Evaluation and Program Planning 58 (2016) 20–27

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

The historical path of evaluation as reflected in the content of *Evaluation and Program Planning*

Abu H. Ayob[a,b,*], Jonathan A. Morell[c]

[a] Faculty of Economics and Management, National University of Malaysia, UKM Bangi, 43600 Selangor, Malaysia
[b] Department of Strategic Management and Globalization, Copenhagen Business School, Kilevej 14, 2000 Frederiksberg, Denmark
[c] Syntek Corporation, 4.669 Evaluation and Planning, 2279 Mershon Drive, Ann Arbor, 48103 MI, United States

ARTICLE INFO

Article history:
Received 13 April 2016
Accepted 8 May 2016
Available online 19 May 2016

Keywords:
Evaluation and program planning
Evaluation research
Citation analysis
Article attributes

ABSTRACT

This paper examines the intellectual structure of evaluation by means of citation analysis. By using various article attributes and citation counts in Google Scholar and (Social) Science Citation Index Web of Science, we analyze all articles published in *Evaluation and Program Planning* from 2000 until 2012. We identify and discuss the characteristics and development of the field as reflected in the history of those citations.

1. Introduction

Evaluation is an inter-disciplinary applied study that involves a process of systematic acquisition and assessment of information in a manner that combines research methodology with sensitivity to political context and the points of view of multiple stakeholders (Trochim & Donnelly, 2001). Evaluation serves as a means to review programs in terms of value, criteria and standards by explaining how the program was implemented, how it operated, what it accomplished, and what would be needed to improve it (Williams, 2000). Or, as another framework would have it, evaluation is concerned with merit, worth and value (Scriven, 1991). By whatever categorization, evaluation activities are diverse, but have historically dealt with cost analysis, process evaluation, performance measurement, impact/outcomes assessment and organizational effectiveness. Initially, evaluation was not a core research area in the social sciences, but has since emerged as a body of knowledge since the 1960s. By now, evaluation inquiry has taken on the legitimacy of other mainstream social science fields (Stern, 2005). That legitimacy has resulted in knowledge that provides policy and program planners with increased explanatory insight to support understanding and practical decision making (Powell, 2006).

Evaluation and Program Planning (EPP), a leading journal in the field, published its first issue in January 1978. According to Social Science Citation Index 2014, it is ranked highly among other comparable journals of *American Journal of Evaluation*, *Evaluation*, *Evaluation Review* and *Zeithschrift Fur Evaluation.*[1]

The main objective of EPP is to assist evaluators and planners to improve their professional practices, to develop skills, and to advance the evaluation knowledge base. For over three decades, EPP has been publishing high-quality research that employs sophisticated techniques and methods of evaluation and planning across different fields such as organizational behavior, human resource development, public health, social services and education. A premise of the journal is that evaluation is an eclectic field, and that evaluators working in different domains can learn from each other.

The objective of this paper is to systematically analyze the path of evaluation since the beginning of the 21st century, as reflected in the contents of EPP. In particular, this study examines the impact of publication in EPP using citation analysis on specific attributes of each article published from 2000 until 2012. Citation analysis has become a widely-accepted approach for many purposes: institution's strategy (Steward & Cotton, 2013), academic promotion (Marsh & Hunt, 2006) and financial rewards for faculty members

* Corresponding author at: Faculty of Economics and Management, National University of Malaysia, UKM Bangi, 43600 Selangor, Malaysia.
E-mail addresses: abuhanifah.ayob@ukm.edu.my, aha.smg@cbs.dk (A.H. Ayob), jamorell@jamorell.com (J.A. Morell).

