Scientometrics (2018) 114:463–479
https://doi.org/10.1007/s11192-017-2599-0

Measuring the stability of scientific collaboration

Yi Bu[1] · Dakota S. Murray[1] · Ying Ding[1,2] · Yong Huang[2] · Yiming Zhao[3]

Received: 16 October 2017 / Published online: 2 December 2017

Abstract Stability has long been regarded as an important characteristic of many natural and social processes. In regards to scientific collaborations, we define stability to reflect the consistent investment of a certain amount of effort into a relationship. In this paper, we provide an explicit definition of a new indicator of stability, based on the year-to-year publication output of collaborations. We conduct a large-scale analysis of stability among collaborations between authors publishing in the field of computer science. Collaborations with medium–high degree of stability tend to occur most frequently, and on average, have the highest average scientific impact. We explore other "circumstances", reflecting the composition of collaborators, that may interact with the relationship between stability and impact, and show that (1) Transdisciplinary collaborations with low stability leads to high impact publications; (2) Stable collaboration with the collaborative author pairs showing greater difference in scientific age or career impact can produce high impact publications; and (3) Highly-cited collaborators whose publications have a large number of co-authors do not keep stable collaborations. We also demonstrate how our indicator for stability can be used alongside other similar indicators, such as persistence, to better understand the nature of scientific collaboration, and outline a new taxonomy of collaborations.

Keywords Scientific collaboration · Stability · Persistence · Scientometrics

Mathematics Subject Classification 62P25

JEL Classification O35

✉ Yi Bu
buyi@iu.edu

[1] School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA

[2] School of Information Management, Wuhan University, Wuhan, Hubei, China

[3] Center for Studies of Information Resources, Wuhan University, Wuhan, Hubei, China

Introduction

Cultural divides, geographic distance, personal differences, linguistic barriers, ephemeral sources of funding, and a constant stream of tasks demanding and diverting attention—these are only a few of the forces that conspire to strain and end a collaborative relationship between researchers. When assessing the benefits of their collaborations, a researcher might weigh these potential stresses against the benefits of collaboration. While there are conflicting evidences that collaboration has a positive impact on research productivity (Lee and Bozeman 2005; Abramo et al. 2009), collaboration and team science is positively correlated with scientific impact (Wuchty et al. 2007; Larivière et al. 2015), and can be used as a strategy to obtain resources necessary for some research (Wray 2002). Multidisciplinary teams of scientists might also have a greater capacity to tackle complex scientific problems that form the bulk of contemporary science (Pennington et al. 2013; Uzzi et al. 2013). Some of these benefits compound as the relationship is maintained and strengthened over a long period of time (Petersen 2015). However, if a researcher decides to maintain a relationship, he then must decide how much effort to invest into the collaboration. A researcher must consider the effects of different degrees of *stability* of their collaborations.

Stability is a concept embedded within many fields of science, each with different interpretations. In physics, the law of *inertia* is a concept that explains the stability of an object's motion. Moving away from natural laws, researchers have also described stability as it relates to linear flow (Ellingsen and Palm 1975), classical solutions (Bogomol'Nyi 1976), and gravitating systems (Fridman and Polyachenko 2012). In biological sciences, the term *homeostasis*, coined by Cannon (1932), reflects the stability that an organism achieves through self-regulation of bodily processes. The absence of homeostasis, the basic processes of life might cease to function. In ecological sciences, stability takes the form of the longstanding concept of the "balance of nature" and "natural equilibrium" whereby species tend towards a balanced or *stable* population size (Simberloff 2014). In sociology, the term "structural inertia" is used to describe the temporal stability in the correspondence between a class of organizations and their environments (Hannan and Freeman 1984).

These various terms and conceptualizations are all close counterparts to stability, but they each describe an inherent characteristic of a natural system, something to be understood, not necessarily to be controlled and manipulated. However, in the study of human organizations and activities, stability can instead be viewed as something to control and optimize in order to produce a given outcome. In the study of business, for example, there is conflicting research that the most stable teams produce the best outcomes for businesses and organizations (Eitzen and Yetman 1972; Smith and Nyman 1939; Nystrom and Starbuck 1984). The effects of stability on success can vary by the types and goals of the teams being studied. Grusky (1963) found that unstable baseball teams, characterized by their high turnover, are less effective, while in product development instability can sometimes be an advantage (Akgün and Lynn 2002). Indeed, when it comes to creative endeavors such as product development or scientific research, the benefits of stability might not be so certain.

As science becomes an activity increasingly dominated by teams (Wuchty et al. 2007), it becomes important to understand the nuances of team stability. Unfortunately, there are few scientometric studies examining the role of collaborative stability to scientific research. At the level of the individual, Ioannidis (2011) highlighted how maintaining a stable flow of publication was important to career success. Ioannidis et al. (2014) extended this concept by introducing a measure of uninterrupted and continuous presence (UCP).

Authors with the greatest degree of UCP were also found to be “top authors”, receiving a disproportionate number of citations (Ioannidis et al. 2014; Wu et al. 2016). Only recently have researchers applied these concepts to the study of stability in collaboration. Petersen (2015) categorized collaborative relationships within physics into three categories, “weak ties”, “strong ties”, and “super ties” based on their long-term co-authorship. Authors who were part of “strong tie” relationships had greater than average productivity and impact, indicating that maintaining stable ties might be universally beneficial. Bu et al. (2017) drew from these to define a measure of *persistence,* computed using the number of skip years without collaboration and the number of time intervals without collaboration. Contrary to Petersen (2015), they found a more nuanced relationship: long-term collaborative relationships are only beneficial to success up to a point, and the most persistent collaborations tend not to produce the highest impact. There are many ways to measure stability in science, depending on the data, the level of analysis, and the purpose of a study, making it difficult to compare and contextualize results.

Definitions of stability vary between scientific disciplines, and even within a single discipline, such as scientometrics. How one defines stability is inherent to the purpose of its use. To some, stability is a characteristic of a natural system, an aspect to be studied and understood. To others, stability is a controllable factor of an activity, something to be optimized in order to produce the best outcomes. Sociologists might be interested in the stability of scientific collaborations as a characteristic of the structure of science, or as evidence of underlying social processes of research. Science policy makers on the other hand, are interested in strategies to increase and improve collaboration—and by extension scientific quality—by creating policies that incentivize optimum stability. However, with so many related yet distinct measures and concepts it can be difficult for either sociologists or policy makers to make sense of stability and collaboration. There is a present need for clarity in the discussions of stability and collaboration.

