes	Yes	Yes	Yes
R	0.013	0.013	0.033	0.014	0.014	0.040
N			521,693			

Note. * P < 0.05; ** P < 0.01; *** P < 0.001; 2-year means 2-year citation, 5-year means 5-year citation

Shared Leadership and Team Scale

In Fig 2a, we divide teams into large teams (more than 3 people) and small teams (less than or equal to 3 people). Large teams receive more attention than small teams. In large teams, heterogenous teams receive more citations than homogenous teams. It suggests that heterogenous leaders with both junior and senior leaders are more adaptive in large teams. In small teams, homogenous teams have a sharp decrease in mean 5-year citation percentile with the increase of leaders' career age. Especially for teams with two senior leaders, they are not suitable to lead small teams. This further proves that senior leaders are more suitable to large teams since they

have visioning, personnel resources and financial resources (Mumford et al., 2007), but not necessarily boost more effective communication and adaptability in small teams (Katzenback & Smith, 1993).

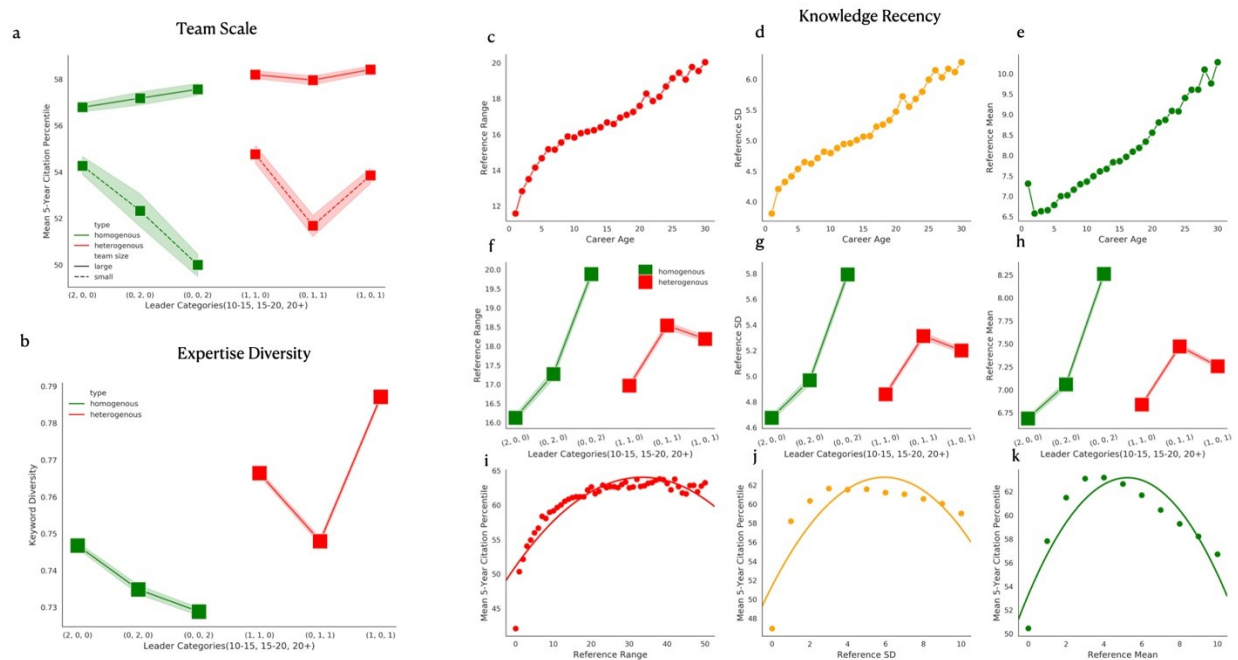


Fig 2. a. Heterogenous leaders in large teams have more citations than homogenous leaders in larger teams. Senior leader combinations in homogenous leaders have less citations. We calculate papers' citation percentile and then aggregate (mean value) in the six different leader categories in different teams within 5 year. Bootstrapped 95% confidence intervals are shown as translucent bands. **b.** Homogenous leaders have higher skill overlap than heterogenous leaders. We calculate the ratio between the same skill sets number and all the skill sets number for two leaders until they publish the focal paper. **c-e.** Researchers who are older in career age have larger (c), reference standard deviation (d) and reference mean (e). Each point in the picture represents the mean values of reference variables grouped by the career age value. **f-h.** Papers with senior leaders have oldest reference, whereas papers with junior leaders have most recent references. We calculate papers' reference range (f), reference standard deviation (g) and reference mean (h) and then aggregate (mean value) in the six different leader categories. Bootstrapped 95% confidence intervals are shown as translucent bands. **i-k.** Curvilinear relationship between reference aging and paper impact. We calculate papers' citation percentile

in the publication year and then aggregate (mean value) based on papers' reference range, standard deviation of mean within 5 years.

Shared Leadership and Expertise Diversity