[1] This statement is based on the 2014 Journal Citation Reports Social Science Edition. The 5-Year Impact Factor for EPP is the second highest = 1.394. The highest is *American Journal of Evaluation* = 2.358. The third is *Evaluation Review* = 1.101, followed by *Zeithschrift Fur Evaluation* = 0.250. *Evaluation* = NA.

http://dx.doi.org/10.1016/j.evalprogplan.2016.05.006
0149-7189/

Table 1
Total articles published by year.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Total articles	41	39	45	41	38	39	45	38	42	41	71	42	63

(Nkomo, 2009). Scholars have used citation analysis to explore the intellectual structure of specific disciplines, e.g. Entrepreneurship (Steward & Cotton, 2013), Information Science (White & McCain, 1998), Marketing (Baumgartner & Pieters, 2003) and Strategic Management (Nerur, Rasheed, & Natarajan, 2008). Likewise, in this study, we examine the evolution and contribution of evaluation research in two dimensions: attributes of published articles and number of citations.

2. Overview of data

We first constructed a database consisting of attributes of each article published in EPP from 2000 until 2012.[2] We only included original research articles and excluded other manuscripts such as commentary and book reviews. Although we could have extended the data until 2015, we decided not to include them because in general, it would take several years for any article to be cited after it has been published. As expected, we found no citation for most articles published in 2013 and 2014. In total, we retrieved 585 articles in both regular and special issues for the period. Table 1 shows the total number of articles published by year, the average being 41 per year. Years 2010 and 2012 were exceptions. They contained 71 and 63 articles, respectively. Details on attributes of the articles will be discussed in the next section.

We searched for the number of citations for each article through two main bibliometric sources: Google Scholar (GS) and Thompson ISI Web of Science (WoS). Although both sources are commonly used for citation analysis, there is a significant difference between them. The former data includes more comprehensive citation coverage of documents: book chapters, books, conference proceedings, and non-US journals. The latter source focuses only on a limited number of journals, mainly English language titles from North America/Western Europe, and disregards other types of documents. As a result, the number of citations in GS is consistently higher than in WoS (Harzing & Van der Wal, 2007; Meho, 2006).

3. Citation analysis

In this section we introduce the citation metrics that we will discuss in detail in the following section. Table 2 shows number of articles with total citation figures in GS and WoS. Most articles have been cited between 1 and 20 times in both sources. There is also a significant number of articles that received between 21 and 40 citations in GS. Also, the data show that nine articles have never been cited in GS and 48 articles that were never cited in WoS; most of the uncited articles are recent publications (2011 and 2012).

Fourteen articles reached more than 100 citations in GS. Four reached 200 and above. Also, two articles received over 100 citations in WoS. These articles shared common traits: old publication (between year 2000 and 2004), research was conducted by practitioners, and they employed a mixed methodology with a heavy use of qualitative data.

Exceptional articles with highest total citations are:

1) Building capacity and sustainable prevention innovations: A sustainability planning model, by Johnson, K., Hays, C., Center, H. & Daley, C. 2004 (GS: 230, WoS: 107).
2) An experimental evaluation of recovery management checkups (RMC) for people with chronic substance use disorders by Dennis, M., Scott, C. K. & Funk, R. 2003 (GS: 184, WoS: 117).
3) A critical analysis for evaluation practice: the Kirkpatrick model and the principle of beneficence by Bates, R. 2004 (GS: 222, WoS: 36).
4) Making public health last: conceptualizing sustainability by Pluye, P., Potvin, L. & Denis J. L. 2004 (GS: 202, WoS: 78).
5) Assessing and improving partnership relationships and outcomes: a proposed framework by Brinkerhoff, J. M. 2002 (GS: 202, WoS: 48).
6) Wilderness challenge programs for delinquent youth: A *meta*-analysis of outcome evaluations by Wilson, S. J. & Lipsey, M. W. 2000 (GS: 186, WoS: 34).

The next analysis determined the overall trend of citations by time since publication. Table 3 shows the results of this analysis. Several findings deserve attention. First, average citations per article dropped significantly from 2000 to 2012, as shown in Fig. 1. In 2000, each article averaged 38 citations in GS and 14 citations in WoS. These numbers dropped steadily, reaching six and two citations respectively in 2012. There is nothing surprising in the finding that citations increase over time, but this analysis provides a specific quantitative estimate for the trend.