In this paper, we aim to draw from previous notions of stability to mathematically define a new indicator of collaborative stability, distinct from other measures. Contrary to our past work that measured *persistence* (Bu et al. 2017), this new measure of stability is sensitive to the changes of the number of collaborations of an author pair over time. We then use this measure to conduct a large-scale quantitative analysis of authors publishing in the field of computer science, and uncover the relationships between collaborative stability and scientific impact. We also provide an example analysis of how such relationships might change based on the diversity of collaborative authors’ scientific ages, scientific impacts, disciplinary topics, and research team sizes. We then produce a new framework that takes into account both *stability* and *persistence* to better understand scientific collaborations.

Methodology

Data

The dataset used in this study is derived from ArnetMiner (Tang et al. 2008a) academic social networking platform. ArnetMiner contains, among other things, information for 2,092,356 papers published between 1936 and 2014 in the field of computer science, 8,024,869 local citation relationships between these publications, and 1,207,061 distinct authors. The authors’ names are disambiguated by the algorithm proposed by Tang et al. (2012), who have shown that this process has satisfactory precision and recall. We select years 2001–2010 as the time period for our analyses, which provide a final dataset of 885,562 unique authors, 449,875 articles, and 606,843 unique local citation relationships.

Methods

We aim to analyze the relationship between the scientific impact of collaborations (success) and the stability of collaboration. We adopt the same framework proposed by Bu et al. (2017), which includes not only the relationship between stability and collaborative success, but also the effect of different *circumstances*, mediating factors that might interact with stability. To measure the scientific performance of a pair of co-authors), we use its yearly average number of citations (YANC) and further consider four circumstances between these two co-authors: degree of transdisciplinarity, diversity of scientific impact and scientific age, and research team size (i.e., the number of their co-authors).

For the purposes of this paper, co-authorship of a publication is used as an indicator of collaboration (Milojević 2010). If two authors co-author a publication, then they are considered to be a *collaborative pair*. We only consider relationships between two authors, thus in the case of more than two authors on a publication, every combination of authors is a distinct collaborative pair. For example, given a publication with three co-authors, (a_1, a_2, a_3), we consider three distinct collaborative pairs: (a_1, a_2), (a_2, a_3), and (a_1, a_3). There are a total of 1,419,990 collaborative pairs in our dataset (in years 2001–2010), and each author belongs to an average of 2.98 collaborative pairs.

The scientific success of a collaborative pair is measured by the average scientific success of all publications co-authored by the pair: the yearly average number of citations (YANC). The YANC of a publication is the number of citations it has received divided by the number of years in a time span. Papers' citation counts are calculated at the end of 2014. For example, given a paper published at the end of 2009 that has accumulated 10 citations in 2010, 12 in 2011, 6 in 2012, 6 in 2013, and 4 in 2014, the YANC will be $\frac{10+12+6+6+4}{2014-2009} = 7.6$. Given a collaborative pair with three co-authored publications, with a YANC of 5, 15, and 12, respectively, the YANC of the collaboration is calculated as $\frac{5+15+12}{3} \approx 10.6667$.

We calculate another indicator for the productivity of collaborative pairs: the yearly average number of publications (YANP). This is simply calculated as the number of publications co-authored by a given collaborative pair, divided by the total number of years of a given time span.

Definition of the stability of scientific collaboration

To define collaborative stability, we begin from intuition: If two authors collaborate frequently each year (Fig. 1a), then they have a stable collaboration relationship. Alternatively, if they maintain a low quantity, but constant, string of publications each year (or even no collaboration in certain years), then their collaboration remains stable, such as in Fig. 1b. However, Fig. 1c, d exhibit two examples of low stability of collaboration, where the number of collaborations between two authors changes abruptly.

Inspired by these examples, we provide an explicit mathematical definition of collaborative stability. Within an arbitrary number of consecutive years, N, where $(N > 2)$ and each year is designated as $y_1, y_2, \ldots, y_N$, suppose that two authors a_m and a_n have co-authored a certain count of publications each year, $c_{mn}^1, c_{mn}^2, \ldots, c_{mn}^N$ ($c_{mn}^1, c_{mn}^2, \ldots, c_{mn}^N \geq 0$, $\sum_{i=1}^{N} c_{mn}^i \neq 0$, $i = 1, 2, \ldots, N$), respectively. The stability of collaboration between authors

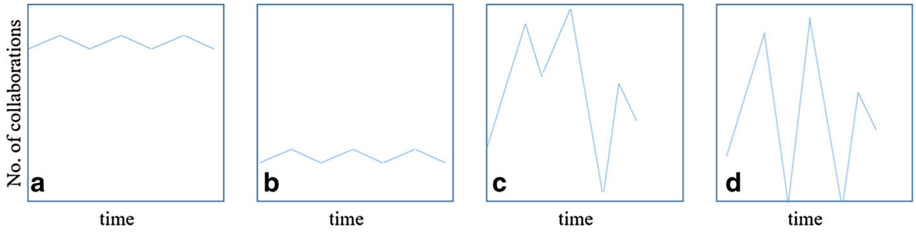

Fig. 1 Diagram of stability. Horizontal axes represent time while vertical axes show the number of collaborations between two authors. Sub-figures **a**, **b** exhibit collaborations with high stability, while **c**, **d** show collaboration of low stability

a_m and a_n, S_{mn_i}, is thus represented using the below equation, where $\max(c_{mn})$ refers to the maximum number of publications co-authored in a year: c_{mn}^1, c_{mn}^2, …, c_{mn}^N.

$$S_{mn\_i} = 1 - \frac{\sum_{i=1}^{N-1} \left| c_{mn}^{i+1} - c_{mn}^{i} \right|}{(N-1) \cdot [\max(c_{mn}) + 1]}$$

From the above formula, we can find that the values of stability are in the range $\left[\frac{1}{\max(c_{mn})+1}, 1\right]$. When a collaborative pair co-authors a consistent number of publications every year, then their measure of stability equals to one. If, however, they alternate between publishing $\max(c_{mn})$ times 1 year, and nothing the next year, then their measurement of stability equals to $\frac{1}{\max(c_{mn})+1}$. Note that due to the categorical characteristics of collaboration count series and the law of limitation,[1] the distribution of stability here shows sparsity within the range of $\left[1 - \frac{1}{N}, 1\right)$.

