

Analyzing Scientific Collaboration with “Giants” based on the Milestones of Career

Yi Bu

Indiana University, U.S.A.

Dakota S. Murray

Indiana University, U.S.A.

Jian Xu *

Sun Yat-sen University, China, issxj@mail.sysu.edu.cn

Ying Ding

Indiana University, U.S.A.

Peng Ai

University of Wisconsin-Madison, U.S.A.

Jinhua Shen

Tongji University, China

Fan Yang

Peking University, China

ABSTRACT

This paper explores how the patterns of collaboration change throughout a scientist’s career, especially their collaboration with so-called “giants”—well-known elite scholars with very high scientific impact. We divide scientist’s careers into five stages based on the number of years since obtaining their Ph.D., and for each stage, we calculate a number of indicators, such as publication count, normalized citation count, count of collaborations, count of collaborations with “giants”, and percentage of uninterrupted and continuous presence. Our analysis shows that although collaborating with “giants” may benefit a young scholar, they are more likely to achieve greater scientific success if they first work independently and only collaborate with “giants” later in their career. Meanwhile, high-impact authors tend to work with “giants” early in their career, but not in the later stages, which forces us to reconsider the established notions of when is best to collaborate with elite scholars. Moreover, we find that high-impact authors tend to have fewer co-authors per article, regardless of the stage of their careers, indicating that collaboration is important but the size of collaborations also matters.

KEYWORDS

Scientific collaboration, scientific career, authoritative author, pattern, scientometrics.

INTRODUCTION

A professional career tends not to be uniform—different career stages may bring changes to a person’s tasks, behavior, and work strategies. For example, at the beginning of the career, a promising junior entrepreneur might be eager to learn how to bring newly designed products to market or acquire venture funding, while senior businesspeople may be more likely to identify trading opportunities instead of indulging in specific product development or marketing tactics. To achieve success throughout each stage of a professional career, diverse strategies are *de facto* required. Similarly, in academia, successful scholars employ distinct strategies at each stage of their career, such as changing their collaboration patterns (Kumar & Ratnavelu, 2016).

Scientific collaboration has become increasingly prevalent in many fields of science (Zhang, Bu, Ding, & Xu, 2018). Collaborations may even lend to some scholar’s success, having been shown to often result in more impactful products (Wuchty, Jones, & Uzzi, 2007), and also breed innovation (Camarinha-Matos & Afsarmanesh, 2005). Nevertheless, while successful scholars may be motivated to collaborate widely, the nature of their collaborations may change over the course of their scientific career. For example, young scholars might take the responsibility to write papers and do empirical studies (Baker & Pifer, 2011; Holgate, 2012), but will gradually start to lead research projects after several years, perhaps taking the role of a corresponding author in their collaborations (Du & Tang, 2013). Alternatively, at the beginning of their careers they may seek to collaborate with so-called “giants”, who are well-known and elite authors in their field (Amjad *et al.*, 2017), but might later work more with students.

“Standing on the shoulders of giants” is a phrase repeated so often and nestled so deeply in the cultural lexicon of science that scholars have authored entire books tracing its attributions and uses (Merton, 1993). Often when scholars employ this adage it is in to reference the cumulative nature of science, with these “giants” taking the shape of famous yet long-gone scholars such as Isaac Newton and Albert Einstein; but others imply that modern scholars also stand on the shoulders of still-living “giants”. Adegbola (2013) described a phenomenon wherein young scholars achieve career success by networking and benefiting from the expertise of their mentors and professional partners. Some scholars, by effort, luck, or both, gain access to the expertise and mentorship of so-called “giants” within their field of study (also referred to as authoritative authors, or AAs, in this paper). Working with AAs provides young scholars the unique opportunity to learn from elites, grow their professional networks, and help them achieve more (Amjad *et al.*, 2017), allowing them to better compete with others for resources, recognition, and employment (Kram & Isabella, 1985; Quatman & Chelladurai, 2008; Adegbola, 2010). Both metaphorically and practically,

standing of the shoulders of individual “giants” is considered valuable which might motivate young scholars to seek out and collaborate with these giants.

While past research has examined the benefit of working with authoritative authors for a young scholar, little previous research has explored whether scholars should also collaborate with authoritative authors later in their career; answering this question could be crucial to help us identify successful collaboration patterns from a career-long perspective. This paper addresses this question by analyzing the entire career of researchers in the field of computer science who collaborate with authoritative authors (AAs), and examines the relationships between these collaborations and their scientific impact at each career stage.

RELATED STUDIES

Different Stages of Scientific Careers

Most researchers start their scientific careers by joining a mentor-mentee program such as doctoral training (Subramanyam, 1983; Hart, 2000), where they publish more papers than those who do not (Muschallik & Pull, 2016). Compared with other career stages, junior researchers are more likely to take the responsibility for writing manuscripts, conducting empirical studies, and implementing ideas (Pachalen & Bhattacharya, 2015) and are less likely to lead the entire research team (DeCastro, Sambuco, Ubel, Stewart, & Jaqsi, 2013). Meanwhile, young scholars during their doctoral studies tend to have more first-authored publications (Costas & Bordons, 2011). A number of researchers are interested in providing feedback to young scholars to help them achieve professional success and evaluate their career development; for example, Buddeberg-Fischer, Stamm, Buddeberg, and Klaghofer (2008) designed a one-dimensional career-success scale targeted to young researchers in the field of physics and proposed several approaches on guiding them towards high achievement. At this early stage of their career, scholars are meant to be gaining experience and preparing themselves for greater achievement later in their careers (Barbuto, Story, Fritz, & Schinstock, 2011; Schlosser & Gelso, 2001).

