Published online: 3-15-2013

Building On and Honoring Forty Years of PBL Scholarship from Howard Barrows: A Scientometric, Large-Scale Data, and Visualization-based Analysis

Hanjun Xian

Purdue University, xianhj@gmail.com

Krishna Madhavan

Purdue University, cm@purdue.edu

IJPBL is Published in Open Access Format through the Generous Support of the [Teaching Academy at Purdue University](#), the [School of Education at Indiana University](#), and the [Jeannine Rainbolt College of Education at the University of Oklahoma](#).

Recommended Citation

Xian, H. , & Madhavan, K. (2013). Building On and Honoring Forty Years of PBL Scholarship from Howard Barrows: A Scientometric, Large-Scale Data, and Visualization-based Analysis. *Interdisciplinary Journal of Problem-Based Learning*, 7(1). Available at: <https://doi.org/10.7771/1541-5015.1325>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](#).

Building On and Honoring Forty Years of PBL Scholarship from Howard Barrows: A Scientometric Large-scale Data and Visualization-based Analysis

Hanjun Xian and Krishna Madhavan

Abstract

Over the past forty years, Howard Barrows' contributions to PBL research have influenced and guided educational research and practice in a diversity of domains. It is necessary to make visible to all PBL scholars what has been accomplished, what is perceived as significant, and what is the scope of applicability for Barrows' groundbreaking findings. As more disciplines recognize Barrows' efforts and adopt PBL in education, it becomes crucial but challenging to sustain community memory so that PBL scholars are kept well informed of research innovations in various domains. In this paper, we review Barrows' scholarly efforts in PBL and reveal the impacts on subsequent studies in various domains. A bibliometrics analysis is conducted on Barrows' PBL publications and the corresponding citations to quantitatively measure Barrows' impact. Our findings demonstrate Barrows' exceptional contributions to PBL and the disciplinary differences in conducting PBL studies based on Barrows' work. It is also revealed that PBL scholars who share similar interests have rarely collaborated with each other. The PBL research community has a real opportunity to connect isolated research groups and reduce the fragmentation so that research innovations in one domain can be disseminated to inform other scholars.

Keywords: problem-based learning, Howard Barrows, scientometric, data mining, citation analysis

Introduction

Since problem-based learning (PBL) was introduced in U.S. medical schools in the 1950s, it has gained increasing attention and has become arguably the most significant innovation in medical education (Boud & Feletti, 1998). It has also reached beyond its intended initial domain, into engineering and science education, in significant ways. Among the tremendous pioneering efforts, Howard S. Barrows has built the fundamentals that scaffold and inspire subsequent studies in PBL. His book *Problem-Based Learning: An Approach to Medical Education* (Barrows & Tamblyn, 1980) has over 2,000 citations according to Google Scholar and his other academic publications have also received wide recognition in the research community. During Barrows' career span, he sustained his commitment to PBL from his earliest academic publication in 1974 to the latest one in 2008. His early publications in 1970s and 1980s continue to be frequently discussed and cited in recent studies.

Given Barrows' key role in PBL, it is essential for the PBL research community to document and highlight his major contributions. A community-wide consensus about definitions, taxonomy, and other fundamental concepts needs to be reached in order to foster effective scholarly communication and collaboration. According to the theory of innovation diffusion, in the process of disseminating research ideas and work, prominent actors are more likely to create, adopt, and spread innovations in the community at an early stage (Liu, Madhavan, & Sudharshan, 2005). With Barrows being a central player in PBL, it is necessary to make visible to all PBL scholars what has been accomplished, what is perceived as significant by the community, and what is the scope of applicability for Barrows' groundbreaking findings.

Although most of Barrows' work on PBL was rooted in medical education, the fundamental concepts and methods he proposed have profound and long-lasting effects on educational research in other domains. For example, according to our analysis of the *Journal of Engineering Education* (one of the top journals in engineering education) over the years 2000–2009, 9 articles (4.8%) have PBL or its varieties (Barrows, 1986) either in their keywords, titles, or abstracts out of a total of 186 research articles. Among these 9 articles, 5 (55.6%) of them have cited Barrows' studies. Similarly, from a completely unexpected problem space, 4 of 11 (36.4%) PBL studies in *Journal of Geography in Higher Education* from 2000 to 2009 have cited Barrows' work. As the applicable scope of PBL expands and more disciplines start to adopt PBL in their educational practices, it becomes increasingly difficult to sustain community memory. It also becomes increasingly crucial to keep PBL scholars well informed of research innovations in various domains that use derivations of PBL.