Consider several author pairs, (a_1, a_2), (a_3, a_4), (a_5, a_6), and (a_7, a_8) with yearly publication counts detailed in Table 1. Based on our definition of stability, the stability measure of (a_1, a_2) is one because there is no change in publication count; the stability of author pair (a_3, a_4) equals to 0.3 $\left(= 1 - \frac{|0-3|+|4-0|+|0-4|+|3-0|}{4\times(4+1)}\right)$. Moreover, we see that, while (a_5, a_6) and (a_7, a_8) have similar fluctuations $(5 - 4 = 10 - 9 = 1)$, (a_5, a_6) have higher stability because the relative fluctuation between ten and nine publications $(\frac{1}{10})$ is less than the fluctuation between five and four publications $(\frac{1}{5})$ for (a_7, a_8).

Measure of circumstances

Here we consider four *circumstances*, a term we use to describe variables that we believe might interact with collaborative stability, which are similar to those outlined in Bu et al. (2017) including transdisciplinarity among authors, scientific career impact, scientific ages of authors, and average research team size.

The first of these measures, the degree of transdisciplinarity (DoT) of a collaborative pair, can be loosely defined as the differences in research topics pursued by each author. In our dataset, we analyze only authors who publish in the field of computer science, so we measure differences in sub-fields (e.g., artificial intelligence vs. operating systems) rather than high-level disciplinary differences (biology vs. computer science). We use the Author-Conference-Topic (ACT) model (Tang et al. 2008b) to generate a 50-dimension vector

[1] When $\max(p_{mn}) \to +\infty$, $S_{mn\_i} \to \left(1 - \frac{1}{N}\right)$.

Table 1 Example of four author pairs' publication counts and stability in 2006–2010

Author pair	2006	2007	2008	2009	2010	Stability
(a_1, a_2)	2	2	2	2	2	1.0000
(a_3, a_4)	3	0	4	0	3	0.3000
(a_5, a_6)	10	9	10	9	10	0.9091
(a_7, a_8)	5	4	5	4	5	0.8000

representing an author's research topic [see further details in Wu et al. (2017)]. Each component of this vector represents a sub-topic of computer science and the probability that author belongs to that sub-topic. The similarity between two author's research topics is calculated as the cosine similarity of their topic vectors. The more similar two author's topic vectors, the lower the DoT.

The scientific career impact of a single author is his/her h-index (Hirsch 2005) calculated using his/her number of publications and citation counts that appear in the ArnetMiner dataset as of 2014. Then, for each collaborative pair we calculate the difference between the two authors' h-indices, and normalize by the maximum difference observed in the entire dataset. We adopt a similar process to calculate the diversity in scientific age between authors in a collaborative author pair, except using the difference between the years of publication of each author's first publication. We again normalize this value by the maximum difference in scientific ages observed in the dataset. Research team size is calculated as the total number of co-authors (not unique) who appear in publications produced by a collaborative pair (including the two authors in the pair), divided by that pair's total number of co-authored publications between 2001 and 2010.

Results

Stability and impact

Past research has found evidences of the so-called "apostle effect", or the benefit to scientific impact and productivity gained from long-term collaborations with "strong" or "super" social ties (Petersen 2015). Our definition of stability is different from that of strong and weak ties, but we can still search for evidence of a similar "apostle effect" as it relates to stability.

Observation of the frequency distribution of the stability of all collaborative pairs concludes that the distribution has a strong negative skew (Fig. 2), with the most common measure of stability ranging between 0.6 and 0.8, which we call middle-high stability. We also name the collaboration pairs with stability 0.0–0.2 as low stable pairs, 0.2–0.4 as middle-low stable pairs, 0.4–0.6 as medium stable pairs, and 0.8–1.0 as high stable pairs.

Figure 2 shows the relationship between the stability and YANC, and between the stability and YANP of all collaborative pairs, grouped into five bins based on their stability measures. Generally, the more stable collaboration, the greater their YANC and YANP. However, those collaborators with a medium–high stability (0.6–0.8) have the greatest YANC, an average of almost two and a half citations greater than the least stable group, and almost two citations greater than the average of the most stable group. YANP is also the highest for the medium–high stable pairs, but quickly decreases in the most stable group.

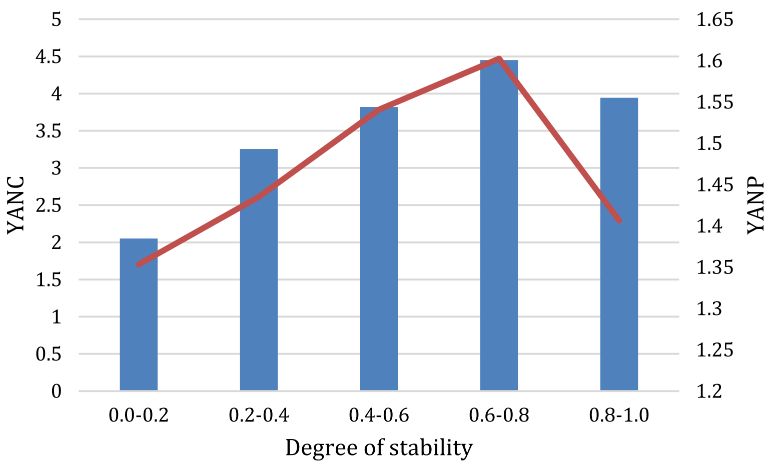

Fig. 2 The relationship between the stability of collaborations and their success, measured as YANC, the yearly average number of citations per article, and YANP, the yearly average number of publications

Stability and transdisciplinarity, research team size, and difference in the scientific age and impact of collaborators

In the previous work we have explored the relationship between persistence, a measure that is similar yet distinct from our current measure of stability, and a set of indicators related to the composition of a collaborative pair (Bu et al. 2017). We call these composition indicators "circumstances", which include a measure of transdisciplinarity, research team size, the difference in the scientific ages of authors, and the difference in the scientific career "impacts" of authors in a collaborative pair. Here we aim to understand how these circumstances relate to stability, and to uncover any patterns of collaboration within our dataset.

Figure 3a shows the values of stability for five bins of research topic similarity , where the closer the research topic similarity is to one, the less transdisciplinarity the collaborative pair. There is little difference between the median of stability of the most and least transdisciplinary pairs. Those collaborative pairs with the lowest research topic similarity have a wider inter-quartile range, suggesting wide diversity in stability among these collaborative pairs. The lowest and second-lowest research topic similarity however has slightly higher median topic similarity than others . Similarly, Fig. 3b–d show the relation between stability and scientific impact/age difference as well as team size. However, all of these differences are not obvious.