Having completed coursework, passed candidacy exams (e.g. qualifying exams) and defended their dissertation, junior researcher are ready to transit from doctoral students into independent scholars (Baker & Pifer, 2011). After they start their new careers, they might have research assistants or postdoctoral scholars working under them (Kumar & Ratnavelu, 2016). Nevertheless, their focus switches from purely publishing articles to writing research proposals, applying for funding, and supervising young scholars (Gulbradsen & Smeby, 2005; Price & Cotton, 2006). At this stage of career, scholars may also begin to experience stress related to productivity, promotion, and tenure—being pushed towards activities that more directly benefit their institutions (Raelin, 1985). When researchers advance to a senior or tenured position, they often serve as interpreters and inter-veners (Jaeger, Sandmann, & Kim, 2011) and may begin to lead research teams (Rowlands & Nicholas, 2006). Gingras, Lari-vière, Macaluso, and Robitaille (2008) found that senior researchers have a smaller ratio of first-authored articles, indicating that they might be managing research teams rather than executing the work. Meanwhile, senior researchers tend to spend less time on research (Kanda & Tomizawa, 2015), finding their research time limited by an increasing load of administrative duties (Barham, Foltz, & Prager, 2014; Kawaguchi, Kondo, & Saito, 2016) and teaching/service commitments (Taylor, Fender, & Burke, 2006). During this period, researchers tend not to seek mentors in an institutionalized or official way, as is usually the case while completing their education (Kram, 1985), but those wishing to cultivate mentor-mentee relationships for career progression purposes still may “seek mentors from within their organizations” (van Eck Peluchette & Jeanquart, 2000, p. 552), allowed for by the strong commitment they have given their institutional, and the shared organizational culture and values (Raelin, 1985). This later career and life stage was also found to have greater stability and less interest in relocation (van Eck Peluchette & Jeanquart, 2000). Each stage of a researcher’s career brings with it distinct focuses, behaviors, and activities.

Scientific Collaboration and Career

Scientific collaboration is often considered as crucial to research in most disciplines and regardless of career stage (Wuchty *et al.*, 2007). However, the nature and frequency of collaborations varies between different career stages. For example, an international survey of 581 researchers in economics found that those who joined their current institution less than one year before have the highest proportion of multi-authored papers, which is untrue for researchers who have stayed with their current institution for ten years or more (Kumar and Ratnavelu, 2016). This study also found that researchers aged 56 or older have significantly fewer multi-authored publications than researchers do aged 45 or younger. Moreover, senior researchers show greater variation in the number of co-authors per publication than younger scholars who tend to have similar numbers of co-authors across articles (Hamermesh, 2015). Bu *et al.* (2018b) divided the high-impact researchers’ careers into two parts: before and after Ph.D graduation, and found different collaboration patterns in each stage, corresponding to findings of previous research (Krapf, 2015). Moreover, in different stages of career, the division of tasks in a collaboration is also different; specifically, as pointed out by Kumar and Ratnavelu (2016), the load of tasks is “predominantly on the shoulders of the mentees” (p. 14) (or junior researchers) instead of mentors (or senior researchers) who instead provide guidance to their research assistants, post-doctoral scholars, and Ph.D. students.

Scientific Collaboration with “Giants” (Authoritative Authors)

We often see that scholars try to meet and work with “giants” in a domain because they might provide knowledge, experience, and resources (Adegbola, 2013). Kumar and Ratnavelu (2016) found that for junior researchers, increasing their productivity is the most important reason to collaborate with “giants”; other reasons are those such as helping young scholars increase the article’s likelihood of acceptance (Bhopal *et al.*, 1997; Bennett & Taylor, 2003). Some scholars have demonstrated the benefits of collaborating with “giants”. For instance, Ebadi and Schiffauerova (2015) highlighted the significant role of social networking to collaboration, demonstrating that a scholar is more likely to secure more funding if they join a large research team and get connected to productive researchers who “have also a good control over the collaboration network and the flow of information” (p. 1). Packalen and Bhattacharya (2015) found that in biomedical sciences, young scholars are more likely to try out new ideas when encouraged and guided by the mentorship of more experienced researchers.

Newton claimed to have stood on the shoulders of giants; Adegbola (2013) used this as a starting point to argue that scholars in the field of healthcare could achieve success by working with “giants” in their domains. Amjad *et al.* (2017) expanded this concept and studied how working with giants influences the success of young scholars. They divided authors into two groups, those who have at any point collaborated with “giants” in their domain and those who had not. They further partitioned the first group into those who had worked with “giants” at the start of their career, and those who collaborated with “giants” later in their career. Their results showed that although collaboration with top researchers has significant positive impact on a future rising star, working alone or with “average” researchers first and then later collaborating with elite authors can also lead to success and high achievement. Based on Amjad *et al.* (2017)’s work, this paper aims to analyze the influence of scientific collaboration with “giants” in different careers using various bibliometric indicators.

METHODOLOGY

Data and Processing

The dataset is a part of ArnetMiner dataset (Tang *et al.*, 2008) including the Association for Computing Machinery (ACM) publications in the field of computer science. ArnetMiner has been widely used and considered as one of the best-curated databases for computer science articles (Dong, Johnson, & Chawla, 2016). We obtained 2,092,356 papers published between 1936 and 2014 that contain 1,207,061 unique authors and 8,024,869 citations. Author name disambiguation was performed according to the algorithms proposed by Tong, Fang, Wang, and Zhang (2012). Figure 1 outlines the procedures of data processing. Our analyses are conducted on scholars whose first publication was published between years 2000 and 2004. For each paper in this dataset, we count the number of citations each paper receives from other papers within ArnetMiner. An author’s h-index (Hirsch, 2005), which has been widely used to evaluate a scholar’s scientific achievements (Lacasse, Hodge, & Bean, 2011; Ciriminna & Pagliaro, 2013), is derived according to the number of his/her publications recorded in the dataset and the number of citations received.

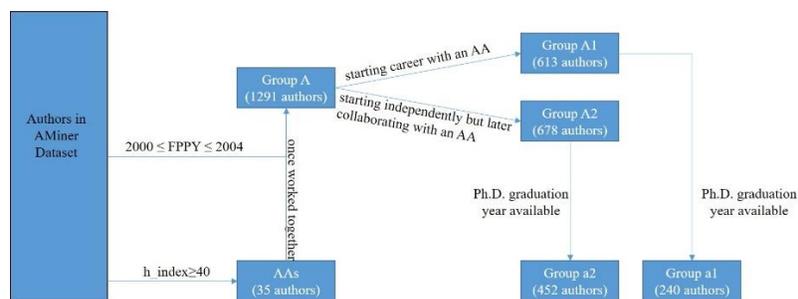


Figure 1. Procedures of data processing. (Note: FPPY = first paper published year).