As we witness the dissemination of Barrows' groundbreaking innovations, it remains unclear what exactly his academic profile looks like and how his publications influence subsequent PBL studies in various domains. For example, what are the major topics he

worked on? What are the various research strands inspired by or based on his studies? What is the nature and topology of the community of researchers evolving around Barrows' work? To answer these questions, we need to develop a comprehensive synthesis of Barrows' and his followers' research efforts in PBL. This requires a holistic large-scale analysis of publications and an effective representation of the results. In this paper, we use a combination of scientometrics, bibliometrics, large-scale data mining, and visualizations to not only pay tribute to Barrows' intellectual contributions, but also to visually showcase the large knowledge communities that have formed and evolved from and around his work.

Literature Review

Meta-analysis of PBL Literature

PBL has a long history. Some previous reviews of PBL studies have used meta-analysis to provide insights—albeit narrow from a large-scale data perspective—for a specific research topic. Albanese and Mitchell (1993) conducted a meta-analysis-like review that categorized about 100 earlier studies over 1972–1992 in medical education by outcome measures such as exams and graduates' perceptions. They then synthesized research findings in each category and demonstrated the effectiveness of PBL against traditional curricula. Other similar studies based on meta-analysis include reviewing effects of PBL on faculty attitudes, student mood, and class attendance (Vernon & Blake, 1993), accumulation of knowledge (Newman, 2003), and participants' satisfaction (Smits, Verbeek, & De Buissonje, 2002). These meta-analysis based reviews offer insightful understanding of the impacts of PBL. However, the scope of these reviews is limited to medical education. Also, the number of papers reviewed is relatively small (100 or less) due to the great effort and unwieldiness of conducting in-depth meta-analysis.

Examining PBL within a non-medical and multidisciplinary framework has recently become a popular research topic. For instance, instead of focusing on medical education, Dochy, Segers, Van den Bossche, and Gijbels (2003) and Gijbels, Dochy, Van den Bossche, and Segers (2005) reviewed empirical work in higher education performed in a real-life classroom setting. The former study revealed the effect of PBL on students' knowledge and skill, whereas the latter compared effect of PBL in three levels of the knowledge structure. With the inclusion of 47 outcomes from non-medical-education disciplines, Walker and Leary (2009)'s study reviewed 82 total papers in PBL and synthesized their assessment level, problem type, and PBL method. Strobel and van Barneveld (2009) used a qualitative meta-synthesis method to compare prior meta-analysis research findings including studies from other domains. These research efforts demonstrate their multidisciplinary

concerns. However, their criterion sampling strategy in searching and selecting literature often results in only a small set of documents. The insufficient coverage of literature limits the capability of showing how PBL is used in different contexts, which, as Ravitz identified (2009), is one of the grand challenges in PBL research.

Bibliometrics for Measuring Academic Performance and Impact

As meta-analysis studies aim to provide an in-depth and insightful review of a small set of highly relevant literature, bibliometrics offer a set of methods to quantitatively evaluate publications, scholars, journals, institutions, and indeed larger population steps such as nations based on large-scale publication metadata (Borgman & Furner, 2002). The most common bibliometrics method is citation analysis. Citations of academic articles have been considered as a measurement tool for various purposes. For example, Hicks, Tomizawa, Saitoh, and Kobayashi (2004) discussed the use of bibliometrics to evaluate research funded by the U.S. federal government. Based on a similar approach to evaluating research using bibliometrics, Campbell et al. (2010) examined projects funded by the National Cancer Institute of Canada. Focusing on both evaluation of research and the impact of governmental funding, Boyack and Börner (2003) analyzed publications in Behavioral and Social Science at the National Institute on Aging primarily measured by citation counts.

Besides being used as a measurement tool for evaluating research and governmental funding allocation, citation information has also been studied to evaluate the impact of journals and conferences. Garfield (1972) presented the results of citation analysis of publications in science and technology indexed by SCl, based on which journals were evaluated by frequency and impact of citations. Brown and Gardner (1985) focused specifically on journals and articles on Contemporary Accounting Research, whereas Sims and McGhee (2003) aimed to rate journals relevant to Ophthalmology. All these studies considered citations as a valid and powerful measurement tool or indicator to evaluate research, scholars, funding strategies, and journal impacts. Some other studies conducted co-citation analyses to identify prominent authors in a specific research area (McCain, 1990). By means of co-citation analyses, researchers characterized main research focuses and trends (Acedo & Casillas, 2005; Chen & Carr, 1999). However, some researchers questioned (Garfield, 1979; Kostoff, 1998) and even opposed the use of citation counts as an evaluation tool and appealed for consideration of the content and context of citations (MacRoberts & MacRoberts, 1996).