Scientific stability, YANC, and the effect of circumstances

In addition to understanding the structure of collaborative stability and each circumstance, it is of interest to policy makers to understand how these circumstances might interact with stability to affect the scientific impact of their publications. In Fig. 4 we show four heat maps, each with the measurements (i.e., one out of four circumstances) and stability grouped into five categories, and each cell represents the YANC of collaborative pairs with the combination of stability/circumstance measurement categories.

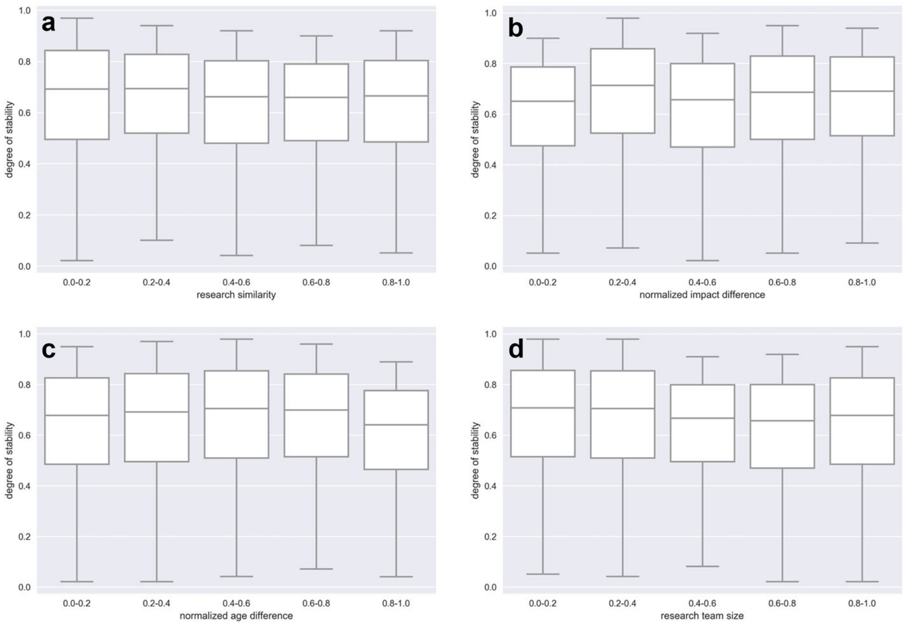

Fig. 3 Boxplots of the values of circumstances for all collaborative pairs within each category of stability. Each shows the relationship between stability with **a** transdisciplinarity, **b** difference in authors' scientific impact, **c** difference in authors' scientific age, and **d** average research team size, respectively

Figure 4a shows the relationship between stability, transdisciplinarity, and research impact. We see that papers with the highest impact are written by collaborative pairs sharing highly similar topics and maintaining high and medium–high stability (0.8–1.0 and 0.6–0.8, respectively). However, highly transdisciplinary collaborations tend to accumulate the more citations when they have low stability (0.0–0.2). Figure 4b plots the average YANC of collaborations along five categories of stability and five categories of difference between collaborator's h-index. The greater the h-indices difference between authors, the closer the measurement is to one. Those collaborative pairs with nearly the same h-indices and high stability have more YANC, while pairs with substantially different h-indices have the greatest YANC when they have either low, medium–high, or high stability, but not medium stability. The YANC of collaborative pairs is lowest for similar h-index authors with the medium stability. Figure 4c shows trends in the diversity in scientific age of collaborators that are similar to those for the difference between collaborator's h-index. Figure 4d shows the average research team size of a collaborative pair, and how this relates to the stability and YANC. Low stable teams are found to have lower average YANC than medium–high stable teams. The highest-impact publications tend to be written by the medium–high stable and large-size teams.

Stability and persistence

Our definition of collaborative stability is one of many similar yet distinct definitions, such as universal continuous presence (Ioannidis et al. 2014), weak/strong/super ties (Petersen 2015), and collaborative persistence (Bu et al. 2017). Here we show how our new indicator

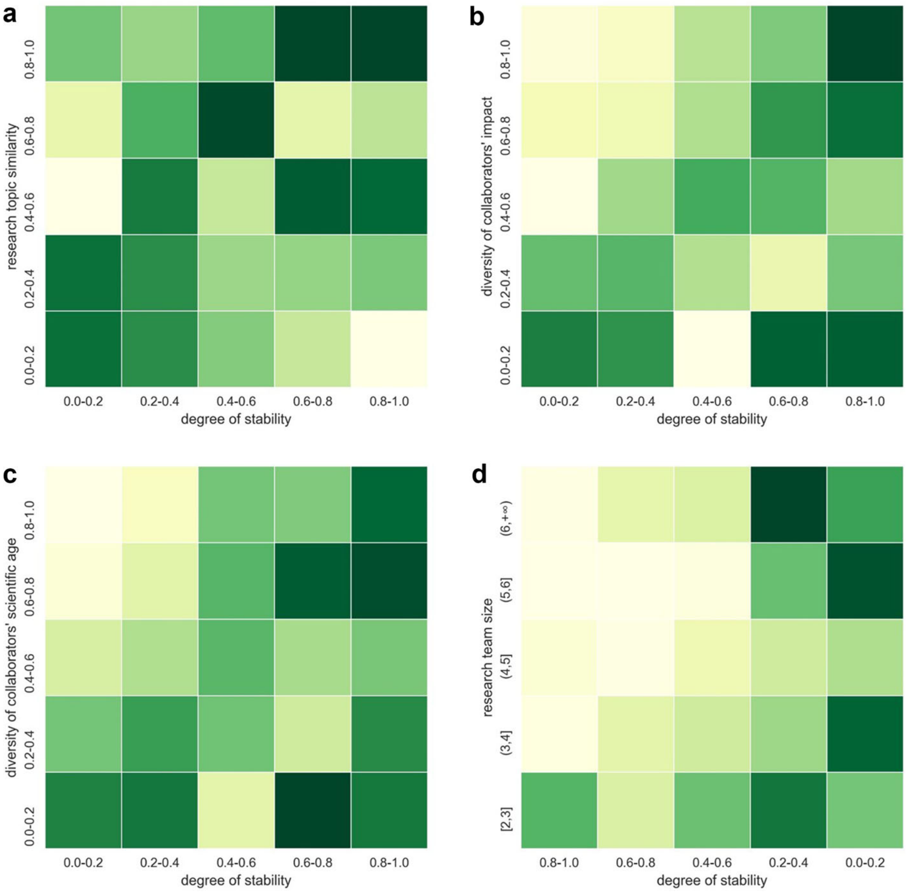

Fig. 4 The relationship between collaboration stability and success in different "circumstances": **a** degree of transdisciplinarity (research topic similarity); **b** diversity of collaborators' impact; **c** diversity of collaborators' scientific age; and **d** research team size. The darkness is proportional to the YANC of the collaboration pairs with the corresponding degree of stability and the value of "circumstance". In **a**, the more similar two collaborators' research topic is, the less their degree of transdisciplinarity is. (Color figure online)

of collaborative stability can be used in conjunction with other indicators to better understand scientific collaboration, specifically as it relates to the scientific impact of publications. To demonstrate this, we use persistence, which was originally outlined in our previous work (Bu et al. 2017).

