Studying Factors that Influence Scholar Retention in Engineering Education Research

Hanjun Xian, Purdue University, West Lafayette

Hanjun Xian is a PhD candidate in the School of Engineering Education at Purdue University. He holds master’s and bachelor’s degrees in Computer Science and started to pursue his doctoral degree in engineering education in 2009. He is currently working as the student team leader with Dr. Madhavan on the Deep Insights Anytime, Anywhere (DIA2) project, where he develops data mining algorithms and visualizations to allow interactive navigation of large-scale bibliographic data in STEM. His research interests include diffusion of research innovations, information visualization, data mining, Bibliometrics, social network analysis, and user study.

Dr. Krishna Madhavan, Purdue University, West Lafayette

Dr. Krishna P.C. Madhavan is an Assistant Professor in the School of Engineering Education at Purdue University. He is also the Education Director and co-PI of the NSF-funded Network for Computational Nanotechnology (nanoHUB.org). He specializes in the development and deployment of large-scale data and visualization based platforms for enabling learning analytics. His work also focuses on understanding the impact and diffusion of learning innovations. Dr. Madhavan was the Chair of the IEEE/ACM Supercomputing Education Program 2006 and was the curriculum director for the Supercomputing Education Program 2005. In January 2008, he was awarded the NSF CAREER award for work on transforming engineering education through learner-centric, adaptive cyber-tools and cyber-environments. He was one of 49 faculty members selected as the nation’s top engineering educators and researchers by the US National Academy of Engineering to the Frontiers in Engineering Education symposium.
Studying Factors that Influence Scholar Retention in Engineering Education Research

Abstract
Engineering Education Research (EER) has become a rapidly growing and diverse community over the past two decades. Scholars from various engineering disciplines, computer science, and other fields such as education have joined to share their experiences and developed fundamentals for this new domain. Among the newcomers, some eventually become active players while some decide to opt out. To sustain scholar diversity and community development, it is essential to understand the overall retention status and factors that influence new scholars’ tendency for leaving EER. In this paper, we perform a large-scale bibliometric analysis on 21,209 FIE/ASEE papers and 8,040 EER-related NSF awards to develop a community-wide understanding of scholar retention in EER. We further recognize and compare two groups: opt-out scholars and retained scholars in terms of their academic profiles such as level of engagement, collaboration, research areas, affiliations, and funding status. Our findings show the decreasing retention rate from 2000 to 2011 and an average of 5.23 years for which new scholars stayed active in the EER community. During their early career in EER, scholars who decided to quickly opt out tended to commit less effort and had fewer original contributions to their research studies in EER. Also, they worked on broader and fundamental topics and received less EER-related financial support than their peers in the retained group. To the best of our knowledge, the present study is the first attempt to study scholar retention at large scale and also the first quantitative effort to measure scholar retention based on bibliometric data within engineering education research.

1. Introduction

Over the past two decades, Engineering Education Research (EER) has drawn the attention of scholars from a variety of disciplines. From 1993 to 2002, it has been reported that Journal of Engineering Education (JEE) authors primarily came from engineering disciplines but no single discipline dominated. Also, among all JEE authors, 23.9% of them have no engineering or computer science background. A recent study describes the opt-in process of how scholars, especially engineering instructors, start to develop interests in EER.

While these three studies seem to depict a diverse and growing community of EER, it is not uncommon that some newcomers eventually decide to opt out and discontinue their research in EER. Compared to drawing more scholars into EER, it is equally important to retain existing engineering education (ENE) scholars and sustain community diversity. The EER community has recognized the importance of scholar retention. Jamieson et al. proposed a sustainable plan that evolved engineering education scholars, engineering instructors, institutions, and funding agencies to foster a better environment for creating EER innovations. Therefore, it is essential to study retention of ENE scholars and understand factors that influence their tendency to quit or stay in EER.

However, it has never been fully explored as to why among these newcomers, some eventually become active contributors and even key players in the EER community, whereas some other
researchers decided to opt out. Scholars’ decisions of whether to continue to pursue EER may be influenced by many factors such as faculty recruitment and reward standards, scholarly collaboration, time limitation, funding support, awareness of EER fundamentals, and interdisciplinarity\textsuperscript{3,4}. In this paper, we aim to provide an overview of scholar retention in EER and compare the academic profile of new ENE scholars who later grew into active practitioners with those who discontinued their EER pursuit. The definitions of new scholars, active practitioners, retained scholars, and opt-out scholars can be found in Section 2.2. The research questions to be answered are:

1. What is the general scholar retention status in EER and how did it change from 2000 to 2011?
2. What is the academic profile of new ENE scholars who sustain active in EER?
3. What is the academic profile of new ENE scholars who discontinue their EER pathways?