Defining a “giant” is crucial in this research. As the h-index distribution for all authors is strongly skewed, we refer to Amjad *et al.* (2017) and define “giants” as those with an h-index of 40 or more and name them “authoritative authors” (AAs); there are 35 AAs in our dataset (0.002% out of all authors). In order to examine the difference of performance between scholars who have collaborated with AAs and those who have not, we partition the authors into two groups: group A contains those authors who collaborated with an AA at least once in their career (1,291 authors) and group B contains the remainder. To explore more granular patterns in their career paths, we further divide Group A authors into two sub-groups: Group A1, which includes 613 authors who published their first articles with an AA, and Group A2, which includes those who did not publish first with an AA, but later collaborated at least once with an AA. By identifying the exact year these authors obtained their doctoral degrees, we can approximate the time they start their scientific careers (Subramanyam, 1983; Hart, 2000). We manually obtain the Ph.D. graduation year of all authors in Group A; we could only find such information for 240 and 452 authors in Groups A1 and A2; we name these two new groups with authors’ Ph.D. graduation years as Groups a1 and a2, which comprises our final population for analysis.

To indicate the research performance of authors and better analyze how their careers have been developed, we refer to Amjad *et al.* (2017) and divide all of the authors in Group a1 and a2 into three types: high-impact authors ($h\text{-index} \geq 10$), medium-impact authors ($5 \leq h\text{-index} < 10$), and low-impact authors ($h\text{-index} < 5$) based on their h-indices. Table 1 shows that the percentage of researchers with h-indices five or more in Group a2 is greater than that in Group a1, indicating that researchers who work alone at the beginning of their career and then collaborate with authoritative authors tend to have better research achievements, supporting the conclusions of Amjad *et al.* (2017).

Table 1. Impact categories of authors in Groups a1 and a2.

author types		high-impact authors ($h\text{-index} \geq 10$)	medium-impact authors ($10 > h\text{-index} \geq 5$)	low-impact authors ($h\text{-index} < 5$)	total
Group a1	Number (%)	50 (20.8%)	76 (31.7%)	114 (47.5%)	240 (100%)
Group a2	Number (%)	158 (35.0%)	189 (41.8%)	105 (23.2%)	452 (100.0%)

Methods

In the United States, many doctoral students majoring in computer science are expected to obtain their degrees in four to six years. Doctoral students tend to spend the first years of their study focused on coursework, leaving the last several years before graduation more devoted to research. After graduating with their Ph.D., researchers are likely to work as a postdoctoral scholar, assistant professor, or junior researchers. If they work in universities as faculty members, tenure promotion usually happens during the fourth to sixth year. If they work in industries or research institutions, they are likely to be promoted to senior researchers. Therefore, we define five stages (milestones) among the researchers' whole careers:

- Four years before Ph.D. graduation to Ph.D. graduation (annotated as Year -4-0): they are likely doctoral students;
- One to five years after Ph.D. graduation (Year 1-5): they are likely to be postdocs, assistant professors, or junior researchers (Kawaguchi *et al.*, 2016; University of California Berkley Career Center, 2017);
- Six to ten years after Ph.D. graduation (Year 6-10): they might be assistant professors/researchers or tenured associate professors/researchers;
- 11 to 15 years after Ph.D. graduation (Year 11-15): they might be associate professors/researchers, full professors, or senior researchers (The Princeton Overview, 2017); and
- 16 to 20 years after Ph.D. graduation (Year 16-20): they tend to be more senior and experience, some of whom are likely to reach the end of their careers (Kawaguchi *et al.*, 2016).

For each of these periods, we calculate several bibliometric indicators. The number of publications and citations are traditional indicators to measure the scholars' research performance (de Solla Price, 1976), thus we incorporate them in this study. The number of publications is simply the number of citable items published by an author during each period. For each period, we then calculate the number of all citations made by other publications in the dataset to citable items published by the author during a given time period. For example, consider an author who publishes two journal articles, paper1 in years one-five, and paper2 during years six-ten; he/she does not publish in years 11-15. In years one-five, paper1 received ten citations. For years six-ten, paper1 received five additional citations, and paper2 received twenty citations. In years 11-15, both papers receive six citations. When calculating the author's citation count, they would have 21 ($=10+5+6$) citation in years one-five, 26 ($=20+6$) citations in years six-ten, and zero citations in years 11-15. After that, we normalize the citation count by using the product of the number of publications and the time difference between the published time and the year of 2014, i.e., the paper's age in 2014, in order to mitigate the potential biases caused by the productivity of the author's and the paper's age.

Another of our research goals is to explore scientific collaboration with authoritative authors and how it impacts a scholar's career; Hence, we will also explore relationships between career stage and the number of distinct collaborators (an indicator of collaboration), distinct AA collaborators (indicating authoritative collaborators), and average h-index of collaborators (indicating the performance of collaborators). Moreover, we are also interested in a scholar's total number of collaborators. We use these indicators to analyze the career trajectory of each author in Groups a1 and a2 along different career stages:

- the number of papers published;
- the number of citations received, normalized by each paper's age;
- the number of distinct collaborators;
- the number of distinct AA collaborators;
- the average h-index of the distinct collaborators; and
- the average number of collaborators per article published.

For each of these indicators, we perform median-based tests of statistical significance on using raw values because of the skewness distribution. We also use an indicator of uninterrupted and continuous presence (UCP) to measure the consistency of a researcher's achievements. Here, degree of UCP is defined as publishing at least one/three/five paper(s) in any consecutive

five years. Ioannidis *et al.* (2014) argued that researchers maintaining UCP in publications could achieve higher citation impacts than those who cannot, and that UCP is a latent indicator for measuring scholars' research activity and professional sociability. Hence, the higher UCP a researcher has, the more achievements and stability they might achieve in their scientific career.