Both persistence and stability measure the characteristics of collaboration counts of a given co-author pair. Persistence, on the one hand, assesses the continuity of collaborations by looking at the distribution of the number of years with and without co-authored publications in a given time span. Specifically, persistence not only considers how many "non-collaboration" years a collaboration pair has, but also calculates whether these "non-collaboration" years are adjacent or not; stability emphasizes the changes of the number of collaborations between two given years, which are not captured by persistence. Meanwhile, persistence regards the series of collaboration counts as binary (simply zero and

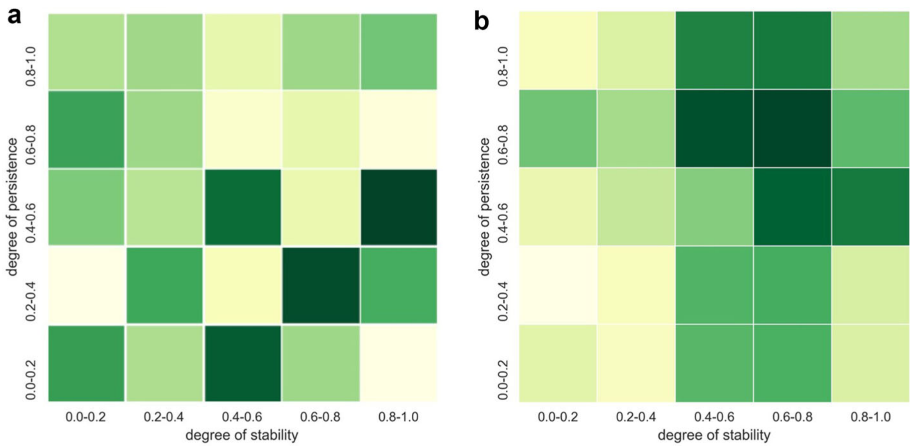

Fig. 5 The relationship between persistence and stability, where the color shows: **a** the frequency of collaborative pairs normalized by the total frequency within each stability bin, and **b** the average YANC of collaborative pairs. Here the degree of persistence has been normalized into [0, 1] for better comparisons. (Color figure online)

one), but stability does not. In practice, two co-authors with high stable collaboration can be either persistent or not, and vice versa (see details in "Discussion" section). Using stability and persistence together can allow us to better make sense of scientific collaboration, as they show two distinct dimensions of collaboration.

We see in Fig. 5a that the most stable collaborative pairs are likely to have medium degree of persistence (0.4–0.6), while never have low persistence (0.0–0.2). Medium–high stability groups, the most common, are likely to have medium–low persistence (0.2–0.4). The least stable groups have a wider variance of likely measures of persistence, but are most likely to also have low (0.0–0.2) or medium–high persistence (0.6–0.8). Examining Fig. 5b, we can see how these structures might affect the scientific impact of publications. Low and medium–low persistence groups tend to have lower impact than high and medium–high persistence groups, and low stability groups tend to have less impact than high persistence groups. The medium and medium–high stable groups tend to have a greatest value of YANC, particularly those that have maintained a medium and medium–high level of persistence.

Discussion

Stability and impact

Stability can offer insights into what types of collaborations are "optimal", and which tends to result in the highest scientific impact. Petersen (2015) is one of the few papers that has examined the effect of "strong", "weak", and "super" ties, a somewhat similar notion to our definition of stability. Most ties between researchers were weak, but those stronger ties constitute substantial investments in social capital that can have long-term implications on a scientist's career. The "apostle effect" is used to describe the benefits accrued to researchers who invested in strong and super ties, an effect that manifests in increased productivity and scientific impact (Petersen 2015). Where our notion of stability most differs from Petersen's is that stability represents the degree of the long-term and consistent investment into a collaborative relationship. We find that the relationship between consistent *investment* and the benefit to scientific impact and productivity is not strictly linear—the most stable relationships have lower average scientific impact and productivity than those high or medium–high stable researchers.

Medium–high stable researchers can be characterized as those who maintain a degree of stability that is less than high stability, but more than those with medium stability. In our dataset medium–high and high stable collaborative pairs are the most common, and similarly to the medium–high and high degrees of persistent collaborations analyzed in our previous work (Bu et al. 2017), Medium–high stability collaborations have the highest average scientific impact. Collaborative relationships that maintain a high level of stability may see a performance benefit resulting from access to shared resources and knowledge (Finholt 1999; Kling and McKim 2000), close and familiar relationships (Hara et al. 2003), and greater team cohesion (Guzzo and Dickson 1996; Kerr and Tindale 2004; Stokols et al. 2008a). But totally stable teams, those that invest a consistent degree of effort each year, may suffer negative returns resulting from high embeddedness (Uzzi 1996), increasing isolation over time (Katz 1982), and diminishing returns from interpersonal trust and relationship intensity (Molina-Morales and Martínez-Fernández 2009). Investment is important, but so is novelty.

The circumstances of stability

The incidence and success of different degrees of collaborative stability is influenced by a host of factors ranging from team demographics to the personality and interpersonal relationships between authors. Here we can only comment on the role of the composition of collaborators along four dimensions: research topic, scientific age, scientific career impact (h-index), and average team size.

Collaborating across disciplinary boundaries brings ample risks and rewards that decide to make the jump. For example, we find that unstable transdisciplinarity teams tend to produce highly impactful research. There are often communication barriers in transdisciplinary research (Emmons et al. 2008), as well as more time consumed by administration and communication (Schaltegger et al. 2013), and more difficulty in acquiring funding (Bromham et al. 2016), all conspiring to end such relationships. We see that those collaborative pairs with the dissimilar topics tend to have greater values of stability, despite the difficulties of maintaining transdisciplinary relationships. Many of these collaborators may have published only one or two papers in a single year, resulting in a high measurement of stability in spite of their ephemeral collaboration. Additionally, our data only includes authors who publish within the field of computer science, and so it is possible that this narrow disciplinary scope excludes more diverse relationships.