To answer these questions, we analyzed 21,209 papers in ASEE Annual Conference and Exposition (ASEE) and Frontiers in Education (FIE) and 8,040 EER-related awards from National Science Foundation (NSF) over 2000-2011. These two conference proceedings are selected because of their popularity among both newcomers and senior scholars. Directorate for Education & Human Resources (EHR) and Engineering Education and Centers (ENG-EEC) in NSF are the primary funding sources for ENE scholars. Therefore, they are appropriate venues for measuring the research activeness of ENE scholars with different length of experience. Based on scholars’ publication history in ASEE and FIE, 703 retained scholars and 2,913 opt-out scholars are recognized among a total of 22,998 authors analyzed.

In this paper, we focus on the following factors and attempt to understand how they correlate to scholar retention: level of engagement, academic collaboration, research interest, affiliation, and funding support. These attributes are selected based on previous findings\textsuperscript{3,4} that show the primary challenges new ENE scholars encounter include working in interdisciplinary problems, insufficient collaboration opportunity, and lack of funding. A scholar’s affiliation, research area, and co-authors are extracted from the scholar’s ASEE and FIE publications. The funding status is acquired from the award data of NSF. We apply a bibliometric analysis on scholars’ publications and NSF awards. Results of this study can be used to guide the development of policies to retain ENE scholars, predict the possibility of new ENE scholars staying or quitting EER in the future, and eventually contribute to a sustainable and diverse community. We understand that the tenor and quality of such studies are improved as a larger number of data sources are included.

2. Literature review

2.1 Scholar retention

Prior studies about scholar retention have focused primarily on identifying groups of academics with a high turnover rate, understanding reasons of scholars’ leaving, and exploring ways to sustain diversity. In these studies, academics who are more likely to experience a high turnover rate are the target population. This includes new faculty\textsuperscript{5,6}, under-represented faculty\textsuperscript{7}, both new and minority faculty\textsuperscript{8}, and faculty in community colleges\textsuperscript{9}. The above studies, however, are all limited in scope to a single institution. Some researchers go beyond an institution to reach a larger population such as new nursing faculty in the U.S.\textsuperscript{10}, women postdocs in South Africa\textsuperscript{11},
and women and minority faculty in the schools of education in the U.S.\textsuperscript{12}. All these studies have clear boundaries of population to be studied. But in engineering education, the group of scholars most likely to quit or opt-out has never been clearly defined. One goal of this paper is to recognize such population within the limits of the data we use and present their academic profiles. Also, the above existing studies based their analyses on a relatively small sample, which poses a potential risk of whether the results can be generalized to the entire population. Instead, the present study covers ENE scholars in such a large number that they can approximately represent the whole community.

Some researchers have investigated scholar retention in a broader scope. Ress\textsuperscript{13} examined the relationship between unionization and faculty retention based on nationwide survey data of over 700,000 faculty members in the U.S. In a more recent study, Nagowski\textsuperscript{14} analyzed the same database but in a different and shorter period to compare faculty retention rates between public and private institutions. Bratsberg et al.\textsuperscript{15} conducted a longitudinal study based on 30 years of data from 238 faculty and found a substantial penalty as faculty stay longer in academia regardless of whether they move to a different institution or a better position and how productive they are. These existing studies calculated retention rates based on scholars’ titles and affiliations. However, these are not applicable in studying EER retention because many scholars, even the most active players, need not leave their current jobs to be affiliated with a program or school that has a clear “engineering education” label. In other words, scholars’ decision of opting in or out of EER usually requires no change in their positions.

2.2 Factors that influence ENE scholars’ research pathway

There is very little research attempting to find factors that affect ENE scholars’ pathways (please note we are not talking about students here. We are discussing only scholars of ENE or ENE researchers). The following attributes are selected in the present study based on literature and theories we draw upon from both ENE and other disciplines.