RESULTS

Figures 2-4 show the results of career analyses on different indicators of Group a1 and a2. Table 2 lists the statistical significance test results for all indicators. Some publications are excluded from these calculations because they are in some way an "extreme", such as having 15 or more co-authors or more than 10,000 citations. We also exclude certain extreme authors, such as those who have published more than 200 publications within any of the five-year time windows. For some groups there is no value for years 11-15 because no data is available for later period(s). For instance, authors who obtain their Ph.D. after 2003 would enter into years 16-20 after 2014, which is not included in our analysis.

From the perspective of the number of papers one author had in different career stages (Figure 2 left and Table 2), we can see difference emerge between Groups a1 and a2 only after year 0, their Ph.D. graduation. This indicates that although the similar-impact authors in Groups a1 and a2 publish similar number of papers before Ph.D. graduation, a2 tends to be more productive after graduation, and thus more active in scientific activities (de Solla Price, 1976). Specifically, the gap in the number of publications between high-impact authors in these two groups increases as authors become older. Meanwhile, within 15 years of obtaining their doctoral degrees, authors with similar impact tend to publish fewer articles in Group a1 than in a2. However, compared with a1, authors in a2 seem more diligent when they become older. Therefore, although Group a1 was lucky enough to work with AAs at the beginning of their careers, they tended not to be as consistently productive as those in Group a2; on the contrary, a2 authors' hard work may have attracted the attentions of AAs who then began to collaborate with them.

It is also worth noting that high-impact authors in Group a2 publish more papers in years 16-20 than in years 11-15. During years 16-20, many researchers could have obtained their tenured position, some perhaps being promoted to full professors. Although previous studies have shown that a researcher's productivity tends to decrease in the later parts of their careers (Kanda & Tomizawa, 2015), our finding demonstrates that high-impact authors in Group a2, those who started their career independently and later have the chance to work with "giants", continuously publish and make consistent scientific contributions. Regardless of group, low impact authors have the fewest publications while the high-impact authors have the most, likely resulting from the fact that the h-index, our measure of impact, is dependent on the number of publications (Hirsch, 2005).

Table 2. Statistical significance test for indicators in Groups a1 and a2 among high-, medium-, and low-impact groups (Note: *: <0.05; **: <0.01; ***: <0.001.).

indicator	high	medium	low	indicator	high	medium	low
no. of publications	***	**	**	no. of distinct AA collaborators	**	*	***
normalized no. of citations	**	*	**	average no. of h-index of collaborators	*	**	***
no. of distinct collaborators	***	***	**	average no. of collaborators per article	***	***	*

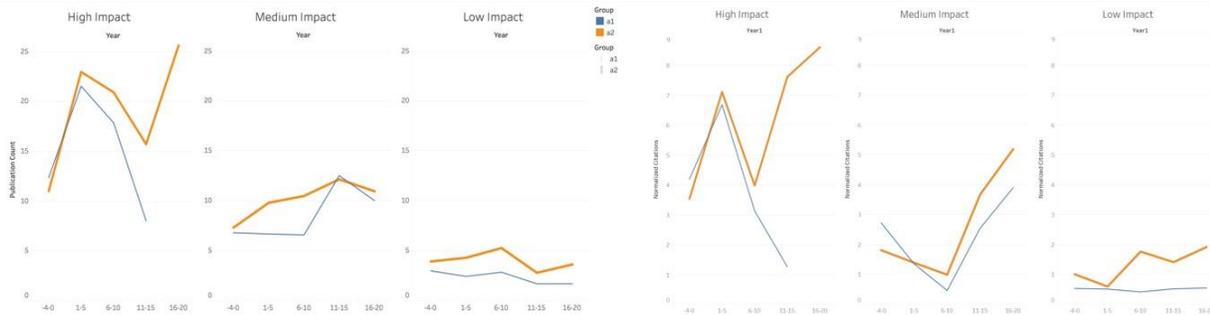


Figure 2. Average number of publications (left) and average normalized number of citations (right) for Groups a1 and a2.

As for the normalized number of citations (Figure 2 right and Table 2), the publications of Group a2 authors tend to receive more citations after we remove the biases of authors' productivity and papers' ages. After receiving their doctoral degrees, authors in Group a2 tend to receive slightly more citations than authors in Group a1, indicating greater scientific success during and immediately after their Ph.D., perhaps a precursor to a successful career. During later career stages, there is greater disparity between high-impact authors of both groups, with high-impact authors in Group a2 vastly outperforming authors in Group a1. These findings are consistent to the results of Amjad *et al.* (2017) who found that Group a2 authors have a tendency to receive more citations than Group a1 authors and such difference becomes larger as their careers progress. Meanwhile, we see a high number of citations in years 16-20 for high-impact authors in Group a2, consistent with the indicator of publication count shown in Figure 2 left. This not only demonstrates the career-long diligence of these authors but also suggests consistent quality of published papers throughout their scientific careers. We can also examine the number of collaborators of authors in each

group, as in Figure 3 left (see significance test results in Table 2, the same below), and we see that the higher the impact of an author, the more collaborators they have, again demonstrating the importance of collaboration (Wuchty *et al.*, 2007). High-impact authors have twice the average number of collaborators than low-impact authors.

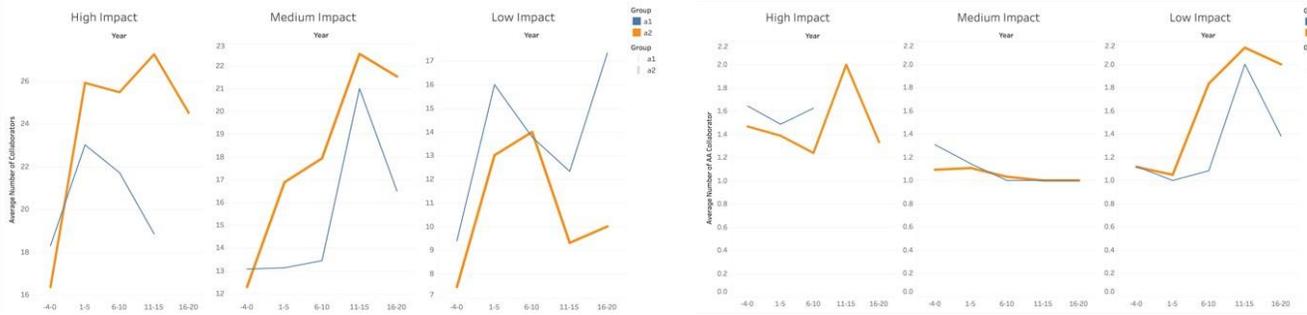


Figure 3. Average number of collaborators (left) and average number of AA collaborators (right) for Groups a1 and a2.