Among the most stable collaborators, the most impactful work tends to come from collaborations with the most similar research topics. These collaborations may not represent a trend between research topic and stability, but instead represent the few highly stable and long-term collaborators who together co-author publications every year. Under our model of transdisciplinary, the more publications that are co-authored by a collaborative pair, the more similar the representation of each author's research topic. Thus the high impact we see among these authors is likely a result of sustained and stable collaboration, rather than being related to disciplinary focus. This manifests in our distribution as a slightly higher average topic similarity and a more narrow spread of values. Although our analysis is complicated by our measurement of transdisciplinarity and stability, it still stands as an example for future research to build upon. Researchers are more than ever crossing disciplinary boundaries (van Noorden 2015), and thus new methods are needed to understand how disciplinary differences might benefit or doom collaborative relationships.

Scientific age, measured using the date of the first publication, and scientific career impact, measured using h-index, are related. More senior and established scholars are likely to be older, as they have likely had more time to build their publication portfolio. We calculate the h-index using the state of a researcher's career at the end of our data, in 2014, while scientific age is calculated using each author's first and last publications. This methodological discrepancy should not cause many significant biases, given that h-index is predicted to increase linearly over time (Hirsch 2005; Guns and Rousseau 2009). There are however variations in the stability distributions of each measure, likely the result of different career trajectories of students. Researchers might change career focus, publishing outside of computer science, or exceed the success of their advisor. Differences in the distribution of stability across scientific age categories likely characterize different types of relationships between academics. Those with low similar career impacts tend to have wide variation in stability, while the collaborative pairs with the most similar ages tend to be more stable, and with a somewhat less "diversity" of stability. On the other hand, those with dissimilar impacts and ages are likely to reflect student-advisor or junior-senior collaborative relationships. Student-advisor relationships are likely to consistently produce with one another for a period of time, leading to the higher measures of stability that we observe. Those with similar ages and career impacts are likely to reflect colleague–colleague relationships, their medium–low stability likely indicative of more sporadic output.

Although the distributions of the occurrence of the difference between our measures scientific age and career impact produce contrasting visualizations, their relationship with scientific impact manifests with similar patterns. When considering differences in scientific age and career impact, the most impactful work tends to come from the most stable relationships. But among these stable collaborative pairs, the highest average impact work tends to come either highly similar, or highly dissimilar co-authors; the former are likely to be medium–high or high stable colleagues, maintaining strong long-term relationships, while the latter are more likely to be stable relationships between junior scholars and mentors. Junior scholars benefit from these relationships because they offer opportunities to learn, develop professional networks, and access resources they may not have otherwise had (Hakala 2009a, b; Straus et al. 2009). Our findings suggest that junior scholars benefit most from medium–high or high stable relationships with senior researchers, giving them ample time to develop into mature and independent researchers. The other group of high-impact collaborations, defined by their low stability and dissimilar ages, may also be students co-authoring with their advisor's colleagues or masters advisors (before attending a Ph.D.), or who publish during a graduate program but later leave academia.

Maintaining large teams can be difficult. Large teams conducting creative work with the goal to "innovative" were found to have poorer team processes, especially when under increased pressure (Curral et al. 2001). Academic research teams tend to be required to "innovate" under high-pressure competitive conditions, and thus may be subject to similar challenges. We see a "U" shape in the median stability of different team sizes. Team sizes with only two or three people tend to have the highest stability, which makes sense, given that these teams would be the easiest to coordinate and maintain. However large teams, with 6+ collaborators, also tend to have relatively high median stability, at least more than those teams with four or five members. Small teams are easy to coordinate and maintain, leading to higher stability, but larger teams may be part of larger labs, in which teams of researchers consistently collaborate on projects.

When we turn to examine how teams relate to impact, we see that large, stable teams tend to produce the most impactful research, likely the result of the survivorship of only the best and most effective relationships. Stable collaborative pairs with large average team size have maintained a close relationship with one another while also managing relationships with a host of other collaborators. Either these collaborative pairs are reaping the benefits of their invested social capital, or else they are stronger researchers, capable of producing high-impact work while navigating tumultuous interpersonal conditions, or some mixture of both; future analysis is needed to understand the direction of causality.

Integrating stability and persistence

This analysis has largely followed from our previous analysis in Bu et al. (2017), and the relationships we find concerning stability are similar to those concerning persistence. But persistence and stability are distinct characteristics of scientific collaborations. In our study we analyzed how stability and persistence might interact to influence the "success" of collaborations, measured by scientific impact, and how these indicators might be used together to draw new insights.

Combining persistence and stability, we can gain a better understanding of the structure of collaboration. One example of their use is the identification likely advisor/advisee relationships in the data. For example, medium-stability collaborative pairs are most likely to be low persistence, likely students consistently publishing two or three papers over the course of their graduate program or post doc, and then ending the relationship, resulting in low persistence. Similar trends repeat for medium–high and high stability groups, where the most common values for persistence reflect publication outputs, and program lengths of various student-advisee relationships. Low stability collaborative pairs tend to have more distributed degrees of persistence, potentially reflecting a greater diversity of short-term, medium-term, and sporadic relationships that a researcher might have over the course of their career. Adding persistence allows for a more careful and nuanced representation of the population that stability alone cannot provide.

The two measures can also be combined to examine how stability relates to impact. In previous work, we found that medium–high persistence collaborative pairs have the highest average scientific impact (Bu et al. 2017), and in this study, we found the same relationship for medium–high stability pairs. Those collaborative pairs at the intersection of these two groups have the highest average impact of all pairs, suggesting that neither stability nor persistence are sufficient on their own—maintaining a reasonably high, but not excessive degree of both is important. But in general, persistent yet unstable groups tend to have more scientific impact than stable yet un-persistent groups, suggesting that while stability is important, persistence remains a key to success in science (Ioannidis et al. 2014).