\textit{Level of engagement}, in this paper, refers to scholarly engagement – meaning, the number of publications produced by a scholar and the degree of contributions (indicated by authorship order) in each publication. In decision making, it is believed that sunk costs lead an individual to avoid loss and therefore tend to continue his/her investment\textsuperscript{16}. Here we do not argue about whether pursuing EER is a rational choice. However, it is reported\textsuperscript{3} that ENE scholars primarily concerned about time limitation and value of doing ENE research. Therefore, we validate whether more investment in time and effort, may lead to a longer stay in EER.

\textit{Collaboration}, in this paper, refers to scholarly collaboration with other ENE academics. According to critical contact theory\textsuperscript{17}, an individual’s career choice is influenced by his/her observation during interacting with personnel of the organization. The significance of collaboration is also recognized in engineering education. Jamieson\textsuperscript{4} appealed for more access to knowledgeable persons and Williams et al.\textsuperscript{3} revealed isolation as one of the difficulties that new ENE scholars faced.

\textit{Research area}, in this paper, means ENE scholars’ research interests based on their past publications. This describes the interdisciplinary aspect when a new scholar moves from his/her original domain to engineering education \textsuperscript{3}. Based on learning theory about knowledge transfer\textsuperscript{19} and adaptive expertise\textsuperscript{20}, an individual may have difficulties adapting past knowledge flexibly in
solving problems in a completely new domain because he/she does not have sufficient knowledge in such a transfer.

**Affiliation**, in this paper, is defined as the institution that a scholar is affiliated with. It has been shown\(^2\),\(^3\) that geographical and institutional boundaries contribute largely to differences in research topics, collaboration, level of activity, and recognition of EER. Therefore, the present study includes affiliation as a potential factor to measure difference in scholar retention.

**Funding**, in this paper, refers to the financial support received by an ENE scholar to carry out research projects in EER. In career choice theories\(^1\), funding is categorized as an influential objective factor in a rational decision-making process. In academia in general, funding is a major incentive for faculty to retain in an institution\(^6\),\(^2\). In EER specifically\(^5\), funding is also listed as the third main challenge encountered by scholars when entering ENE.

3. **Methodology**

In this paper, we perform a bibliometric analysis on ASEE and FIE publications and NSF awards in years 2000-2011 to study scholar retention in EER. First, we collect detailed publication data from Engineering Village\(^8\) and insert them into a database. Second, we develop a program to disambiguate author names and affiliation names so that different variations of the same entity are grouped to represent one. Third, a bibliographic analysis is performed to characterize the general status of scholar retention during 2000-2011. Fourth, new ENE scholars are recognized and divided into a group of opt-out researchers and a group of retained researchers based on their publication history. We compare these two groups in terms of their publication production, authorship order, collaboration, research areas, affiliations, and funding support.

![Figure 1. The workflow used to analyze ASEE and FIE papers and NSF awards to understand scholar retention](http://www.engineeringvillage.com)

3.1 **Data sampling and collection**

The ASEE and FIE conference proceedings papers are chosen because they are popular publication venues for both new and senior scholars. In future studies, we expect to add more data sources. Therefore, studying ASEE and FIE authors’ publication history is an approximate measure of ENE scholars’ overall research activeness regardless of their length of experience in EER. To demonstrate the broad scholarship coverage of ASEE and FIE, we compare authors in

\(^a\) Engineering Village: http://www.engineeringvillage.com
these two conference proceedings with those publishing in Journal of Engineering Education (JEE) and Advances in Engineering Education (AEE) during 2000-2011, as illustrated in Figure 2. Out of a total of 1,254 JEE authors, only 37% of them have never published in ASEE/FIE. If we exclude authors who have published only once (named single-paper authors in this paper), only 9% of JEE authors have not published in ASEE or FIE. Similarly, out of 215 AEE authors, 29% of them have no ASEE/FIE paper and it decreased to 2% when single-paper authors are excluded. The reason why authors with only one paper are excluded is because we do not consider scholars publishing one paper as a real attempt to enter EER. Including such occasional efforts by these authors introduces a significant amount of noise in data analyses. The fact that most of JEE and AEE authors have published in ASEE and FIE shows the validity of using ASEE and FIE authors to represent the ENE community. As an alternative, we may also include journal papers such as JEE and AEE in this study. However, journal papers usually have longer and varied review cycles. This makes time series analysis inconsistent and inaccurate if mixed with conference proceedings papers.