We use the keywords of each paper to represent authors' expertise. For each leader in each paper, we trace all the keywords of one's publications before publishing the focal one. For two leaders in each paper, we measure the percentage of the skill overlap $\frac{a \cap b}{a \cup b}$, where a and b are the number of skills of two leaders (with duplication, recount the skill number when the same keyword appears in different papers). We conclude that the average keyword overlap percentage of homogenous leaders is larger than that of heterogenous leaders (Fig 2b). It suggests that heterogenous leaders have more diverse expertise and they might be more complementary in knowledge. Our results are similar to the findings in the open source software teams, low overlap in expertise can help teams have active activities and updates (Giuri et al., 2010). Paper publishing teams and software building teams are similar since their tasks both require novelty. Innovative teams benefit more from the expertise diversity from heterogenous leaders than the same expertise from homogenous leaders.

Shared Leadership and Knowledge Recency

Usually, a paper has a list of references, which reflect authors' depth and range in knowledge. We cannot directly differentiate the reference is cited by whom when the publication is finished by a team. Thus, we choose 776,614 sole-author publications. We can find that in Fig 2c-e, senior researchers have larger reference range, reference standard deviation and mean values than junior researchers. The range of references, for example, a paper published in 2020 has two references, the earliest reference is 2010, the newest reference is 2019. Then we calculate the time distance of reference year to the publication year, 10 and 1 separately. Thus, the range of references is 9. Similarly, we calculate the standard deviation and mean value of references. In Fig 2f-h, both senior leaders (0,0,2) have the largest reference range, standard deviation and mean value. It verifies senior leaders can help increase the depth of papers and trace older references. Both junior leaders (2,0,0) have the smallest reference range, standard deviation and mean values. It proves that junior leaders always keep the trace of newest research and thus cite more recent papers. But the results tell us neither too old references nor too recent references will

increase papers' impact. In Fig 2i-k, there is an inverted U-shape between reference aging and paper impact. Papers with medium reference range have higher citation impact than papers with both small or large range. This is the same for reference standard deviation and mean values. This result indicates heterogeneous leaders with both junior and senior leaders follow newest and trace oldest ideas meanwhile have more citations. We find the same patterns that senior researchers cite older references than junior ones (Gingras et al., 2008; Merton, 1973). Prior research suggests that when senior researchers work with junior researchers, reference aging decreases and disruptive novelty increases (Cui et al., 2022). Our results demonstrate the assemble of junior leader and senior leader in shared leadership can also bring high citation influence for teams.