However, these are raw numbers and do not show the rate of change in citations per article per year. To determine that rate, we normalized change in citation counts over time by years since publication, using the formula:

N = Average per year/$(2013 - x)$; where x is publication years of 2000 until 2012.

This formula is based on the assumption that the total number of citations will increase proportionately with the number of years since publication. The data are presented in numerical form in Table 3, and in graphic form in Fig. 2.

4. Article attributes

In this section, we present characteristics of each of the article's attributes and their citation analysis.

4.1. Regular or special issues

In addition to the regular issues of EPP, a few volumes are allocated for special issues every year. Each special issue is dedicated to articles on a specific theme. The reason for publishing special issues is to present studies on controversial, emerging or under-researched topics. Special issues are produced by editors who are chosen both for their unique expertise in the topic and for their knowledge of where to find the best work. Thus, we might expect special issues to show a particularly high number of citations. On the other hand, the topics are often specialized and narrow, a fact that might result in a fewer number of citations.

Table 4 shows the total articles published in regular and special issues collapsed over all years. *T*-test analysis shows that articles in regular issues are cited significantly more frequently than articles in special issues. For GS, on average, regular issues have been cited 24.5 times compared to 18.05 in special issues ($p = 0.016$). There are

[2] The data is correct as of 31st January 2015.

Table 2
Number of articles with total citations.

Total citations	0	1–20	21–40	41–60	61–80	81–100	100–199	200 and above
GS	9	377	110	43	21	10	10	4
WoS	48	481	44	8	2	0	2	0

Table 3
Average citations: Per year, and Normalized.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Total articles	41	39	45	41	38	39	45	38	42	41	71	42	63
Total GS	1562	1003	1458	1184	1328	952	922	1028	782	816	1201	597	411
Total WoS	585	317	471	522	495	393	327	400	298	286	412	212	127
Average GS	38.10	25.72	32.4	28.89	34.95	24.41	20.49	27.05	18.62	19.90	16.92	14.21	6.52
Average WoS	14.27	8.13	10.47	12.73	13.03	10.08	7.27	10.53	7.10	6.98	5.80	5.05	2.02
–													
N GS	2.93	2.14	2.95	2.89	3.88	3.05	2.93	4.51	3.72	4.98	5.64	7.11	6.52
N WoS	1.10	0.68	0.95	1.27	1.45	1.26	1.04	1.76	1.42	1.75	1.93	2.53	2.02

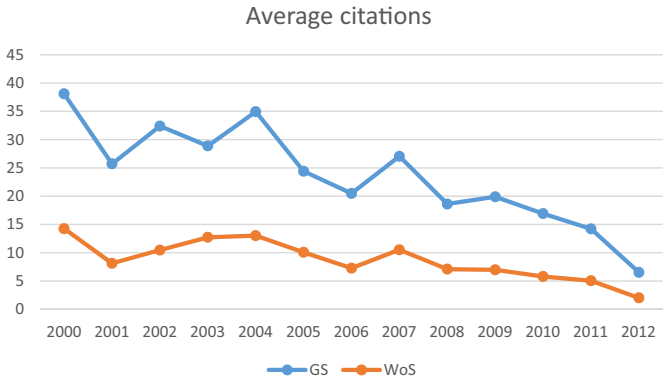

Fig. 1. Average citations per article by year.

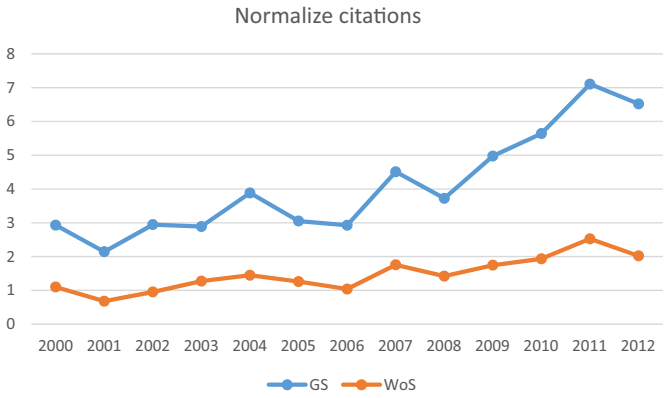

Fig. 2. Normalize citations by year.

similar findings for WoS, regular issues have been cited 9.25 times compared to 5.89 in special issues ($p = 0.001$).