In this study, we aim to develop a broad understanding of Barrows' efforts and impacts. Evaluation of individual academic efforts can be quantified by bibliometrics measures such as *h*-index (Hirsch, 2005) and *m*-index (Bornmann, Mutz, Neuhaus, & Daniel, 2008). *h*-index is defined as follows (Hirsch, 2005): "A scientist has index *h* if *h* of

his or her Np papers have at least h citations each and the other $(Np - h)$ papers have $\leq h$ citations each". h -index has been shown (Hirsch, 2007) to have a better predictive power of a scholar's productivity and future achievement than number of citations, number of papers, and mean citations per paper. As a variant of h -index, m -index (Bornmann et al., 2008) is defined as "the median number of citations received by papers in the Hirsch core" and Hirsch core refers to "papers ranking smaller than or equal to h ". m -index is a stronger predictor of peer assessment of scientists than h -index (Bornmann et al., 2008) and therefore is also appropriate to showcase Barrows' contributions to PBL. According to Konur's (2012) recent analysis of 179,832 educational research manuscripts over the past 30 years, the top 20 most prolific authors publish between 72 and 172 papers, have citations per paper ranging from 0.69 to 26.39, and have h -index from 4 to 33

Also, instead of using meta-analysis to perform an in-depth content analysis of a small number of publications, we attempt to understand Barrows' efforts in their entirety across domains. This requires citation analysis among publication data from many different domains, requiring us to introduce a methodology based on large-scale data analysis. As far as we know, such a study that brings together data from multiple domains has never been attempted in the PBL community before, in large part due to complexity of integrating the necessary data together. Furthermore, the algorithmic advances needed to attempt this study are under-utilized in learning research.

Research Questions

In this paper, we review Barrows' scholarly contributions to PBL and the impacts in various domains. The research questions we aim to address in this study are:

1. What were Barrows' major research foci in PBL? How did he shift from one topic to another during his career span, representing intersecting impacts on the construction and propagation of PBL as a paradigm?
2. What is Barrows' scholarly impact measured by h -index (Hirsch, 2005) and m -index (Bornmann et al., 2008)?
3. Who was Barrows collaborating with directly?
4. How did a secondary community of researchers utilizing Barrows' work evolve?
5. What are the major research topics built upon Barrows' work? How have they changed over time?

The first three research questions focus on demonstrating Barrows' academic profile. The last two questions aim to show his broad impacts in all related disciplines that have PBL studies built upon his work.

Methods

The research questions of the present study are to demonstrate Barrows' academic profile in problem-based learning and his impact on research work from a wide range of domains. Figure 1 illustrates the workflow used for (I) collecting, sampling, and processing data and (II) analyzing data and representing results. In step 1, we collect 412 Barrows' publications from multiple resources such as ERIC, PubMed, and Web of Science. In step 2, among these articles, 37 PBL-related ones are selected including duplicates and relevance is determined by the occurrence of PBL in title, abstract, and

Figure 1. The workflow used in the present study to collect, sample, analyze data and represent results.

author-supplied keywords. In step 3, document titles are scanned and compared with each other automatically to remove duplicate documents. In step 4, the 26 resulted publications run through our author name disambiguation algorithm to make sure that the same researchers with name variations are counted only once. For articles without author-supplied keywords, we extract keyphrases based on the full text or abstracts of the documents.

Based on the cleaned data set, we perform a bibliographic analysis to demonstrate Barrows' academic profile, such as his *h*-index, *m*-index, major research foci, collaborators, and most cited articles in step 5. Next, in steps 6 and 7, we choose his top five most cited articles and acquire 987 citations using Google Scholar and Microsoft Academic Search to present a full comparative picture of the relative statistics. Again, author name disambiguation and keyphrase extraction are performed in step 8 prior to data analysis. Finally in step 9, the analysis results are represented in visual forms such as tag clouds and network visualizations to make interpretation of our findings easier.

Data Collection and Sampling

We search for PBL literature authored by Howard Barrows in multiple data sources. The following name variations are included in the search: Howard S. Barrows, Howard S Barrows, Howard Barrows, H. S. Barrows, H S Barrows, H. Barrows, and H Barrows. Each of these names is used to search ERIC, PubMed, Web of Science, Academic Medicine, Microsoft Academic Search, and Google Scholar. We further narrow the sample to PBL-related literature by selecting publications that contain problem-based or PBL in title, abstract, or author-supplied keyword. Only books, journal papers, and conference papers are analyzed in the present study. Data came in a variety of formats including BibTex, RIS, EndNote, and XML. We developed a parser for each of these formats to extract metadata fields such as title, abstract, authors, publications venues, and keywords and save them into a MySQL¹ relational database.

Barrows' publications in PBL have produced a broad impact on medical education and other disciplines. To track their exact influence on other researchers and subsequent studies, we analyzed papers that cite Barrows' five most cited publications. Citations of these five publications are acquired using the Microsoft Academic Search API. Using search engines such as Google Scholar (GS) and Microsoft Academic Search (MAS) has the advantage of broader literature coverage and lower cost while retaining a strong correlation to impact statistics with traditional publication indexing sites such as Web of Science and Journal Citation Report (Harzing & Van der Wal, 2008). In collecting the citation data, we have observed that GS in general has a more comprehensive coverage than MAS. However, we choose MAS because GS results do not have author-supplied keywords or abstracts,

1 MySQL—<http://www.mysql.com>

which are crucial to our topical analysis. We implemented a JSON-RPC client program to query the MAS database and save the returned result to our publication database.