![Figure 2. Author coverage in ASEE and FIE against JEE and AEE over 2000-2011 when (a) all authors are considered and (b) only authors with two or more publications are included](image-url)

The publication metadata of ASEE and FIE conference proceedings papers are downloaded from Engineering Village. For each paper, the following attributes are available in the metadata: title, authors, author affiliations, terms, publication year, publication venue, reference counts, abstract, language, sponsors, and page count. The last four fields are irrelevant to this study and therefore not used in later steps. All other attributes are extracted from the metadata and inserted into a mySQL database. Terms are assigned to each paper by Compendex, an engineering literature database, based on a taxonomy built upon all its archives. Due to the problem of missing author-supplied keywords in many papers, we choose systematically assigned but still highly descriptive terms for the purpose of understanding the major topics of each paper. Some of these terms, such as engineering discipline, graduate students, and active learning, may be too broad to characterize precisely a paper’s major foci. However, manual assignment of accurate key terms for a large number of documents by domain experts is infeasible. It also leads to a very diverse range of topics that makes it almost impossible to recognize repeating patterns. Automatic key term extraction by algorithm inevitably introduces a significant amount of noise (irrelevant and meaningless terms). Therefore, this paper uses key terms generated based on a human-curated taxonomy by Compendex and allows a certain degree of abstraction in research topics in EER.
3.2 Name disambiguation

Name ambiguity problem refers to either one entity having two or more name variations or multiple entities share the same name representation\(^\text{24}\). For example, John Smith may have papers under names J. Smith, John Smith, and John K. Smith, whereas many individuals may share the same name Jane Doe. Failure to disambiguate names can lead to distortion in data analysis\(^\text{25}\). For example, we may mistakenly recognize an active researcher John Smith as quitting EER if all his subsequent papers are under a different name without appropriate disambiguation. Similar problems also occur in institution names, which may be abbreviated, misspelled, or include college/school names. To solve the name ambiguity problem, we first insert each author and affiliation name as a separate entry into the database. This means that all names are initially assumed different from each other. Then we develop a token-based disambiguation algorithm based on the well-known Levenshtein Distance\(^\text{26}\) to compute the similarity between names. Two author names are detected as duplicates if they meet all of the following conditions:

1. Two last names are identical;
2. If one name has an abbreviated first name, the first letter must be the same as the other name. Otherwise, two first names must be at least 80% similar. This means, for two first names of 5 to 9 letters in length, only 1 letter different is allowed; and
3. One author name has a middle name while the other one has not. If both have middle names, two middle initials must be identical.

Similarly, for institution names, the following rules are applied:

1. One name is a sub-string of the other one. For example, “School of Engineering Education, A University” and “A University”; and
2. If two institution names vary less than 3 letters in length, the two names must be at least 80% similar. For example, “A University, DC” and “A University - DC”

3.3 Characterizing the general retention status

Similar to the definition of student retention rate\(^\text{27}\), the timely retention rate of scholars is computed by:

\[
R(t_i) = \frac{\text{Num}(\text{Author}_{t_{i-1}} \cap \text{Author}_{t_i})}{\text{Num}(\text{Author}_{t_{i-1}})}
\]

Where \(t_i\) refers to a time period such as a year and \(\text{Author}_{t_i}\) refers to authors publishing in the time period \(t_i\). \(\text{Num}\) is the total number of authors. Therefore, the retention rate for the current period is the number of authors publishing papers both in the current and previous periods divided by the total number of authors publishing in the previous period. A scholar may sometimes skip one or two years and resume paper production afterwards. In such cases, it may not be appropriate to flag the scholar as opt-out because of the skipped years. Therefore, besides yearly retention rate, we also compute retention rate for every two, three, four, and six years. These periods divide exactly the 12 years of publications reviewed in the present study. Besides calculating retention rate, we also measure active period for each author. An author’s active period is defined as a
year range between the year of his/her first publication and that of the latest publication. This is a local measure particular to the data set in consideration.

3.4 Recognizing and comparing opt-out scholars and retained scholars

In this study, we define new scholars as ASEE or FIE authors who have the earliest publication in or after 2002. New here denotes that a scholar joined EER during 2000-2011 (once new to EER) instead of indicating that a scholar recently started to do research in ENE. According to this definition, those publishing in 2000 or 2001 will not counted as “new” because some of them may have publications prior to 2000, which are not covered in our study. Although this may also exclude scholars who actually began their EER career in 2000 or 2001, we choose to exclude all authors who published in 2000 or 2001 to ensure the integrity of the new scholar group.