The special issues with the highest number of citations are:

1) HIV outreach and substance abuse treatment for Latino drug users: implications for program planning. 2008, Volume 31, No 1.
2) Ethics, evaluation and for-profit corporations. 2004, Volume 27, No. 3.
3) Research on implementing evidence based practices in community based addiction treatment programs: policy and program implications. 2011, Volume 34, No. 4.

4.2. Authors: academics or/and practitioners

Evaluation is a field where academically based professionals and non-academic practitioners do work that speaks to both scholarly and practical applications. Authorship of articles in EPP reflects this situation.

Table 5 and Fig. 3 show the percentage of articles authored solely by academics, solely by practitioners, and by academic-practitioner collaborations by year, as determined by the stated affiliations of authors. While academics are the largest group over the years, almost half of the articles are authored by either sole practitioner or by academic-practitioners collaborations. Assuming that authorship is an indicator of public interest, these data indicate that EPP is relevant for both academic scholars and to those who do evaluation outside of academics.

While representation of authorship shows a joint academic/practitioner orientation, citation analysis shows that articles authored solely by academics are cited significantly more frequently than articles written either solely by practitioners or by practitioner/academic collaborations. (ANOVA comparing the three groups yields $p = 0.047$). On average, papers authored only by academics have 25.37 GS citations, compared to 20.39 GS citations for practitioners and 18.64 GS citations for practitioner/academic collaborations.

4.3. Fields

Because evaluation is inter-disciplinary and transcends the boundaries of specific fields (Stern, 2005), it is challenging to classify evaluations into clearly specified disciplines. Nonetheless, a careful reading of the articles did reveal eight categories that do a reasonably good job of describing content areas:

1) Social work and community
   Example
   Capacity for effectiveness: the relationship between coalition structure and community impact by Hays, C. E., Hays, S. P., DeVille, J. O. & Mulhall, P. F. 2000 (GS: 93, WoS: 42).

Table 4
Total articles in special and regular issues by year.

Year	Regular issues	Special issues
2000	41	0
2001	39	0
2002	23	22
2003	33	8
2004	22	16
2005	33	6
2006	16	29
2007	32	6
2008	35	7
2009	27	14
2010	37	34
2011	36	6
2012	42	21
Total	416	169

Table 5
Percent of articles by academics, practitioners, and academic/practitioner collaborations.

Year	Academic	Practitioner	Academic-practitioner collaborations
2000	23	6	12
2001	19	11	9
2002	18	15	12
2003	14	19	8
2004	19	11	8
2005	20	13	6
2006	19	11	15
2007	31	2	5
2008	26	5	11
2009	26	6	9
2010	41	15	15
2011	22	5	15
2012	36	10	17
Total	314	129	142

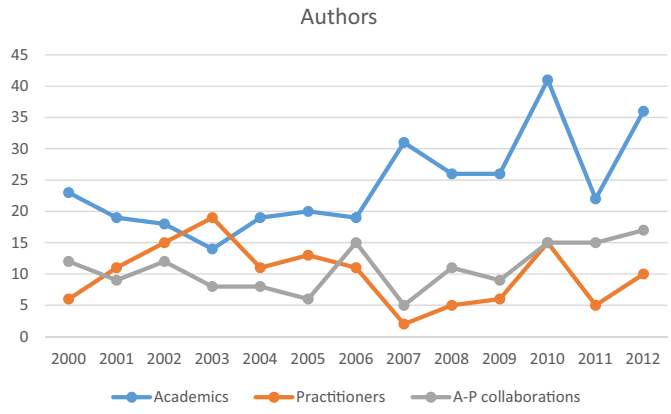

Fig. 3. Graphical view: Percent of articles by academics, practitioners, and academic/practitioner collaborations.