Conclusion

Leadership is evolving from “individualistic, hierarchical, one-directional and de-contextualized notions” to shared leadership (DeRue, 2011, p. 125). With increasing complexity, uncertainty and knowledge-intensity in scientific tasks, a single leader cannot play all leading functions, which requires multiple leaders participate in the leadership activities. Shared leadership is a widely discussed topic in diverse creative tasks, such as R&D teams in companies (Gumusluoglu & Ilsev, 2009), open source community (Zhu et al., 2011), brainstorming and negotiating (Woolley et al., 2010). In this paper, we extend the concept of shared leadership to scientific context. Scientific teams are driven by creativity and knowledge, which might be different from companies driven by profits and routine affairs. Knowledge workers have the professional pursuit and desire for leadership impact (Pearce et al., 2004). Social identity theory proposes that acting as leaders makes them feel important, have a stronger belonging sense to the group and meanwhile boost team innovation (Hogg, 2008). Our research enriches the social identity theory since we apply that into the shared leadership in scientific teams. In the science of science background, we consider shared leadership as a kind of informal leaders, rather than strict definition of formal supervisors. We distinguish different leaders from junior, mid, to senior leaders in important positions, and six different combinations of these leaders. We compare the difference between homogenous leaders and heterogeneous leaders. Our main finding is that heterogeneous leaders outperform homogenous leaders with higher citation ranks. The difference

of two leaders' career age is positively associated with paper's high citations, even when we control time fixed effect and team size. The result that heterogenous leaders are better is still consistent for all students of different career age. We further explore possible mechanisms underlying this pattern. Firstly, heterogenous leaders with both senior and junior leaders are more adaptive in large and small teams. Heterogenous shared leadership can equip teams with both the benefits of senior researcher (e.g., vision, impact) and junior researcher (e.g., details, novelty). Secondly, senior leaders cite older reference whereas junior leaders follow more recent work. The complementary of this literature knowledge can build projects with depth and recency. This can help explain why heterogenous leaders are better than homogenous leaders. Thirdly, senior leaders and junior leaders have diverse expertise, whereas peer leaders have large skill overlap. Heterogenous leaders can bring heterogenous skills and expertise for teams, which is beneficial for team performance.

Our project suggests that the combination of senior leaders and junior leaders can maximize the team performance, in comparison with leaders with similar age. It will produce long-term effect to junior scientists who want to gain leadership in science and students who start to begin the research career. How junior scientists win a seat at the decision-making table is a heated topic in science (Powell, 2021). Although junior scientists are still at the early stage of career, their participation in leadership can bring fresh ideas. Working with senior leaders in the form of shared leadership is a way to help junior scientists feel ownership and make up for the lack of experience. Heterogenous leadership can boost student-centered mentorship. Listening different voices from multiple leaders can help them develop critical thinking and learn fuse diverging ideas (Phillips & Pugh, 2015). We still have some limitations in this paper. We choose leaders whose career age are above 10. The definition of leaders is relatively simple. It is possible that they are still junior researchers though they are above 10, or excellent talents who are young but are actually leaders. Also, the distinguish between junior, mid, and senior based on career age is relatively arbitrary. Career age has limits in defining experience, and power of leaders. These identified leaders are not necessarily formal mentors. Instead, they are informal leaders. The name disambiguation might have difficulty in identifying Asian names, and thus some wrong identifications exists. In the future, we plan to consider the order of leaders in authorship, the ranking of senior leader or mid leader or senior leader. We will also consider the demographic

variables of leaders, such as gender, race, institutions, and the styles of leaders to see how these variables influence the combination of different leaders.

References

- Abrams, D., Randsley de Moura, G., Marques, J. M., & Hutchison, P. (2008). Innovation credit: When can leaders oppose their group's norms?. *Journal of personality and social psychology*, 95(3), 662.
- Acuna, D. E., & Liang, L. (2021). Are AI ethics conferences different and more diverse compared to traditional computer science conferences?.
- Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, 357(6351), 583-587.
- Ali, A., Wang, H., & Johnson, R. E. (2020). Empirical analysis of shared leadership promotion and team creativity: An adaptive leadership perspective. *Journal of organizational behavior*, 41(5), 405-423.
- Azoulay, P., Fons-Rosen, C., & Graff Zivin, J. S. (2019). Does science advance one funeral at a time?. *American Economic Review*, 109(8), 2889-2920.
- Bantel, K. A., & Jackson, S. E. (1989). Top management and innovations in banking: Does the composition of the top team make a difference?. *Strategic management journal*, 10(S1), 107-124.
- Barber, B. (1961). Resistance by Scientists to Scientific Discovery: This source of resistance has yet to be given the scrutiny accorded religious and ideological sources. *Science*, 134(3479), 596-602.
- Barnett, G. A., & Fink, E. L. (2008). Impact of the internet and scholar age distribution on academic citation age. *Journal of the American Society for Information Science and Technology*, 59(4), 526-534.
- Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of applied psychology*, 88(2), 207.