Figure 3 right shows that high-impact authors in these two groups have relatively more AA collaborators before their Ph.D. graduations than after. Collaborating with AAs enables young scholars the chance to “stand on the shoulder” of the giants (Amjad *et al.*, 2017, p. 307), giants who might guide them on how to conduct research and prepare for future work. However, it does not seem that the more collaborations with AAs, the more successful the authors, especially five and ten years after their Ph.D. graduation. One of the interpretations concerning this finding is that there is potential for the “halo effects” of AAs to “outshine” junior scholars and to some extent fetter their development (Bu *et al.*, 2018b). This finding warns us that collaborations with “giants” are not always as beneficial as we might think, especially at the beginning of our careers. Figure 4 left depicts the average h-index of collaborators for Groups a1 and a2 of authors in order to examine the impact of their collaborators. The result shows that the average h-index of collaborators declines as their careers progress, regardless of which group an author is in. This finding confirms our intuitive assumptions: At the beginning of their career, a scholar is likely to be a student, with their main collaborator being their advisor, supervisor, or mentor, who will tend to have a higher h-index than themselves; as time goes by, they may become an established researcher and their collaborators could include more students or junior researchers, decreasing the average h-index of their collaborators. The average number of collaborators per article, shown in Figure 4 right, ranges between two and five. Krapf (2015) points out that the skills and research capability of researchers can be exercised and enhanced by performing different types of tasks, such as designing research, conducting empirical studies, reviewing literature, writing manuscripts, and analyzing data. Exercise and experience with these distinct tasks could help them achieve their future success (Krapf, 2015). Nevertheless, low-impact authors, no matter which group they belong, are found to have more co-authors per article, suggesting a point where too many co-authors ceases to be beneficial.

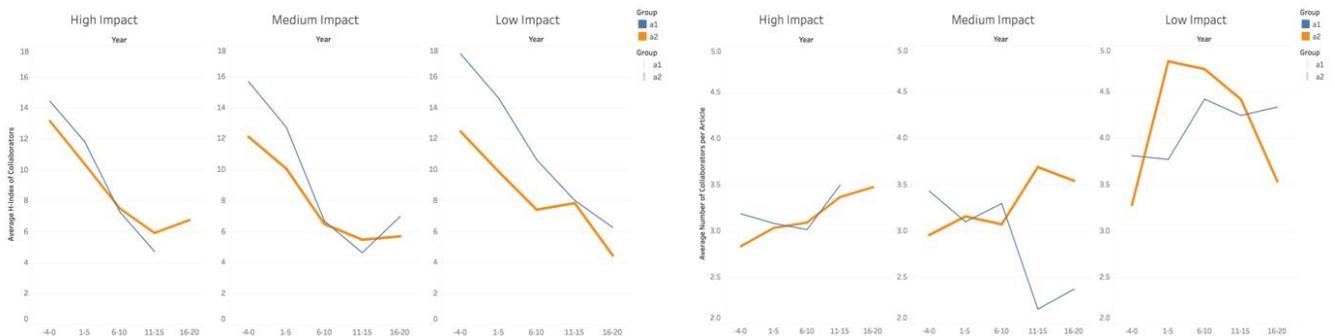


Figure 4. Average h-index of collaborators (left) and average number of collaborators per article (right) for Groups a1 and a2.

Figure 5 shows the UCP results based on different criteria (one, three, and five papers per year), in which we can see that the authors’ UCP percentage in Group a2 is almost twice as many as that in Group a1. Specifically, about 70%, 28%, and 12% of the Group a2 authors can maintain a UCP publication flow for one, three, and five papers per year, while only 39%, 12%, and 5% of Group a1 authors can do so. This high value for UCP indicates that Group a2 has more consistent scholarly activities throughout their careers, corroborating the results of Amjad *et al.* (2017). For all of the authors satisfying UCP criteria in these two groups, we partition them into one of three categories based on whether they achieve their UCP status before, during, or after their Ph.D. graduation. *Before Ph.D. graduation* is defined as the periods before obtaining their Ph.D., including any publications made before they obtain their doctoral degree. *After Ph.D. graduation* includes all papers published by an authors after their Ph.D. graduation. The period of *during Ph.D. graduation* overlaps with before/after Ph.D. graduation; the starting time of during Ph.D. graduation should be before the graduation year of doctoral degree and the ending time should be after the graduation year of doctoral degree. That is to say, the period of *during Ph.D. graduation* should cover one author’s Ph.D.

graduation year. For example, if an author begins his/her Ph.D. in 2005 and completes it in 2010, then *before Ph.D. graduation* would be defined as any time before 2010, *during Ph.D. graduation* would be defined as the period that covers 2010, and *after Ph.D. graduation* as any time after 2010. Figure 6 shows a low percentage of author UCP *before Ph.D. graduation* in both groups. *During the Ph.D.*, the percentage of author UCP in Group a2 is more than that in Group a1, and after graduating with their Ph.D., the gap doubles or even triples. This large discrepancy indicates dramatic variations of career trajectories between groups. Researchers without a UCP record are found more likely to leave academia and research (Ioannidis *et al.*, 2014); hence, Group a2 tends to be successful probably due to their “stable” scientific careers.

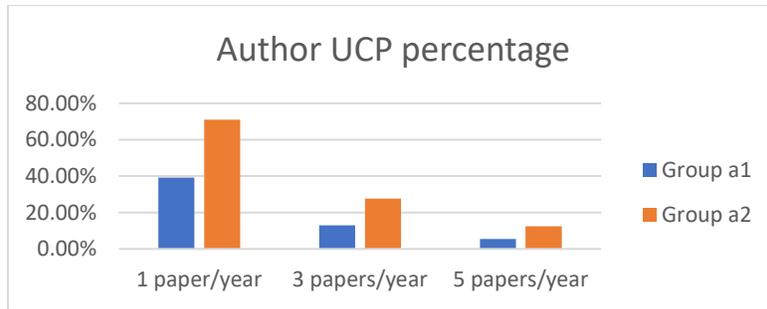


Figure 5. Authors' uninterrupted and continuous presence (UCP) percentage in Groups a1 and a2.