2) Public health
   Example
   Treatment outcomes among adolescents with substance abuse problems: the relationship between comorbidities and post-treatment substance involvement by Shane, P. A., Jasiukaitis, P. & Green, R. S. 2003 (GS: 88, WoS: 49).
3) Organization management and performance
   Example
   Measuring outcomes and managing for results by Schalock, R. L. & Bonham, G. S. 2003 (GS: 95, WoS: 35).
4) Education and knowledge
   Example
   Performance evaluation of extension education centers in universities based on the balanced scorecard by Wu, H. Y., Lin, Y. K. & Chang, C. H. 2011 (GS: 72, WoS: 17).
5) Public policy and government affairs
   Example
   Predicting intermediate outcomes for prevention coalitions: a developmental perspective by Florin, P., Mitchell, R., Stevenson, J. & Klein, I. 2000 (GS: 85, WoS: 44).
6) Business and enterprise management
   Example
   Differences in stakeholder perceptions about training evaluation: a concept mapping/pattern matching investigation by Michalski, G. V. & Cousins, J. B. 2000 (GS: 53, WoS: 13).
7) Agriculture, environment and ecology
   Example
   Using conceptual models as a planning and evaluation tool in conservation by Margoluis, R., Stem, C. & Salafsky, N. & Brown, M. 2009 (GS: 64, WoS: 28).
8) Economics and cost
   Example
   A multiple account framework for cost-benefit analysis by Campbell, H. F. & Brown, R. P. C. 2005 (GS: 32, WoS: 12).

Table 6 shows the numbers of articles in each category by year. Most of the evaluations described in the articles fall into three areas: social work/community (30%), public health (27%), and organizational management/performance (24%). Each of the other areas represent 10% or less. While the numbers in some categories are small, the existence of eight separate categories does reflect the success of EPP's stated goal of presenting a broad, inter-disciplinary perspective on evaluation.

In term of citations count, ANOVA found no significant differences in the number of citations per article across the eight categories of article content. Despite this lack of statistical significance, it is worth noting that the three most studied areas did receive the highest number of citations per article: social work/community (8.25, GS: 21.83), public health (WoS 9.64, GS: 20.84), and organizational management/performance (WoS 8.76, GS: 27.17).

4.4. Locations of evaluations

Table 7 shows the region of origin for articles by year. In all years, evaluations conducted in North America (primarily in the US) have dominated, and this trend has not changed over time.[3] (Once there is enough data for a more current analysis, this finding may be different, as EPP has been implementing a deliberate policy to publish evaluations from a larger number of countries.)

4.5. Methodology

Careful reading of the articles revealed seven common methodologies:

1) Survey
   Data from distributed questionnaires (Fink, 2003).
2) Secondary data

[3] We found that 74% of authors' affiliation are in the US, followed by Europe (9%). Six percent of the articles were collaborative works between authors in different regions, ahead of authors solely from Africa, Oceania and Asia (3%).

Table 6
Numbers of articles by topic per year.

Year	Social work/ community	Public health	Organizational management/ performance	Education/ knowledge related	Public policy/ government affairs	Business/enterprise management	Agriculture/ environment/ ecology	Economics/cost related
2000	20	12	0	4	3	2	0	0
2001	13	9	9	3	5	0	0	0
2002	13	12	11	2	3	4	0	0
2003	7	8	22	2	2	0	0	0
2004	8	15	6	4	1	4	0	0
2005	10	8	13	0	5	1	0	2
2006	15	14	13	2	1	0	0	0
2007	10	7	16	5	0	0	0	0
2008	13	16	8	5	0	0	0	0
2009	15	3	5	7	3	0	3	5
2010	22	27	7	12	2	0	1	0
2011	10	13	11	2	3	0	2	1
2012	14	14	18	11	4	0	2	0
Total	170	158	139	59	32	11	8	8

Table 7
Countries of origin for articles by year.