Name Disambiguation, Citation Counts, and Keyphrase Extraction

In collecting Barrows' publications, the same publication may appear in the search result from more than one data source. Detecting duplicate documents is achieved by a token-based name disambiguation algorithm based on the Levenshtein Distance (Levenshtein, 1966). We apply this technique to compute similarity between document titles with a threshold of 0.75. It means that two titles must be at least 75% similar to be considered as duplicates. Once the threshold is reached, we treat them as duplicates if they are published in the same year with the same issue and volume numbers, if applicable.

The detected document duplicates are then merged into one. However, publication data obtained from different sources may be slightly different from each other. For example, one may spell out the journal name such as Medical Education, whereas another abbreviates it as Med Edu. One may contain volume and issue numbers while the other one does not. When duplicate documents are merged into one, for each data field, the longer, more descriptive values are always kept so as to make the collected data as complete as possible. Number of citations for each publication is computed using the following rule: If different data sources do not agree on the citation count, use the mean; If only one data source provides a citation count higher than zero, rely on this number when merging documents; and if all data sources suggest zero citation found, assume the number of citations to be zero. Measurement of citation counts can be used to calculate Barrows' *h*-index (Hirsch, 2005) and *m*-index (Bornmann et al., 2008).

Barrows' collaborators are defined in this study as his co-authors in PBL publications. Although authorship information is usually available from the data sources, overlooking the name ambiguity problem can cause severe errors in data analysis. Name ambiguity problem is a common problem in bibliographic databases where the same author may publish papers under different name variations and the same author name may refer to different individuals (Smalheiser & Torvik, 2009). For instance, one of Barrows' collaborators, Cindy Hmelo-Silver, has publications under names *Cindy Hmelo*, *Cindy E. Hmelo*, *C. E. Hmelo-Silver*, and *Cindy E. Hmelo-Silver*. Without disambiguating duplicate names, these four names will be considered as four individual researchers. As a result, this error yields an incorrect number of collaborators and leads to other imprecise statistics. We built a supervised name disambiguation system that computes similarity between names based on multiple attributes: full name, institution, past publication, and co-author. Our program automatically identifies potential duplicates and shows them as recommendations for disambiguation. Upon approval from the administrator, duplicate names are grouped to represent one entity from that point forward.

Key terms of a publication usually refer to author-supplied keywords. However, some papers do not contain author-supplied keywords and therefore, we must have a mechanism to generate key terms for these publications. We developed a keyphrase extraction algorithm based on GenEx (Turney, 2000) to identify the top 5 keyphrases for each document based on the full text. Here *keyphrase* is defined as a sequence of up to four words that have relatively high occurrences in a document but appear much less frequently in other documents. Therefore, compared to the most common word-frequency approaches, our keyphrasing algorithm not only filters common terms such as *is*, *of*, *the* but more importantly, it extracts phrases that are more descriptive and retain more contextual meaning than single words. For instance, a common word *learning* is not as specific as *learning environment* and *learning outcome* in describing a document's context. The rest of this paper will use keyphrases to represent both the author-supplied keywords and machine-generated keyphrases.

Bibliographic Analysis

The above search strategy, sampling criteria, and data cleaning process result in a total of 26 PBL publications authored by Barrows. For each publication, we collect metadata such as title, abstract, publication venue, year, authors, volume, issue as well as full text in PDF. Barrows' *h*-index and *m*-index are computed using the formula discussed in the literature review of this paper. Aggregation of the keyphrases helps characterize Barrows' major research foci and the topical evolution during his career span. The disambiguated result of documents and author names can precisely recognize Barrows' direct collaborators in PBL.

Based on the collected citations, the present study provides a statistical overview such as the number of unique authors, list of major research topics, and the collaboration network among all co-authors. The collected citations are then grouped into the following categories: medical education, engineering education, science education, general PBL, and others. Publications that fall within medical education and related fields are grouped and labeled as "medical education". An article must contain at least one of the following terms in its title, abstract, or keyphrase to be considered as being in this category: *medical education*, *clinical education*, *patient*, *psychological*, *nurse*, *dental education*, *pharm-*, and *medicine*. Similarly, we identify PBL publications in engineering education by searching using the term *engineer*. We choose *science*, *physics*, *chemistry*, and *math* for recognizing papers in science education. Papers that do not fall into any of the above categories but have mentioned *problem-based learning* or *PBL* in the title, abstract, or keyphrase are marked as general PBL. The rest of the articles are categorized as others. The above keyphrase lists for categorizing citations are derived from our automatic categorization algorithm based on training data. We randomly select 100 citations and manually categorize them into medical education, science education, engineering education, general PBL, and others.