Single-paper authors are ASEE or FIE authors who published only one paper during 2000-2011. In our analysis, single-paper authors are excluded because as mentioned earlier, we do not view scholars publishing one paper as a real attempt to enter EER.

Opt-out scholars are defined as new scholars who (1) have the latest publication in or before 2009; (2) have at least two papers; and (3) sustain active for two or three years. Based on the same reason of not knowing whether a scholar will continue to publish in the future, those publishing in 2010 or 2011 are not considered as opt-out scholars. One may argue that some of these opt-out scholars have never truly opted in because they only contributed to their ENE research projects to a very limited extent. However, distinguishing and measuring the exact contribution by each author among all papers analyzed in the present study are infeasible because of the cost of conducting a qualitative study of over 20,000 authors. The definition above guarantees a certain degree of engagement, which is used as an indicator of scholars opting in.

Accordingly, retained scholars are those new scholars who (1) have at least six publications and (2) sustain active for at least six years. We choose six years as a dividing line because it is long enough to be distinguished from the opt-out group, the usual time for earning tenure, longer than average time for earning a PhD degree, while short enough to keep a sufficient number of scholars for analysis.

The opt-out group and the retained group are compared in terms of the following measures: population size, number of papers produced, average authorship order, average number of co-authors per paper, research topics, affiliation, and NSF funding related to EER.

4. Results

4.1 Overview

A total of 16,439 ASEE papers and 4,770 FIE papers were analyzed over the period of 2000-2011. Out of the 21,209 papers, 22,998 unique authors have been identified. On average, an ASEE/FIE paper has 2.66 authors. There are also 1,092 unique organizations and 43,851 unique terms identified from this dataset.
4.2 General retention status

Out of 22,998 authors, 15,636 (67.99%) of them published only one paper over 2000-2011, whereas 56 (0.24%) authors produced at least one paper every year. The average active period of an author is 2.35 years (5.23 years if single-paper authors are excluded). It means that ENE scholars stayed on an average 5.23 years before they stopped publishing in the data outlets we considered. The number of authors against active periods is presented in Figure 3 (with all single-paper authors excluded). From the eighth year on, the number of scholars becomes stable, which means that scholars who have pursued ENE research for eight years will probably stay in EER in the long run. The reason why there are 319 scholars who sustained 12 years whereas only 56, as mentioned above, have at least one paper per year is because according to our definition of active period in Section 3.4, an author’s active period is determined by the earlier and latest publication years and does not depends on the number of publications and missing years in between.

![Figure 3. Number of authors with active periods from 2 years to 12 years.](image)

In terms of timely trends, Figure 4 shows the retention rate for every year, two years, three years, and four years. The yearly retention rate has two troughs in around 2003 and 2008 but on average, 43.6% of all authors publishing a FIE/ASEE paper in a year will continue to publish in the next year. A similar pattern is observed from the retention rate of every two years. In Figure 4(c) and 4(d), however, it becomes clear that the retention rate continued to drop from 2000 to 2011. The decreasing retention rate seems even more problematic if we combine it with the fact that the number of papers per year in FIE/ASEE was increasing during the same period. That means, as the two publication venues encouraged more work to be published, fewer scholars sustained their interests in EER. Also, as the year range becomes larger (from Figure 4(a) to 4(d)), the average retention rate gets higher: from 43.6%, 50.4%, to eventually the six-year retention rate of 2006-2011 against 2000-2005 as 57.9%. This can be explained by the observation that authors sustained activity for 5.23 years but did not constantly produce papers every year. For example, an author who has papers in 2001, 2004, and 2007 will be considered retained if a six-year range is chosen but viewed as an opt-out if authors are aggregated based on a smaller period. This shows a pattern that some scholars published one or two papers, quitted for years, and came back to publish again. The infrequent activities may indicate their low degree of commitment to EER. An additional study is needed to understand the cause of this pattern.
Retention rate (Yearly)

(a) Retention rate (yearly): Average 43.6%

Retention rate (Every 2 years)