- Bu, Y., Ding, Y., Xu, J., Liang, X., Gao, G., & Zhao, Y. (2018a). Understanding success through the diversity of collaborators and the milestone of career. *Journal of the Association for Information Science and Technology*, 69(1), 87-97.
- Bu, Y., Murray, D. S., Xu, J., Ding, Y., Ai, P., Shen, J., & Yang, F. (2018b). Analyzing scientific collaboration with “giants” based on the milestones of career. *Proceedings of the Association for Information Science and Technology*, 55(1), 29-38.
- Carpenter, M. A., & Fredrickson, J. W. (2001). Top management teams, global strategic posture, and the moderating role of uncertainty. *Academy of Management journal*, 44(3), 533-545.
- Chinchilla-Rodríguez, Z., Sugimoto, C. R., & Larivière, V. (2019). Follow the leader: On the relationship between leadership and scholarly impact in international collaborations. *PloS one*, 14(6), e0218309.
- Costas, R., & Bordons, M. (2011). Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective. *Scientometrics*, 88(1), 145-161.
- Cui, H., Wu, L., & Evans, J. A. (2022). Aging Scientists and Slowed Advance. *arXiv preprint arXiv:2202.04044*.
- DeRue, D. S. (2011). Adaptive leadership theory: Leading and following as a complex adaptive process. *Research in organizational behavior*, 31, 125-150.
- Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of informetrics*, 5(1), 187-203.
- Eisenhardt, K. M., & Bourgeois III, L. J. (1988). Politics of strategic decision making in high-velocity environments: Toward a midrange theory. *Academy of management journal*, 31(4), 737-770.
- Fernandes, J. M., Costa, A., & Cortez, P. (2022). Author placement in Computer Science: a study based on the careers of ACM Fellows. *Scientometrics*, 127(1), 351-368.
- Firth, A., & Martens, E. (2008). Transforming supervisors? A critique of post-liberal approaches to research supervision. *Teaching in Higher Education*, 13(3), 279-289.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... & Barabási, A. L. (2018). Science of science. *Science*, 359(6379), eaao0185.

- Fuyuno, I., & Cyranoski, D. (2006). Cash for papers: putting a premium on publication. *Nature*, 441(7095), 792-793.
- Gingras, Y., Lariviere, V., Macaluso, B., & Robitaille, J. P. (2008). The effects of aging on researchers' publication and citation patterns. *PloS one*, 3(12), e4048.
- Giuri, P., Ploner, M., Rullani, F., & Torrasi, S. (2010). Skills, division of labor and performance in collective inventions: Evidence from open source software. *International Journal of Industrial Organization*, 28(1), 54-68.
- Giuri, P., Ploner, M., Rullani, F., & Torrasi, S. (2010). Skills, division of labor and performance in collective inventions: Evidence from open source software. *International Journal of Industrial Organization*, 28(1), 54-68.
- Gumusluoglu, L., & Ilsev, A. (2009). Transformational leadership, creativity, and organizational innovation. *Journal of business research*, 62(4), 461-473.
- Haeussler, C., & Sauermann, H. (2020). Division of labor in collaborative knowledge production: The role of team size and interdisciplinarity. *Research Policy*, 49(6), 103987.
- Hambrick, D. C., Cho, T. S., & Chen, M. J. (1996). The influence of top management team heterogeneity on firms' competitive moves. *Administrative science quarterly*, 659-684.
- Haythornthwaite, C. (2006). Learning and knowledge networks in interdisciplinary collaborations. *Journal of the American society for information science and technology*, 57(8), 1079-1092.
- Helgesson, G., & Eriksson, S. (2019). Authorship order. *Learned Publishing*, 32(2), 106-112.
- Hirst, G., Van Dick, R., & Van Knippenberg, D. (2009). A social identity perspective on leadership and employee creativity. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 30(7), 963-982.
- Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, 117(17), 9284-9291.
- Hogg, M. A. (2008). Social identity theory of leadership. *Leadership at the crossroads*, 1, 62-77.
- Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological review*, 65(2), 117.