The categorized citations then become training data and other citations are automatically categorized based on the probability of keyphrase occurrence in the training data in each category. For each of these disciplinary categories, we generate a detailed view of main research topics based on keyphrases of publications in the category. We also show the collaboration network in the disciplinary community to illustrate the current collaboration status among scholars who have cited Barrows' PBL work.

Results and Analyses

Overview of Barrows' Contributions

The analysis of Barrows' 26 publications demonstrates his constant efforts in PBL since 1974. Figure 2(a) shows the number of his publications in each decade over 1974–2008. In the 1970s and 1980s Barrows played a leading role in most of his research studies, whereas during 1990–2008 he co-authored more than before. Barrows has a total of 17 collaborators in his PBL publications and Figure 2(b) illustrates his active collaboration with other scholars in each decade. Barrows' collaborators in PBL publications are listed in Table 1.

Barrows' commitment to PBL research has received wide recognition from other scholars. This can be demonstrated by a total of 4,508 citations of his PBL publications. Among all Barrows' 26 PBL publications, he has an h -index of 21 and m -index of 105. Out of his 16 first-author publications, his h -index is 14 and m -index is 125. As introduced earlier, the top 20 scholars in educational research have citations per paper ranging from 0.69 to 26.39 and have h -indices from 4 to 33. This means that, although Barrows has fewer publications on educational research than the top 20 authors in Konur's (2012) study, Bar-

Figure 2. Based on Barrows' PBL publications over 1974–2008: (a) Number of publications and (b) Number of collaborators

Num. of co-authored papers with Barrows	Researchers
3	L. Distlehorst, P. Feltovich, T. Koschmann, A. Myers, R. Tamblyn
2	E. Dawson, C. Hmelo-Silver, R. Robbs, R. Williams
1	S. Crooks, A. Kelson, D. Mitchell, E. Moticka, V. Neufeld, T. Cameron, R. Coulson

Table 1. Barrows' collaborators in PBL publications

Title	Authors, Published year
Problem-based learning: An approach to medical education	(Barrows & Tamblyn, 1980)
A taxonomy of problem-based learning-methods	(Barrows, 1986)
How to design a problem-based curriculum for the pre-clinical years	(Barrows, 1985)
Problem-based learning in medicine and beyond: A brief overview	(Barrows, 1996)
The 'McMaster Philosophy': An approach to medical education	(Neufeld & Barrows, 1974)

Table 2. Barrows' top five most cited publications in PBL.

Barrows enjoys relatively high citation counts and *h*-index. This demonstrates his remarkable impact on the scholarly community and his constant production of high quality work that was heavily cited in secondary literature. Table 2 shows Barrows' top five publications in terms of citation counts.

During Barrows' academic career in PBL, he primarily focused on the following topics: *higher education, medical education, problem solving, teaching methods, simulation, patients, medical schools, and educational strategies*. Keyphrases in Barrows' publications are grouped into the following categories: general, medical-related, teaching & instruction, professional development, assessment, curriculum, active learning, collaborative learning, and others. Figure 3 shows Barrows' primary research interests and specific topics within each category. The circle size indicates the total number of occurrences of keyphrases in Barrows' publications. The keyphrases within a category are sorted from the most frequently mentioned keyphrases to the least.

Research Studies Built upon Barrows' Work

Overview

The 987 publications that cite Barrows' PBL efforts are authored by 1,926 unique authors. As presented in Figure 4(a)(b), the number of authors and articles that cite Barrows' top five PBL publications has increased exponentially from 1970s to 2000s. Although all

Figure 3. Barrows' research foci and main topics **Figure 4.** An increasing number of (a) authors and (b) academic articles cite Barrows' top five PBL publications

five of Barrows' classic articles in Table 2 were published before 2000, they continue to gain wider recognition in academia. Although an increasing number of scholars conduct studies based on Barrows work, these scholars have rarely collaborated with each other. Figure 5 shows the fragmented collaboration network among authors who cite Barrows' work. The largest network is composed of only 43 authors, which is only 2.2% of the total

number of authors. Also, the largest network relies heavily on only a few key researchers (cut-points), without which the network will further fall apart. There are 636 groups of researchers working in isolation and an author has only 3.08 collaborators on average. Essentially, here we provide an indication of why various modifications and adaptations to Barrows' original work are not gaining mainstream usage and propagation. This type of fragmentation is characteristic of many educational innovations and perhaps is the basis for why educational innovations do not diffuse very quickly and easily.

Figure 6 shows the main research topics based on all these publications. Unlike regular Wordle (Viégas, Wattenberg, & Feinberg, 2009) representations that visualize single words and their frequencies, this tag cloud takes each *keyphrase* as a whole and the font size indicates how many articles have used a specific keyphrase. In essence, our approach is statistically, analytically, and algorithmically significantly superior to the simple word clouds. We choose the commonly used word cloud type of presentation for interpretability. However, the algorithmically identified keyphrases are a key component.