To understand scientific collaborations, a host of perspectives is needed. Each indicator has the ability to reveal some new insights, while also obscuring others. Here we showed how indicators of collaborative persistence and stability might be used together to better understand the structure of scientific collaboration and collaboration "success". Moving forward, we can use these two indicators to create a new taxonomy to characterize scientific collaborations. Figure 6 presents this new framework for characterizing the nature of the scientific collaboration, diving them into four types based on their measures of stability and persistence. Each type represents a distinct type of collaboration, likely with its own demographics, topic focuses, and characteristics. Using this framework, as well as the indicators of collaborative stability introduced here and the indicator of persistence,

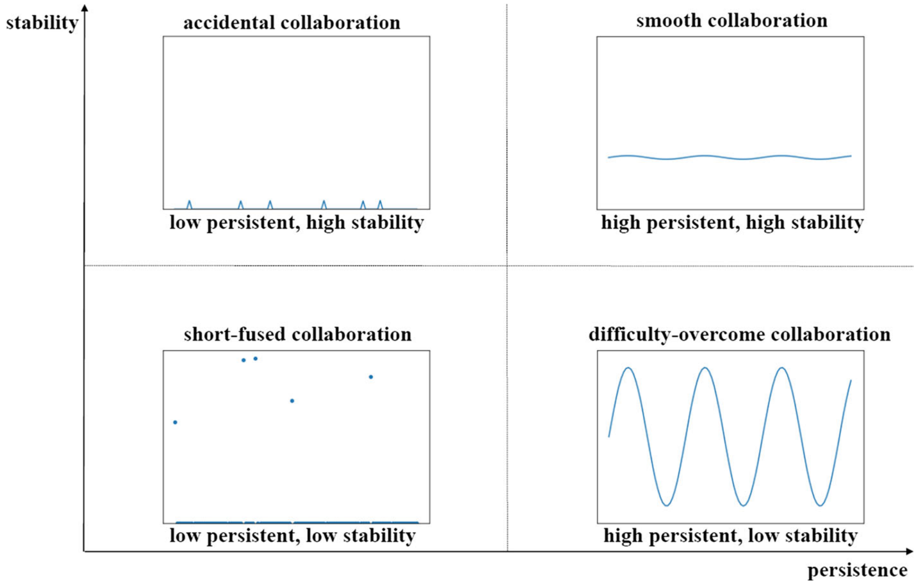

Fig. 6 Diagram of four types of collaborations based on their persistence and stability

future research can seek a better understanding of the different types of scientific collaboration.

Limitations

This study is subject to several limitations that must be considered when interpreting results and drawing conclusions. Some of these limitations have been discussed in previous sections, but others need to be stated or re-stated.

One limitation is that the stability indicator itself might sometimes produce unintuitive results. For example, the stability will be the same for a collaborative pair that published once in the first year of analysis, and then never again, and of another collaborative pair that publishes one publication every year, but nothing in the last year of analysis. Unintuitive results often occur in edge cases, and so care must be taken to understand and contextualize the results of any analysis using this indicator of collaborative stability, especially if being used to inform policy decisions.

ArnetMiner was an effective dataset that allowed us to demonstrate the utility of measuring collaborative stability, but it also limits our analysis. Our data includes only publications appearing in the field of computer science; our analysis thus does not consider an author's career publication or collaboration portfolio outside computer science, citations they might receive from journals in other disciplines, or how collaborative dynamics might differ across disciplines. While our indicator of stability is generalizable to all disciplines, the external validity of our analysis is limited, and care is needed to generalize our findings outside of computer science.

A limitation of most scientometric works is that they rely on citations counts as an indicator of publication quality and scientific success. Citations have often been criticized as being poor indicators of quality, especially for short-term analyses (Leydesdorff et al. 2016). Our measures of scientific success, YANC and h-index, both rely on short-term citation counts and thus should be considered as only a convenient measure for our preliminary analysis, requiring further contextualization and discussion of limitations in future analyses.

Finally, the most pressing issue of this analysis is that we do not control for authors entering or leaving the population at different points. For example, to calculate stability, we divide by the total span of our analysis, no matter when a collaborative pair first appears in the dataset. Thus, a collaborative pair that publishes their first papers in 2009 will have the same stability as one who published once in 2001, even though the 2009 pair has less time to demonstrate stability. Collaborations begin at different time, and while the addition of persistence can help control for some of the effects, more work is needed to extend this preliminary analysis. Using our study as a starting point, more work is needed to expand upon these findings that can consider the dynamics of the population.

Conclusions

In this study we introduced a new indicator, a measure of collaborative stability, which describes the degree of sustained and consistent investment in a collaboration relationship, measured by publication output. We used this indicator to analyze collaborative stability within a dataset of researcher publishing in the field of computer science. We find that the population of collaborative pairs that maintains a medium–high measure of stability are the most common collaborative pairs, and also have the highest average scientific impact. We show how this relationship differs based on a host of factors related to the composition of the collaborative pair, including similarity in research topic, difference in h-index and scientific age, and average team size. We then show how our indicator of stability can be used in conjunction with another indicator we introduced in a previous study, persistence (Bu et al. 2017), to gain a deeper understanding of the characteristics of scientific collaboration. Finally, we introduce a new framework for classifying scientific collaborations using both persistence and stability. With this new framework, we provide a starting point upon which future work can build, and thus to gain a deeper and more nuanced understanding of scientific collaborations.

Scientific collaborations between researchers are a complex phenomenon, vulnerable to a myriad of factors inside and outside of each researcher's control. In a world where both the incidence and size of collaborations are increasing (Larivière et al. 2015), it is increasingly important to understand the nuances of research collaboration. Stability is one characteristic of collaborations, and knowledge about how to promote stability, and how stability affects the success of collaborations might prove relevant to many stakeholders. Research managers, administrators, and those otherwise involved in research teams might want to know how to make their teams more stable. Policy makers may be interested in policies that promote a certain level of collaborative stability while also incentivizing a certain amount of exploratory behaviors so that researchers do not become isolated from outside knowledge or new ideas and techniques.

Researchers have a finite amount of time, and must make careful decisions about how to invest their resources, whether they are physical, such as funding and equipment, or intangible, such as in social capital and human relationships. If a researcher hopes to maximize their scientific impact, they may decide to invest their time in continuously publishing papers (Ioannidis et al. 2014), in collaborating with others in order to solve more complex problems (Stokols et al. 2008b), and to establish strong and super ties with some of these collaborators (Petersen 2015). With our findings however, we show that it is not simply the *amount* that a researcher invests in persistence and collaboration that is important, but also the consistency. To draw from a cliché, a scientific career is like a marathon that requires a constant and continuous pace, and this research is some of the first to investigate how to set this pace.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant Numbers: 71420107026 and 71403190).

References

Abramo, G., D'Angelo, C. A., & Costa, F. D. (2009). Research collaboration and productivity: Is there correlation? *Higher Education, 57*(2), 155–171.