Year	Asia	Europe	North America	South America	Oceania	Africa	Cross-region
2000	1	2	34	0	3	1	0
2001	0	0	38	0	0	1	0
2002	0	6	17	0	2	18	2
2003	0	4	34	1	2	0	0
2004	0	1	34	0	0	1	2
2005	0	2	33	1	2	0	1
2006	1	2	37	1	0	1	3
2007	1	4	25	1	1	3	3
2008	1	1	30	1	2	3	4
2009	3	10	20	2	3	3	0
2010	4	4	58	1	1	2	1
2011	5	4	27	1	2	2	1
2012	7	10	40	0	4	2	0
Total	23	50	427	9	22	37	17

Data from existing or published resources such as annual reports, books, census data, clinical records, company records, government publications and periodicals.

3) Interview
Qualitative data from verbal interaction with respondents (in-person, telephone, and on-line).

4) Case study
Data from comprehensive examination of specific settings of varying units of analysis (e.g. people, companies, events) over a sustained period of time (Creswell, 2009).

5) Experiments and quasi-experiments
A systematic comparison among groups to provide a test under controlled conditions that is made to demonstrate a known truth, examine the validity of a hypothesis, or determine the efficacy of something previously untried (Shadish, Cook, & Cambell, 2002).

Table 8
Methodologies used in research by year.

Year	Survey	Secondary data	Interview	Case Study	Experiment	Grounded theory	Mixed
2000	5	8	3	3	1	2	19
2001	1	12	1	2	2	0	21
2002	0	5	6	3	0	0	31
2003	3	1	6	2	0	0	29
2004	2	6	10	2	0	0	18
2005	1	7	2	1	0	0	28
2006	5	7	2	2	0	0	29
2007	3	10	4	2	0	0	19
2008	5	6	10	1	0	0	20
2009	5	6	1	0	0	0	29
2010	8	32	6	0	0	0	25
2011	5	12	5	1	0	0	19
2012	1	19	4	0	0	0	39
Total	44	131	60	19	3	2	326

6) Grounded theory
   Deriving meaning and structure as revealed by an ongoing analysis of data, rather than by beginning with a specific hypothesis (Hussein, Hirst, Salyers & Osuji, 2014).
   Data from systematic procedures for developing an inductively derived theory from the data (Corbin & Strauss, 1990).
7) Mixed method
   Data from more than one of the above (Yin, 2009).

Table 8 shows the number of methodologies used by year. More than half of the evaluations employed mixed methods. Forty four percent of studies used secondary data, followed by interview data (10%), survey data (7%) and case study (3%). Although only two articles employed grounded theory, it is worth noting that the average number of citations for these two articles received the highest number of citations (Average GS: 42.5, WoS: 18.5), compared to the average number of citations for other methodologies. Those two articles are:

1. Concept mapping in mental health: uses and adaption by Johnsen, J. A., Biegel, D. E. & Shafran, R. 2000 (GS: 68, WoS: 31).
2. Assessing program impact using latent growth modeling: a primer for the evaluator by Hess, B. 2000 (GS: 17, WoS: 6).

Just behind grounded theory, the next highest average citation count methodology types were: Surveys (Average GS: 26.05, WoS: 10.5) and interviews (Average GS: 24.23, WoS: 9.97). The two articles employing these methodologies that received the highest number of citations are:

1. Consumer based quality of life assessment: a path model of perceived satisfaction by Schalock, R. L., Bonham, G. S. & Marchand, C. B. 2000 (GS: 101, WoS: 39).
2. Stigma, HIV/AIDS and prevention of mother-to-child transmission in Zambia by Bond, V., Chase, E. & Aggleton, P. 2002 (GS: 174, WoS: 72).