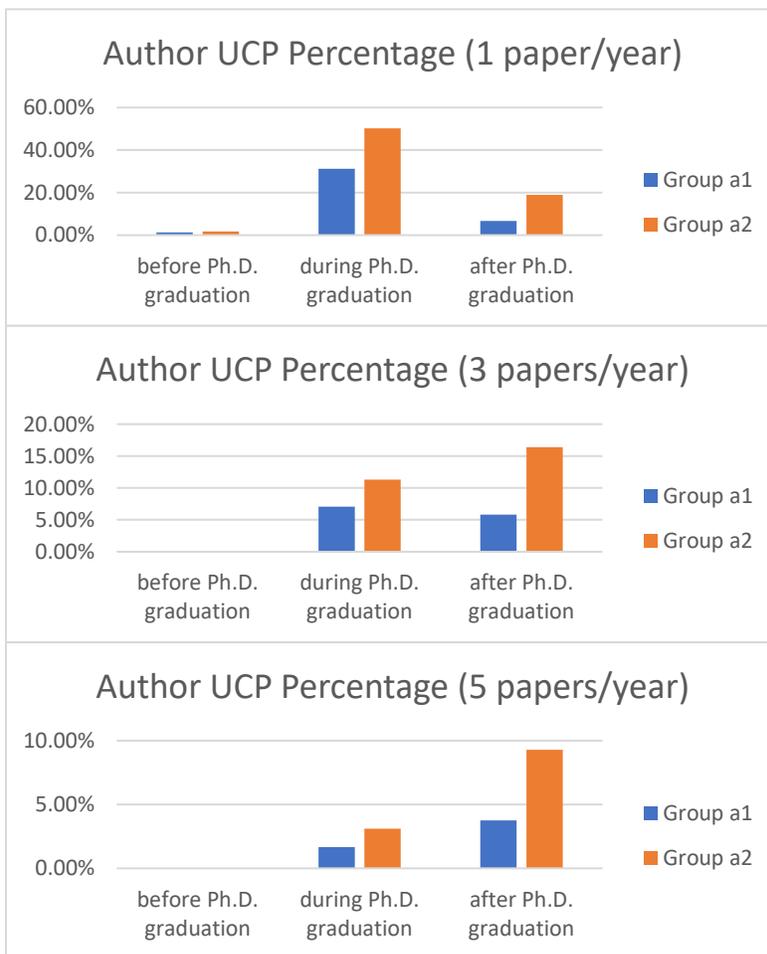


Figure 6. Author UCP percentage in different periods with different criteria: 1, 3, and 5 paper(s)/year.

DISCUSSION AND CONCLUSIONS

Robert Merton detailed the “halo effect” (1968), which implies that the prestige and respect of successful scholars may “rub off” on their close counterparts, but our findings paint a more complicated image. This paper investigates the dimension of time in regards to collaborating with authoritative authors, or “giants”, and explores how working with giants at different stages of a scientific career might lead to distinct outcomes. We divide the careers of researchers into different stages and calculate a number of indicators for each stage (e.g. number of publications, normalized number of citations, number of collaborations, and UCP percentage, etc.). Our results show that although working with “giants” may have benefits and appeal for ambitious

young scholars who prioritize career achievement and productivity; it is hardly a requirement for long-term success. Instead, working hard to become an established researcher, and only later attracting the attention of “giants” to work with and learn from tends to lead to greater professional achievement. The greater determinant of career success seems to be persistent hard work and continuous scientific contribution, evidenced by both our work and that of past researchers (Ioannidis *et al.*, 2014).

The Matthew effect in science (Merton, 1968) implies that if a young scholar dreams of success and prestige, then they would benefit from building cumulative advantage early in their career by finding and working with an elite and prestigious “giant” in their field. Our findings provide evidence that shows this may not be the case. While prestige has benefits, attentive mentorship may be more important for a young scholar just starting their research career. A “giant” may have little time for careful guidance of individuals, and is instead likely has many professional engagements, such as giving external lectures, serving on professional organization committees, mentoring doctoral students, and tending to heavy teaching and service loads (Barham *et al.*, 2014; Taylor, Fender, & Burke, 2006). In such circumstances, a young scholar may benefit from a more junior mentor who can devote more time to individual mentorship.

We need to reconsider the traditional wisdom of academia, and instead realize that there is a right time to prioritize prestige, and a right time to prioritize mentorship. Working with an authoritative author at the right time, when a scholar is still young but with an established career can position them to most benefit from the prestige and experience of such collaborations. Timing “forcefully moves them from inertia to dynamic and explosive scholarly heights” (Adegbola, 2013, p.17). Mentor-mentee program directors and conference organizers may be able to use these findings so that they can design conferences, workshops, and other opportunities that allow beginning, junior, and senior scholars to communicate and interact in a way that most benefits them. When it comes to mentorship and collaboration, there is no universal prescription; rather, policies and programs should reflect the fact that for best results, one must do the right thing at the right time.

This paper also explored the shifting styles of collaboration between levels of impact and throughout a career. We found that high-impact authors tend to have fewer coauthors per article than do medium- and low-impact authors, regardless of the stage of their career; while collaboration is important, this finding also indicates that the total number of collaborations also seems to have an effect. Too many collaborators in a research project may dilute the effectiveness of the team, and result in weaker publications. The increasing of the number of coauthors in one publication also hints at decreasing relative workloads for each author (Schneider & Doyen, 2005). As noted by Curral, Forrester, Dawson, and West (2011), too large of a team may limit performance, which then could limit the quality of the paper. Conducting research in large teams and on many projects may lead team-members to be less engaged with each individual project, and the loss efficiency by spending a disproportionate amount of time organizing the team, rather than improving the research (Vrabie, 2013). This finding adds to the literature of collaboration and offers the insight that more is not always better, a truism that seems to hold throughout a scholar’s career.