(b) Retention rate (every 2 years): Average 50.4%

Retention rate (Every 3 years)

(c) Retention rate (every 3 years): Average 52.8%

Retention rate (Every 4 years)

(d) Retention rate (every 4 years): Average 54.4%

Figure 4. Timely retention rate of ASEE/FIE authors (a) every year, (b) every two years, (c) every three years, and (d) every four years.
4.3 Opt-out scholars vs. Retained scholars

Based on our definitions of different scholar groups in Section 3.4, we identify 2,917 opt-out scholars and 703 retained scholars. We compare their academic profiles for their early career in EER. During the first two years of their active period, the opt-out scholars produced on average 2.56 papers whereas the retained scholars have 3.25 papers. The average number of authors per paper is 3.45 for the opt-out group and 3.27 for the retained group. The average authorship order (1 being the first author) is 2.50 for the opt-out group and 2.22 for the retained group. We performed a pairwise t-test on each of the above comparisons and the results show a significant difference between the two scholar group ($p<.001$) in every case. These statistics show that during the first two years of a scholar’s EER career, the opt-out group tended to be less productive, collaborate with more scholars, but are less likely to initiate or significantly contribute to their research projects.

The opt-out scholars and the retained scholars share the same top seven research topics: Graduate students, Student learning, Undergraduate students, Engineering programs, Faculty members, and active learning. This shows the common research interests in engineering students in higher education. However, there are also topics that are more favored by one group. In general, the opt-out group put more efforts on broader topics such as Engineering disciplines, Learning environments, Teaching and learning, Design, and Engineering degree. In contrast, the retained group focused more on topics such as Technical literacy, Teaching strategy, Program management, Oral presentations, Laboratory exercises, Leadership skills, and MATLAB. This may imply that if new scholars start with topics that are more tangible and closer to their major disciplines, they will slowly develop more interests and decide to stay longer.

In terms of comparison by institution, we sort institutions by the number of scholars affiliated with them. Appendix A shows the most common institutions for the opt-out group, retained group, and both, along with number of authors affiliated for each institution. The majority of these institutions have a large number of retained and opt-out scholars. But there are also a few institutions that are more favored by one group or the other. There is also a large gap between the two groups in the percentage of scholars awarded by NSF for conducting EER-related research. We map authors in the two groups to PI’s/coPI’s who had been awarded by EHR or ENG-EEC in NSF. It is found that 44.81% of retained scholars had EHR or ENG-EEC awards during 2000-2011, as opposed to only 18.96% of opt-out scholars. This particular point on funding for core EER work through NSF provides one form of validation to show that how we grouped authors (as being opt-out vs. retained) is correct.

5. Implications

Our results present a general trend of decreasing retention rate over 2000-2011. There is a great need for changes in the EER community to sustain scholar diversity. This includes not only helping current scholars to continue their interests in EER (from 5.23 to 8 years) but also inviting more new scholars to join.

Our analysis of the opt-out and retained scholars further elaborates possible solutions to the decreasing retention rate problem. First, a previous survey of new scholars’ pathways to EER revealed that legitimacy is the most mentioned challenges encountered by new scholars. Our results also show low paper productivity and low authorship order for opt-out scholars in their
early EER career. Scholars who are uncertain about the real value of EER research often hesitate to commit efforts and eventually leave the community. This suggests a need for making explicit connections between new scholars’ work to the EER fundamental theories and concepts, helping them witness the impact of their work, and encouraging them to lead original research efforts. Although new scholars have more collaborators in their publications than average ASEE/FIE authors, it is essential to collaborate with more experienced and knowledgeable peers who can help them achieve the above goals.

Second, we also find that new scholars who study more specific topics in their early career in EER are more likely to stay longer than those starting with broad problems. The contextualized projects provide new scholars more opportunities to contribute and apply their domain knowledge. Such publications about their experiences and practices in EER should be valued by the community so that they feel their commitments are rewarded. As mentioned earlier, helping new scholars ground practices in EER fundamentals is necessary to let new scholars realize the value of their research, but starting with more specific projects get them more engaged and prepared in their long-term career.

Third, affiliation also seems to have a great effect on influencing retention. A further study is needed to understand exactly why scholars in some institutions tend to sustain their efforts much longer than others. Institutions that have set up a school/department/program specific for engineering education perform well in the number of publication production and scholars retained. But these institutions also have great potential to engage a great number of opt-out scholars.