- Hsiehchen, D., Espinoza, M., & Hsieh, A. (2015). Multinational teams and diseconomies of scale in collaborative research. *Science advances*, 1(8), e1500211.
- Hu, X. (2009). Loads of special authorship functions: Linear growth in the percentage of “equal first authors” and corresponding authors. *Journal of the American Society for Information Science and Technology*, 60(11), 2378-2381.
- Hunter, L., & Leahey, E. (2008). Collaborative research in sociology: Trends and contributing factors. *The American Sociologist*, 39(4), 290-306.
- Katzenback, J. R. & Smith, D.K. (1993). The discipline of teams. *Harvard Business Review*, March/April.
- Kraut, R. E., Egido, C., & Galegher, J. (2014). Patterns of contact and communication in scientific research collaborations. *Intellectual teamwork*, 163-186.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (Vol. 111). University of Chicago Press: Chicago.
- Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., & Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social studies of science*, 46(3), 417-435.
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323-1332.
- Lee, K., Brownstein, J. S., Mills, R. G., & Kohane, I. S. (2010). Does collocation inform the impact of collaboration?. *PloS one*, 5(12), e14279.
- Liu, S., Hu, J., Li, Y., Wang, Z., & Lin, X. (2014). Examining the cross-level relationship between shared leadership and learning in teams: Evidence from China. *The leadership quarterly*, 25(2), 282-295.
- Liu, X. Z., & Fang, H. (2014). Scientific group leaders' authorship preferences: An empirical investigation. *Scientometrics*, 98(2), 909-925.
- Manjunath, A., Li, H., Song, S., Zhang, Z., & Kumar, I. (2021). Comprehensive analysis of 2.4 million patent-to-research citations maps the biomedical innovation and translation landscape. *Nature Biotechnology*, 39(6), 678–683.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

- Meyer, E., & Schroeder R. (2015). Knowledge machines: Digital transformations of the sciences and humanities. MIT Press, Cambridge, Massachusetts.
- Miles, J. R., & Kivlighan Jr, D. M. (2010). Co-leader similarity and group climate in group interventions: Testing the co-leadership, team cognition-team diversity model. *Group dynamics: Theory, research, and practice*, 14(2), 114.
- Milojević, S. (2012). How are academic age, productivity and collaboration related to citing behavior of researchers?. *PloS one*, 7(11), e49176.
- Milojević, S., Radicchi, F., & Walsh, J. P. (2018). Changing demographics of scientific careers: The rise of the temporary workforce. *Proceedings of the National Academy of Sciences*, 115(50), 12616-12623.
- Mumford, T. V., Campion, M. A., & Morgeson, F. P. (2007). The leadership skills strataplex: Leadership skill requirements across organizational levels. *The Leadership Quarterly*, 18(2), 154-166.
- Nielsen, M. W., Bloch, C. W., & Schiebinger, L. (2018). Making gender diversity work for scientific discovery and innovation. *Nature human behaviour*, 2(10), 726-734.
- O'mahony, S., & Ferraro, F. (2007). The emergence of governance in an open source community. *Academy of Management Journal*, 50(5), 1079-1106.
- O'Toole, J., Galbraith, J., & Lawler III, E. E. (2002). When two (or more) heads are better than one: The promise and pitfalls of shared leadership. *California management review*, 44(4), 65-83.
- Paulus, P. B., Kohn, N. W., Arditti, L. E., & Korde, R. M. (2013). Understanding the group size effect in electronic brainstorming. *Small Group Research*, 44(3), 332-352.
- Pearce, C. L. (2004). The future of leadership: Combining vertical and shared leadership to transform knowledge work. *Academy of Management Perspectives*, 18(1), 47-57.
- Pearce, C. L., & Conger, J. A. (2002). *Shared leadership: Reframing the hows and whys of leadership*. Sage Publications.
- Phillips, E., & Pugh, D. (2015). *EBOOK: How to Get a PhD: A Handbook for Students and their Supervisors*. McGraw-Hill Education (UK).
- Powell, K. (2021). How junior scientists can land a seat at the leadership table. *Nature*, 592(7854), 475-477.