4.6. Quantitative and qualitative data

Unlike the previous discussion of the use of mixed methods, here we focus on a particular type of method mixing, i.e. the combination of qualitative and quantitative data. Table 9 and Fig. 4, show numbers of studies that relied primarily on quantitative and qualitative data, by year. That analysis reveals that most of the evaluations are qualitative in nature. Yet, while qualitative analyses predominate, there is ambiguous evidence concerning the appeal of these methodologies to the evaluation community.

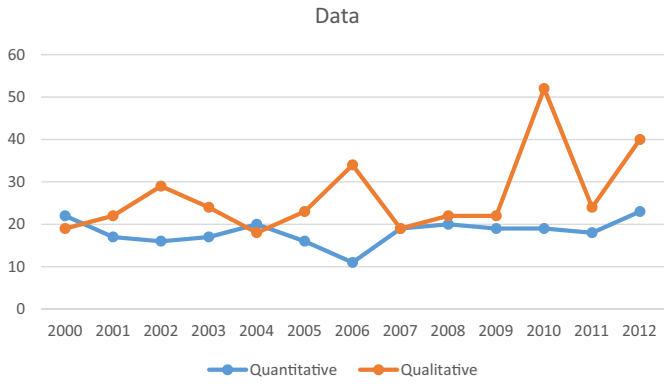

Fig. 4. Graphical view: Numbers of quantitative and qualitative studies by year.

GS shows a greater number of qualitative methodologies (average 23.05 versus 22.04), while WoS shows more citations for quantitative studies (9.35 versus 87.55).

4.7. Cross-sectional and longitudinal data

Table 10 and Fig. 5 show the number of cross-sectional and longitudinal studies by year. Data show that evaluations using cross-sectional data represent double or more than the evaluations that used longitudinal data in all years of publication. In terms of citation counts, our *t*-test shows no significant difference between these two methodologies.

4.8. Independent or funded research

There is some reason to believe that publications based on funded research receive higher citation counts than those received by publications based on unfunded research. For instance, this has been shown to be the case in the field of library and information science (Zhao, 2010). (Of course this begs the question of what it means to say that a project was not funded, because ultimately everything is paid for. But, presumably, there is a difference between research that is done in the normal course of someone's work, and research that has dedicated resources behind it.) Might we find a similar pattern in the field of evaluation? To address this question we classified all articles as funded or unfunded based on whether a funding source is acknowledged by an article's authors. The data are presented in Table 11 and Fig. 6. Unlike the pattern that might be expected, in the case of evaluation the funded/not funded distinction does not seem to matter. Differences both for the WoS data and the GS data are not statistically significant.

Table 9
Numbers of quantitative and qualitative studies by year.

Year	Quantitative	Qualitative
2000	22	19
2001	17	22
2002	16	29
2003	17	24
2004	20	18
2005	16	23
2006	11	34
2007	19	19
2008	20	22
2009	19	22
2010	19	52
2011	18	24
2012	23	40
Total	237	348

Table 10
Numbers of cross-sectional and longitudinal studies by year.

Year	Cross-sectional	Longitudinal
2000	27	14
2001	25	14
2002	33	12
2003	27	14
2004	25	13
2005	25	14
2006	33	12
2007	27	11
2008	25	17
2009	22	19
2010	60	11
2011	31	11
2012	37	26
Total	397	188

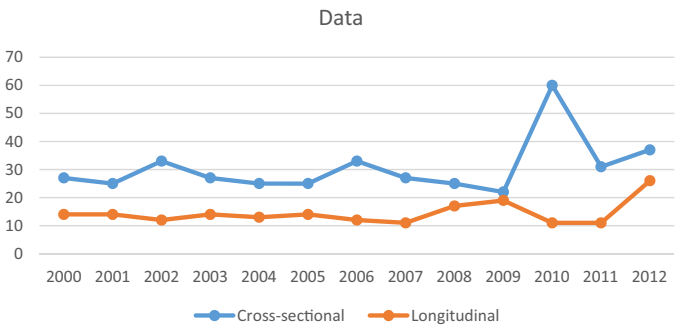

Fig. 5. Graphical view: Numbers of cross-sectional and longitudinal studies by year.