The top keyphrase *problem based learning* is removed from the tag cloud to increase the readability of less significant phrases. Besides problem based learning, the top ten topics are *problem solving*, *medical education*, *student learning*, *case study*, *learning environment*, *teaching and learning*, *medical students*, *self directed learning*, *active learning*, and *collaborative learning*. These keyphrases are generally consistent with Barrows' publications in Figure 2 such as collaborative learning, problem solving, higher education, and self

Figure 5. The co-author network among authors that cite Barrows' top five PBL publications

Figure 9. The co-author network among authors studying PBL in the context of medical education

Figure 10. Main research topics in science education built upon Barrows' top five PBL publications

Citations in Engineering Education

In engineering education, the top ten keyphrases in Figure 12 (*problem based learning*, *problem solving* and *engineering education* removed) are: *software engineering*, *case study*, *student learning*, *teaching and learning*, *software development*, *collaborative learning*, *teaching methods*, *active learning*, *biomedical engineering*, and *engineering design*. Based on

- volume 7, no. 1 (Spring 2013)

Figure 11. The co-author network among authors studying PBL in the context of science education

Figure 12. Main research topics in engineering education built upon Barrows' top five PBL publications

this topical analysis, PBL is more likely to be adopted by software engineering, electrical engineering, and biomedical engineering. The emphasis on design education echoes the conclusion of a prior study that reviewed PBL in engineering education (Mills & Treagust, 2003). The collaboration network for the engineering education community in terms of PBL research is similar to the one in science education. A relatively large network begins to emerge but in general, the community remains largely fragmented.

Figure 13. The co-author network among authors studying PBL in medical education **Figure 14.** Main research topics in the general PBL category built upon Barrows' top five PBL publications

Citations in General PBL

In the general PBL category, the top ten keyphrases in Figure 14 (*problem based learning* and *problem solving* removed) are: *self directed learning*, *student learning*, *higher education*, *learning environment*, *case study*, *learning experience*, *teaching and learning*, *collaborative learning*, *active learning*, and *instructional design*. Keyphrases in this category resemble the overall topical pattern. Authors in this group are the least collaborative—perhaps because of the diversity of the overall problem space represented by these authors.

Other PBL Citations

In the other PBL category, the top ten keyphrases in Figure 16 (*problem based learning* and *problem solving* removed) are: *learning environment*, *case study*, *instructional*

Figure 15. The co-author network among authors studying PBL without disciplinary orientation

Figure 16. Main research topics in the “other PBL” category built upon Barrows’ top five PBL publications

design, collaborative learning, student learning, teaching and learning, higher education, active learning, learning activities, and learning resource. Instructional design, learning resource, and online learning have gained more interest among scholars in this

Figure 17. The co-author network among other authors studying PBL

community than average. Again, authors have worked in isolation with very limited degree of collaboration with each other.

Discussion

In 1996, Barrows described his observation that PBL had been disseminated to a large number of medical schools, which developed their own variation of PBL in the curricula (Barrows, 1996). The above result further reveals the dissemination of PBL to other disciplines such as software engineering and biomedical engineering. Such a broader applicable scope and greater varieties in adapting PBL inevitably incur two problems (Barrows, 1996). First, although many schools claim to use PBL, many are poorly designed and not carried out according to the core PBL model (Barrows, 2003). Second, such imprecise implementation of PBL makes teachers believe that PBL only offers very limited advantage over what they already have. More efforts are needed to understand whether PBL is defined and implemented correctly in different disciplinary contexts and how such practices are evaluated.

Other than the concerns recognized by Barrows, dissemination of PBL to a broader range of disciplines poses new challenges to the PBL community. The network topology shown in the last section demonstrates the lack of collaboration among core contributors in medical education, among scholars in other disciplines, and between core researchers and medical-unrelated practitioners. According to the adoption process in the Diffusion

of Innovation theory (Rogers, 2003), for scholars and educators in non-medical disciplines, adoption of PBL requires not only their first exposure to PBL core concepts and taxonomy but also the perceived advantages of implementing PBL in their own contexts. Disciplines that have started to adopt PBL such as software engineering can be used as a case study to understand the motivation and process of a discipline adopting PBL. Performing such a case study can guide the dissemination of PBL to other potential disciplines with similar characteristics. For those that already decide to adopt PBL, the challenge becomes how to correctly tailor and implement PBL and finally report their practices and experience back to the PBL research community to inform future studies. For example, one may compare the definition, theoretical framework, and implementation in these practices with the fundamental PBL studies to evaluate the appropriateness of these PBL adaptations. It is also essential for core PBL researchers to publish papers in a language that can be easily understood by non-PBL scholars. As shown in the present study, Barrows' publications are frequently cited by scholars outside medical education and therefore, they are good examples of how to communicate ideas with a larger audience.