- Sie, R. L., Drachsler, H., Bitter-Rijkema, M., & Sloep, P. (2012). To whom and why should I connect? Co-author recommendation based on powerful and similar peers. *International Journal of Technology Enhanced Learning*, 4(1-2), 121-137.
- Sperber, S., & Linder, C. (2018). The impact of top management teams on firm innovativeness: a configurational analysis of demographic characteristics, leadership style and team power distribution. *Review of Managerial Science*, 12(1), 285-316.
- Staats, B. R., Milkman, K. L., & Fox, C. R. (2012). The team scaling fallacy: Underestimating the declining efficiency of larger teams. *Organizational Behavior and Human Decision Processes*, 118(2), 132-142.
- Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational identity: A reader*, 56(65), 9780203505984-16.
- Tang, J., Fong, A. C., Wang, B., & Zhang, J. (2011). A unified probabilistic framework for name disambiguation in digital library. *IEEE Transactions on Knowledge and Data Engineering*, 24(6), 975-987.
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157), 468-472.
- Van Knippenberg, D. (2011). Embodying who we are: Leader group prototypicality and leadership effectiveness. *The leadership quarterly*, 22(6), 1078-1091.
- Voss, B. L. (1999). Co-Leaders: The Power of Great Partnerships. *Journal of Business Strategy*, 20(5), 45-45.
- Wagner, C. S., Whetsell, T. A., & Mukherjee, S. (2019). International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Research Policy*, 48(5), 1260-1270.
- Watts, J. H. (2010). Team supervision of the doctorate: Managing roles, relationships and contradictions. *Teaching in Higher Education*, 15(3), 335-339.
- Weiss, C. H., Cambone, J., & Wyeth, A. (1992). Trouble in paradise: Teacher conflicts in shared decision making. *Educational Administration Quarterly*, 28(3), 350-367.
- Whitley, R. (2000). *The Intellectual and Social Organization of the Sciences* (2nd ed.). Oxford: Oxford University Press.

- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *science*, 330(6004), 686-688.
- Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378-382.
- Xu, F., Wu, L., & Evans, J. (2022a). Flat teams drive scientific innovation. *Proceedings of the National Academy of Sciences*, 119(23), e2200927119.
- Xu, H., Bu, Y., Liu, M., Zhang, C., Sun, M., Zhang, Y., ... & Ding, Y. (2022b). Team power dynamics and team impact: New perspectives on scientific collaboration using career age as a proxy for team power. *Journal of the Association for Information Science and Technology*.
- Yin, Y., Dong, Y., Wang, K., Wang, D., & Jones, B. F. (2022). Public use and public funding of science. *Nature Human Behaviour*, 1-7.
- Zeng, A., Fan, Y., Di, Z., Wang, Y., & Havlin, S. (2021). Fresh teams are associated with original and multidisciplinary research. *Nature Human Behaviour*, 5(10), 1314-1322.
- Zhang, C., Bu, Y., & Ding, Y. (2016). Understanding scientific collaboration from the perspective of collaborators and their network structures. *ICConference 2016 Proceedings*.
- Zhu, H., Kraut, R. E., Wang, Y. C., & Kittur, A. (2011). Identifying shared leadership in Wikipedia. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3431-3434).
- Zhu, J., Liao, Z., Yam, K. C., & Johnson, R. E. (2018). Shared leadership: A state-of-the-art review and future research agenda. *Journal of Organizational Behavior*, 39(7), 834-852.
- Zucker, L. G., & Darby, M. R. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences*, 93(23), 12709-12716.