Fourth, financial support from NSF may be another determining factor, as also stated in an earlier study. However, the correlation between funding and retention can be interpreted in two different ways. On the one hand, it may be because of the NSF grant that scholars are convinced and financially ready to carry on EER. On the other hand, the scholars may have conducted exceptional research in ENE for a long time that finally get acknowledged by NSF. Regardless, our findings confirm the importance of funding for new scholars to sustain in EER.

6. Limitations and future work

Our study has the following limitations. The coverage of publication venues and funding agencies selected in this study may affect the validity of our findings. However, we have shown by comparing our dataset to those in popular journals that ENE scholars included in the present study can approximately represent all members in the EER community. Although it is generally believed that NSF is the primary funding source for EER, we would like to include awards from other funding agencies such as Department of Education in future studies. This study measures scholars’ degree of engagement in EER by their number of EER papers but a ratio to their total number of publications may be a better indicator. Also, scholars’ publication histories only provide a partial view of their research pathway. Scholars’ shift of research interests happens rarely with clear evidence. Besides bibliographic data, there are other ways such as survey and interview to capture such changes and explore other possible factors. In fact, in other disciplines, scholars’ motivations to pursue a certain career path have been studied using survey and interview and factors such as learning-oriented, performance-oriented, and lifestyle have been revealed. These qualitative methods are so costly that it is infeasible for studying a large number of scholars. But such qualitative analyses can complement the present study by answering
questions such as when new scholars decide their career paths, whether educators perceive time and financial support differently from graduate students in the ENE department. In a future study, we plan to draw a small sample of scholars from the present study and conduct a more in-depth analysis to understand their pathway preferences.

7. Conclusion

In this paper, we perform a bibliometric analysis on ASEE and FIE papers and NSF awards from 2000 to 2011 to study the retention of ENE scholars. Based on 21,209 papers, we characterize the general retention status of 22,998 ENE scholars. Our results present the trend of decreasing retention rate from 2000 to 2011. We also find that during this period, new scholars stayed in EER for an average of 5.23 years. Then we recognize (1) opt-out scholars who discontinued efforts in EER quickly after they started and (2) retained scholars who sustained interest in EER topics for a longer time. We compare their academic profiles including level of engagement, collaboration, topics, affiliation, and funding status. During their early career in EER, scholars who preferred to opt out quickly tended to commit less and had fewer original contributions. They also worked on broader topics and received less funding support than the retained group. To the best of our knowledge, the present study is the first attempt to study scholar retention in large scale and also the first quantitative effort to measure scholar retention based on bibliometric data.

8. Acknowledgement

This project is supported through NSF-DUE 1123108.

Appendix A. Top 10 most common institutions for the opt-out scholars, retained scholars, and both (marked with asterisk), sorted by number of affiliated authors in descending order.

<table>
<thead>
<tr>
<th>Opt-out</th>
<th>Num. of affiliated authors</th>
<th>Retained</th>
<th>Num. of affiliated authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purdue University*</td>
<td>328</td>
<td>Purdue University*</td>
<td>373</td>
</tr>
<tr>
<td>Arizona State University*</td>
<td>155</td>
<td>Virginia Tech</td>
<td>217</td>
</tr>
<tr>
<td>Virginia Tech</td>
<td>140</td>
<td>Arizona State University*</td>
<td>188</td>
</tr>
<tr>
<td>Iowa State University*</td>
<td>132</td>
<td>Pennsylvania State University*</td>
<td>183</td>
</tr>
<tr>
<td>University of Leeds</td>
<td>107</td>
<td>Rowan University*</td>
<td>101</td>
</tr>
<tr>
<td>Rowan University</td>
<td>97</td>
<td>Texas A and M University</td>
<td>97</td>
</tr>
<tr>
<td>Pennsylvania State University*</td>
<td>91</td>
<td>Northeastern University</td>
<td>89</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>87</td>
<td>Purdue University, Calumet</td>
<td>85</td>
</tr>
<tr>
<td>Michigan Technological</td>
<td>81</td>
<td>Michigan Technological</td>
<td>83</td>
</tr>
<tr>
<td>University*</td>
<td></td>
<td>Iowa State University*</td>
<td>77</td>
</tr>
<tr>
<td>Georgia Institute of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References