This study is one of the first to explore the idea of “doing the right thing at the right time” in academia, and being an exploratory study, suffers from several limitations—most of which result from its scale. Identifying the collaborators of authoritative authors along with their career milestones, such as Ph.D. graduation requires lengthy manual work, and limits the samples size of researchers to only a few hundred. With such a small sample, we cannot control for many of the confounding factors in our work. For example, researchers defined in Group a1, those who begin their career with giants, may be working on their Ph.D. but after graduation may pursue a non-research intensive career, publishing less or not at all, and thus lowering the average of the group’s indicators. Conversely, researchers in Group a2, those who collaborate with giants later in their career are likely to be those who stay in academia and research enough to catch the attention of these giants and are therefore, almost by definition, likely to be stronger and more committed researchers. Differences in sub-discipline may also play a role; we were forced to define a cutoff with which to identify “giants”, however different subfields vary greatly in the number of citations and publications they may receive (Qian, Rong, Jiang, Tang, & Xiong, 2017). A threshold at an h-index of 40 may restrict the number of “giants” in our sample; in the future, we plan to refer to Hohberger (2016) and define “giants” by using the top one and/or two percentile h-index authors. The h-index of 40 threshold might limit the sources of “giants” to a small number of highly productive or high-citation sub-fields which may have distinct patterns of collaboration, mentorship, and career trajectories after Ph.D. graduation; these patterns are likely to diverge even more so between disciplines, such as in the life sciences or humanities. Moreover, because we simply aim to analyze collaboration with “giants” in various careers, this paper does not set as a baseline those authors who never have had any opportunities to work with “giants”. Also, the citation count and h-index are both based on the records within the dataset, which might cause some biases. These limitations lead us to make our conclusions cautiously; while our findings offer evidence for distinct trends and nuances in collaboration over a career, more work is necessary to present a thorough understanding of the field.

Failing to take other co-authors into consideration is one of the limitations of the current paper. It is likely that the papers completed by a certain research and an AA are also co-authored with others. We will try to address this problem in the near future. On the other hand, ignoring the asymmetry between two co-authors is also imperfect. A potential approach to deal with it is to consider the preferred partners of a given author, such as employing Affinity Index and Probabilistic Affinity Index

(Chinchilla-Rodríguez *et al.*, 2018). Meanwhile, we hope to work to increase and diversify our sample of researchers and control for these confounding factors. We also hope to expand the range of disciplines in order to understand whether the patterns of collaboration we find in the careers of computer scientists are similar in other fields. Moreover, we will expand our suite of indicators, which while presently useful for creating a rough understanding of a scholar's career achievement, do not capture the nuances of scientific impact. More indicators will also allow us to explore other aspects of collaboration and career stage, such as changes in impact diversity, research topic, collaborator diversity, and the degree of persistence (Bu *et al.*, 2018a) and stability (Bu *et al.*, 2018c) regarding scientific collaborations. This work lays the foundation for future exploration of how collaboration and other aspects of a researcher's career may affect future success, and to identify not only how to shape a successful career, but also when to prioritize different tasks and behaviors.

ACKNOWLEDGEMENT

The authors would like to thank three anonymous reviewers for their improving the quality of this paper.

REFERENCES

- Adegbola, M. (2010). Nurses collaborating with cross disciplinary networks: starting to integrate genomics into practice. *Journal of National Black Nurses Association*, 21(1), 46-49.
- Adegbola, M. (2013). Scholarly tailgating defined: A diverse, giant network. *The ABNF Journal: Official Journal of the Association of Black Nursing Faculty in Higher Education, Inc.*, 24(1), 17-20.
- Amjad, T., Ding, Y., Xu, J., Zhang, C., Daud, A., Tang, J., & Song, M. (2017). Standing on the shoulders of giants. *Journal of Informetrics*, 11(1), 307-323.
- Baker, V.L., & Pifer, M.J. (2011). The role of relationships in the transition from doctoral student to independent scholar. *Studies in Continuing Education*, 33(1), 5-17.
- Barbuto, J.E., Jr., Story, J.S., Fritz, S.M., & Schinckel, J.L. (2011). Full range advising: Transforming the advisor-advisee experience. *Journal of College Student Development*, 52(6), 656-670.
- Barham, B.L., Foltz, J.D., & Prager, D.L. (2014). Making time for science. *Research Policy*, 43(1), 21-31.
- Bennett, D.M., & Taylor, D.M. (2003). Unethical practices in authorship of scientific papers. *Emergency Medicine Australasia*, 15(3), 263-270.
- Bhopal, R., Rankin, J., ..., & Rodgers, H. (1997). The vexed question of authorship: Views of researchers in a British medical faculty. *British Medical Journal*, 314(7086), 1009-1012.
- Bu, Y., Ding, Y., Liang, X., & Murray, D.S. (2018a). Understanding persistent scientific collaboration. *Journal of the Association for Information Science and Technology*, 69(3), 438-448.
- Bu, Y., Ding, Y., Xu, J., Liang, X., Gao, G., & Zhao, Y. (2018b). 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., Ding, Y., Huang, Y., & Zhao, Y. (2018c). Measuring the stability of scientific collaboration. *Scientometrics*, 114(2), 463-479.
- Camarinha-Matos, L.M., & Afsarmanesh, H. (2005). Collaborative networks: A new scientific discipline. *Journal of Intelligent Manufacturing*, 16(4-5), 439-452.
- Chinchilla-Rodríguez, Z., Bu, Y., Robinson-García, N., Costas, R., & Sugimoto, C. R. (2018). Travel bans and scientific mobility: Utility of asymmetry and affinity indexes to inform science policy. *Scientometrics*.
- Ciriminna, R., & Pagliaro, M. (2013). On the use of the h-index in evaluating chemical research. *Chemistry Central Journal*, 7(1), 132.
- 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.
- 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.
- DeCastro, R., Sambuco, D., Ubel, P.A., Stewart, A., & Jaqsi, R. (2013). Mentor networks in academic medicine: Moving beyond a dyadic conception of mentoring for junior faculty researchers. *Journal of the Association of American Medical Colleges*, 88(4), 488-496.
- Dong, Y., Johnson, R.A., & Chawla, N.V. (2016). Can scientific impact be predicted? *IEEE Transactions on Big Data*, 2(1), 18-30.
- Du, J., & Tang, X. (2013). Perceptions of author order versus contribution among researchers with different professional ranks and the potential of harmonic counts for encouraging ethical co-authorship practices. *Scientometrics*, 96(1), 277-295.
- Ebadi, A., & Schiffauerova, A. (2015). How to receive more funding for your research? Get connected to the right people! *PLoS One*, 10(7), e0133061.
- Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J.-P. (2008). The effects of aging on researchers' publication and citation patterns. *PLoS One*, 3(12), e0004048.
- Gulbradsen, M., & Smeby, J.-C. (2005). Industry funding and university professors' research performance. *Research Policy*, 34(6), 932-950.
- Hamermesh, D.S. (2015). Age, cohort and co-authorship. Cambridge, MA: National Bureau of Economic Research.
- Hart, R.L. (2000). Co-authorship in the academic library literature: A survey of attitudes and behaviors. *Journal of Academic Librarianship*, 26(5), 339-345.
- 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.