Akgün, A. E., & Lynn, G. S. (2002). Antecedents and consequences of team stability on new product development performance. *Journal of Engineering and Technology Management, 19*(3), 263–286.

Bogomol'Nyi, E.B. (1976). The stability of classical solutions. *Soviet Journal of Nuclear Physics (English Translation)*, *24*(4).

Bromham, L., Dinnage, R., & Hua, X. (2016). Interdisciplinary research has consistently lower funding success. *Nature, 534*(7609), 684–687.

Bu, Y., Ding, Y., Liang, X., & Murray, D. S. (2017). Understanding persistent scientific collaboration. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.23966.

Cannon, W. B. (1932). *Homeostasis. The wisdom of the body*. New York: Norton.

Curral, L. A., Forrester, R. H., Dawson, J. F., & West, M. A. (2001). It's what you do and the way that you do it: Team task, team size, and innovation-related group processes. *European Journal of Work and Organizational Psychology, 10*(2), 187–204.

Eitzen, D. S., & Yetman, N. R. (1972). Managerial change, longevity, and organizational effectiveness. *Administrative Science Quarterly, 17*(1), 110–116.

Ellingsen, T., & Palm, E. (1975). Stability of linear flow. *The Physics of Fluids, 18*(4), 487–488.

Emmons, K. M., Viswanath, K., & Colditz, G. A. (2008). The role of transdisciplinary collaboration in translating and disseminating health research: Lessons learned and exemplars of success. *American Journal of Preventive Medicine, 35*(2), S204–S210.

Finholt, T. (1999). Collaboratory life: Challenges of Internet-mediated science for chemists. In National Research Council (Ed.), *Impact of advances in computing and communications technologies on chemical science and technology*. Washington, DC: The National Academies Press.

Fridman, A. M., & Polyachenko, V. L. V. (2012). *Physics of gravitating systems I: Equilibrium and stability*. Berlin: Springer.

Grusky, O. (1963). Managerial succession and organizational effectiveness. *American Journal of Sociology, 69*(1), 21–31.

Guns, R., & Rousseau, R. (2009). Simulating growth of the h-index. *Journal of the American Society for Information Science and Technology, 60*(2), 410–417.

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology, 47,* 307–333.

Hakala, J. (2009a). Socialization of junior researchers in new academic research environments: Two case studies from Finland. *Studies in Higher Education, 34*(5), 501–516.

Hakala, J. (2009b). The future of the academic calling? Junior researchers in the entrepreneurial university. *Higher Education, 57*(2), 173.

Hannan, M. T., & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review, 49*(2), 149–164.

Hara, N., Solomon, P., Kim, S. L., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology, 54*(10), 952–965.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569–16572.

Ioannidis, J. P. (2011). More time for research: Fund people not projects. *Nature, 477*(7366), 529–531.

Ioannidis, J. P. A., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLoS ONE, 9*(7), e101698.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly, 33*(1), 81–104.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55,* 623–655.

Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science, 51*(14), 1306–1320.

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, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673–702.

Leydesdorff, L., Bornmann, L., Comins, J., & Milojević, S. (2016). Citations: Indicators of quality? The impact fallacy. Retrieved from http://arxiv.org/abs/1603.08452.

Milojević, S. (2010). Modes of collaboration in modern science: Beyond power laws and preferential attachment. *Journal of the American Society for Information Science and Technology, 61*(7), 1410–1423.

Molina-Morales, F. X., & Martínez-Fernández, M. T. (2009). Too much love in the neighborhood can hurt: How an excess of intensity and trust in relationships may produce negative effects on firms. *Strategic Management Journal, 30*(9), 1013–1023.

Nystrom, P. C., & Starbuck, W. H. (1984). To avoid organizational crises, unlearn. *Organizational Dynamics, 12*(4), 53–65.

Pennington, D. D., Simpson, G. L., McConnell, M. S., Fair, J. M., & Baker, R. J. (2013). Transdisciplinary research, transformative learning, and transformative science. *BioScience, 63*(7), 564–573.

Petersen, A. M. (2015). Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences of the United of America, 112*(34), E4671–E4680.

Schaltegger, S., Beckmann, M., & Hansen, E. G. (2013). Transdisciplinarity in corporate sustainability: Mapping the field. *Business Strategy and the Environment, 22*(4), 219–229.

Simberloff, D. (2014). The "Balance of Nature"—Evolution of a panchreston. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.1001963.

Smith, E. D., & Nyman, R. C. (1939). *Technology and labor: A study of the human problems of labor saving*. New Haven, CT: Yale University Press.

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008a). The science of team science: Overview of the field and introduction to the supplement. *American Journal of Preventive Medicine, 35*(2), S77–S89.

Stokols, D., Misra, S., Moser, R. P., Hall, K. L., & Taylor, B. K. (2008b). The ecology of team science: Understanding contextual influences on transdisciplinary collaboration. *American Journal of Preventive Medicine, 35*(2), S96–S115.

Straus, S. E., Chatur, F., & Taylor, M. (2009). Issues in the mentor–mentee relationship in academic medicine: A qualitative study. *Academic Medicine, 84*(1), 135–139.

Tang, J., Fong, A. C. M., Wang, B., & Zhang, J. (2012). A unified probabilistic framework for name disambiguation in digital library. *IEEE Transaction on Knowledge and Data Engineering, 24*(6), 975–987.

Tang, J., Jin, R., & Zhang, J. (2008b). A topic modeling approach and its integration into the random walk framework for academic search. In *Proceeding of the eighth IEEE international conference on data mining* (pp. 1055–1060). Pisa, Italy.

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008a). ArnetMiner: Extraction and mining of academic social networks. In *Proceedings of the fourteenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 990–998). Las Vegas, NV.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review, 61*(4), 674–698.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. F. (2013). Atypical combinations and scientific impact. *Science, 342*(6157), 468–472.

Van Noorden, R. (2015). Interdisciplinary research by the numbers. *Nature News, 525*(7569), 306.

Wray, K. B. (2002). The epistemic significance of collaborative research. *Philosophy of Science, 69*(1), 150–168. https://doi.org/10.1086/338946.

Wu, S., Xu, J., & Ding, Y. (2017). Discover citation topic distribution patterns of highly cited papers. In *iConference 2017 global collaboration across the information community* (pp. 739–743). Wuhan, Hubei, China.

Wu, Y., Venkatramanan, S., & Chiu, D. M. (2016). Research collaboration and topic trends in Computer Science based on top active authors. *PeerJ Computer Science, 2,* e41.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science, 316*(5827), 1036–1039.