Conclusion

This paper demonstrates Howard Barrows' contributions to PBL research and his scholarly impact on a broad range of disciplines. A scientometric analysis is conducted based on Barrows' PBL publications and the citations of his five highest cited publications. We apply data mining techniques to collect bibliographic data, disambiguate names, and assign keyphrases to documents without author-supplied keywords. Our results show Barrows' pioneering and leading role in the PBL community. Clearly, we demonstrate the growing importance of his work in multiple branches and sub-spaces of education research. His academic work has also produced a significant impact on non-medical education domains such as science and engineering education. This paper recognizes topics that are almost evenly popular across disciplines such as critical thinking and problem solving. We also identify the differences in research focus across different disciplines. It is further revealed that scholars who share similar interests have rarely collaborated with each other. The PBL research community has a real opportunity to connect isolated research groups and reduce the fragmentation so that research innovations in one domain can be disseminated to inform other scholars.

The methodology introduced in this paper involves integration of bibliographic data from across multiple domains, application of data mining techniques in filtering and assuring data quality, and use of diverse visualizations for presenting the results. In future work, we plan to follow the same workflow to reveal the general picture based on all PBL studies. Also, we are going to perform a content analysis (Spiegel-Rosing, 1977) on PBL papers to not only identify which paper gets cited, but also reveal why and in what context

it is cited. The content analysis result can provide more details about the rationales and extent of PBL adoption in various domains. The methods presented in this paper have broader applicability and can be adapted for use in other areas of education research. For example, we can demonstrate how collaborative learning was first formed, evolved, and adopted in various disciplines. We can also show the scholarly impact of Vygotsky and his “Zone of proximal development” on education research.

Acknowledgments

This project is supported through National Science Foundation Grant TUES-1123108.

References

- Acedo, F. J., & Casillas, J. C. (2005). Current paradigms in the international management field: An author co-citation analysis. *International Business Review*, 14(5), 619–639. <http://dx.doi.org/10.1016/j.ibusrev.2005.05.003>
- Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68(1), 52.
- Barrows, H. S. (2003). Response to “the problem with problem-based medical education: Promises not kept” by R. H. Glew. *Biochemistry and Molecular Biology Education*, 31(4), 255–256. <http://dx.doi.org/10.1002/bmb.2003.494031040269>
- Barrows, H. S. (1985). *How to design a problem-based curriculum for the preclinical years*. New York: Springer.
- Barrows, H. S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, 20(6), 481–486. <http://dx.doi.org/10.1111/j.1365-2923.1986.tb01386.x>
- Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*. <http://dx.doi.org/10.1002/tl.37219966804>
- Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education* (Vol. 1). New York: Springer.
- Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36, 3–72.
- Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8(1), 93–102. <http://dx.doi.org/10.3354/esep000084>
- Boud, D., & Feletti, G. (1998). *The challenge of problem-based learning*. London: Kogan Page.
- Boyack, K. W., & Börner, K. (2003). Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology*, 54(5), 447–461. <http://dx.doi.org/10.1002/asi.10230>

- Brown, L. D., & Gardner, J. C. (1985). Using citation analysis to assess the impact of journals and articles on contemporary accounting research (CAR). *Journal of Accounting Research*, 23(1), 84–109. <http://dx.doi.org/10.2307/2490908>
- Campbell, D., Picard-Aitken, M., Côté, G., Caruso, J., Valentim, R., Edmonds, S., Williams, G. T., et al. (2010). Bibliometrics as a performance measurement tool for research evaluation: The case of research funded by the National Cancer Institute of Canada. *American Journal of Evaluation*, 31(1), 66–83. <http://dx.doi.org/10.1177/1098214009354774>
- Chen, C., & Carr, L. A. (1999). Trailblazing the literature of hypertext. *Proceedings of the tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots returning to our diverse roots - HYPERTEXT '99* (pp. 51–60). Presented at the the tenth ACM Conference, Darmstadt, Germany. <http://dx.doi.org/10.1145/294469.294486>
- Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13(5), 533–568. [http://dx.doi.org/10.1016/S0959-4752\(02\)00025-7](http://dx.doi.org/10.1016/S0959-4752(02)00025-7)
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science*, 178(4060), 471–479. <http://dx.doi.org/10.1126/science.178.4060.471>
- Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1(4), 359–375. <http://dx.doi.org/10.1007/BF02019306>
- Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, 75(1), 27. <http://dx.doi.org/10.3102/00346543075001027>
- Harzing, A.-W. K., & Van der Wal, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*, 8(1), 61–73. <http://dx.doi.org/10.3354/esep00076>
- Hicks, D., Tomizawa, H., Saitoh, Y., & Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation*, 13, 76–86(11). <http://dx.doi.org/10.3152/147154404781776446>
- 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. <http://dx.doi.org/10.1073/pnas.0507655102>
- Hirsch, J. E. (2007). Does the *h* index have predictive power? *Proceedings of the National Academy of Sciences*, 104(49), 19193–19198. <http://dx.doi.org/10.1073/pnas.0707962104>
- Konur, O. (2012). The evaluation of the global research on the education: A scientometric approach. *Procedia - Social and Behavioral Sciences*, 47, 1363–1367. <http://dx.doi.org/10.1016/j.sbspro.2012.06.827>
- Kostoff, R. (1998). The use and misuse of citation analysis in research evaluation. *Scientometrics*, 43(1), 27–43. <http://dx.doi.org/10.1007/BF02458392>
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10, 707.
- Liu, B. S.-C., Madhavan, R., & Sudharshan, D. (2005). DiffuNET: The impact of network structure on diffusion of innovation. *European Journal of Innovation Management*, 8(2), 240–262. <http://dx.doi.org/10.1108/14601060510594701>