- Hohberger, J. (2016). Does it pay to stand on the shoulders of giants? An analysis of the inventions of star inventors in the biotechnology sector. *Research Policy*, 45(3), 682-698.
- Holgate, S.A. (2012) How to collaborate. *Science (Careers)*, DOI: 10.1126/science.caredit.a1200082.
- 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.
- Jaeger, A.J., Sandmann, L.R., & Kim, J. (2011). Advising graduate students doing community-engaged dissertation research: The advisor-advisee relationship. *Journal of Higher Education Outreach and Engagement*, 15(4), 5-25.
- Kanda, Y., & Tomizawa, H. (2015). Changes in the ratio of time spent on work activities by university and college faculty members: A comparison of results of the “Survey of full-time equivalency data at universities and colleges” of 2002, 2008 and 2013. NISTEP Research Material 236 (in Japanese). Retrieval from <http://data.nistep.go.jp/dspace/handle/11035/3027>.
- Kawaguchi, D., Kondo, A., & Saito, K. (2016). Researchers' career transitions over the life cycle. *Scientometrics*, 109(3), 1435-1454.
- Kram, K.E., & Isabella, L.A. (1985). Mentoring alternatives: The role of peer relationships in career development. *Academy of Management Journal*, 28(1), 110-132.
- Koch, C.G. (2007). *The science of success: How market-based management built the world's largest private company*. Hoboken, NJ: Wiley.
- Kumar S., & Ratnavelu, K. (2016). Perceptions of scholars in the field of Economics on co-authorship associations: Evidence from an international survey. *PLoS One*, 11(6), e0157633.
- Lacasse, J.R., Hodge, D.R., & Bean, K.F. (2011). Evaluating the productivity of social work scholars using the h-index. *Research on Social Work Practice*, 21(5), 599-607.
- Merton, R.K. (1968). The Matthew Effect in science. *Science*, 159(3810), 56-63.
- Merton, R.K. (1993). *On the Shoulders of Giants*. Chicago: University of Chicago Press.
- Muschallik, J., & Pull, K. (2016). Mentoring in higher education: Does it enhance mentees' research productivity? *Education Economics*, 24(2), 210-223.
- Packalen, M., & Bhattacharya, J. (2015). Age and the trying out of new ideas. National Bureau of Economic Research. DOI: 10.3386/w20920.
- Qian, Y., Rong, W., Jiang, N., Tang, J., & Xiong, Z. (2017). Citation regression analysis of computer science publications in different ranking categories and subfields. *Scientometrics*, 110(3), 1351-1374.
- Quatman, C., & Chelladurai, P. (2008). Social network theory and analysis: A complementary lens for inquiry. *Journal of Sport Management*, 22(3), 338-360.
- Raelin, J. (1985). *The clash of cultures*. Boston: Harvard Business School Press.
- Rowlands, I., & Nicholas, D. (2006). The changing scholarly communication landscape: an international survey of senior researchers. *Learned Publishing*, 19(1), 31-55.
- Schlosser, L.Z., & Gelso, C.J. (2001). Measuring the working alliance in advisor-advisee relationships in graduate school. *Journal of Counseling Psychology*, 48(2), 157-167.
- Schneider, M.E., & Doyen, B. (2005). *The everything guide to writing a book proposal: Insider advice on how to get your work published*. San Diego, California: Adams Media Corporation.
- de Solla Price, J. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5), 292-306.
- de Solla Price, J., & Cotten, S.R. (2006). Teaching, research, and service: Expectations of assistant professors. *The American Sociologist*, 37(1), 5-21.
- Subramanyam, K. (1983). Bibliometric studies of research collaboration: A review. *Journal of Information Science*, 6(1), 33-38.
- 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., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). 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, August 24-27, 2008, Las Vegas, NV., U.S.A.
- Taylor, W., Fender, B., & Burke, K. (2006). Unraveling the academic productivity of economists: The opportunity costs of teaching and service. *Southern Economic Journal*, 72(4), 846-859.
- The Princeton Overview. (2017). A day in the life of a professor. Retrieved from <https://www.princetonreview.com/careers/127/professor>.
- University of California Berkley Career Center. (2017). The transition from graduate student to assistant professor. Retrieved from <https://career.berkeley.edu/PhDs/PhDtransition>.
- van Eck Peluchette, J., & Jeanquart, S. (2000). Professionals' use of different mentor sources at various career stages: Implications for career success. *The Journal of Social Psychology*, 140(5), 549-564.
- Vrabie, A. (2013). Productivity and team size: Less is more. Retrieved from <http://blog.sandglaz.com/productivity-and-team-size-less-is-more/>
- Wuchty, S., Jones, B. F., Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036-1039.
- Zhang, C., Bu, Y., Ding, Y., & Xu, J. (2018). Understanding scientific collaboration: Homophily, transitivity, and preferential attachment. *Journal of the Association for Information Science and Technology*, 69(1), 72-86.