- MacRoberts, M. H., & MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics*, 36(3), 435–444. <http://dx.doi.org/10.1007/BF02129604>
- McCain, K. W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science*, 41(6), 433–443. [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<433::AID-ASI11>3.0.CO;2-Q](http://dx.doi.org/10.1002/(SICI)1097-4571(199009)41:6<433::AID-ASI11>3.0.CO;2-Q)
- Mills, J. E., & Treagust, D. F. (2003). Engineering education: Is problem-based or project-based learning the answer? *Australasian Journal of Engineering Education*, 3, 2–16.
- Neufeld, V. R., & Barrows, H. S. (1974). The “McMaster Philosophy”: an approach to medical education. *Academic Medicine*, 49(11), 1040. <http://dx.doi.org/10.1097/00001888-197411000-00004>
- Newman, M. (2003). A pilot systematic review and meta-analysis on the effectiveness of problem-based learning. *On behalf of the Campbell Collaboration Systematic Review Group on the Effectiveness of Problem-based Learning*. Newcastle upon Tyne, UK: Learning and Teaching Support Network-01, University of Newcastle upon Tyne.
- Ravitz, J. (2009). Introduction: Summarizing findings and looking ahead to a new generation of PBL research. *Interdisciplinary Journal of Problem-based Learning*, 3(1), 2. <http://dx.doi.org/10.7771/1541-5015.1088>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.
- Sims, J. L., & McGhee, C. N. (2003). Citation analysis and journal impact factors in ophthalmology and vision science journals. *Clinical & Experimental Ophthalmology*, 31(1), 14–22. <http://dx.doi.org/10.1046/j.1442-9071.2003.00610.x>
- Smalheiser, N. R., & Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology*, 43(1), 1–43. <http://dx.doi.org/10.1002/aris.2009.1440430113>
- Smits, P. B. A., Verbeek, J. H. A. M., & De Buissonje, C. D. (2002). Problem based learning in continuing medical education: A review of controlled evaluation studies. *Bmj*, 324(7330), 153. <http://dx.doi.org/10.1136/bmj.324.7330.153>
- Spiegel-Rosing, I. (1977). Science studies: Bibliometric and content analysis. *Social Studies of Science*, 97–113. <http://dx.doi.org/10.1177/030631277700700111>
- Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-based Learning*, 3(1), 4. <http://dx.doi.org/10.7771/1541-5015.1046>
- Turney, P. D. (2000). Learning algorithms for keyphrase extraction. *Information Retrieval*, 2(4), 303–336. <http://dx.doi.org/10.1023/A:1009976227802>
- Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic medicine*, 68(7), 550. <http://dx.doi.org/10.1097/00001888-199307000-00015>
- Viégas, F. B., Wattenberg, M., & Feinberg, J. (2009). Participatory visualization with wordle. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6), 1137–1144. <http://dx.doi.org/10.1109/TVCG.2009.171>
- Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-based Learning*, 3(1), 6. <http://dx.doi.org/10.7771/1541-5015.1061>

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. Correspondence concerning this article should be addressed to Hanjun Xian at the School of Engineering Education, Purdue University, Rm. 457, 207 S. Martin Jischke Dr., #203, West Lafayette, Indiana 47907; Email: hxian@purdue.edu

Dr. Krishna Madhavan is an assistant professor in the School of Engineering Education at Purdue University. He is the principal investigator for the National Science Foundation TUES Central Resource Project entitled *Deep Insights Anytime, Anywhere* (DIA2). Dr. Madhavan is also co-principal investigator and education director for nanoHUB.org. Prior to his arrival at Purdue, he was an assistant professor with a joint appointment in the School of Computing and the Department of Engineering and Science Education at Clemson University. Dr. Madhavan also served as a research scientist at the Rosen Center for Advanced Computing, Information Technology at Purdue University. He specializes in the development and deployment of large-scale data and visualization based platforms for enabling personalized learning and portfolio impacts analyses. 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 at the Frontiers in Engineering Education symposium.