Table 11
Numbers of funded and non-funded research by year.

Year	Funded	Non-funded
2000	19	22
2001	19	20
2002	16	29
2003	23	18
2004	22	16
2005	10	29
2006	20	25
2007	15	23
2008	23	19
2009	23	18
2010	23	48
2011	23	19
2012	31	32
Total	267	318

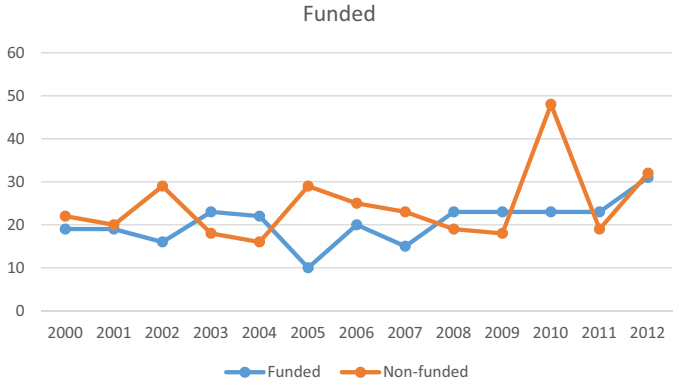

Fig. 6. Graphical view: Numbers of funded and non-funded research by year.

5. Conclusions

EPP began publication shortly after evaluation in the modern sense of the term gained currency among social scientists and policy makers. Since then its trajectory has been determined by an interaction between the journal's editorial policy, and the work produced by people who self-identified as evaluators, and who thus saw it as worthwhile to share their work with others who similarly self-identified. In this sense, the identity of EPP can be seen as part of the organic growth of the field, as the field has developed from engagements among evaluation practitioners, evaluation theorists, and those who seek evaluation services (Morell, 2013). That development is reflected in our data in several ways.

5.1. Practitioners and academics

The number of articles authored by practitioner/academic collaborators indicates a rich collaboration among people whose primary professional responsibility is the production of evaluation, and those who respond to a knowledge production function about evaluation that speaks to the traditional academic reward system. The mix of academics and practitioners (both as sole authors and as collaborators) is a sign that evaluation as a field can be thought of as an endeavor whose success over time requires a particularly close collaborative interaction between intellectual input that has a primary objective of practical application, and intellectual input that also includes some combination of student training and contributions to a scientific field. We do not know if this kind of close collaboration is needed to fuel other fields, but it certainly seems to be needed in evaluation.

On average, articles published solely by academics are cited more frequently than those published either solely by practitioners, or through practitioner/academic collaborations. This finding carries meaning for understanding the development of the field of evaluation. Discussions of the academic and practitioner mindsets hold that academic researchers are primarily rewarded for the pursuit of knowledge, while practitioners have a primary interest in practical solutions to specific problems (Brennan & Ankers, 2004). So, it certainly seems reasonable to believe that articles published solely by practitioners will be focused on specific, narrow, evaluation concerns. It is less obvious, but plausible, to believe that any time an article includes a practitioner author, the motivation behind the work is relatively narrowly focused on a specific practical evaluation need. Reasonable, but not without challenge because it is also true that of 585 articles in our sample, the 20 with the greatest number of citations were authored by practitioners.

Thus we have three findings. First, that the history of publication in evaluation has a particularly rich history of collaboration between practitioners and academics. Second, that on average, articles published solely by academics are cited most frequently. Third, that if an article has an extraordinarily large number of citations, it was authored by a practitioner. Given these findings, what is the implication for the mix of academic and practitioner perspectives in the field? We believe these findings reinforce the importance of practitioner/academic collaboration not in terms of individual articles, but over time, for the development of the field as a whole. Over time: 1) academic/practitioner collaboration is important for "getting the news out there", 2) in terms of impact on the field, on average, sole academic authorship has the most impact, but that 3) practitioners have the greatest potential to do work that draws the attention of the evaluation community.

5.2. Open submission and special issues

Submissions from open sources reflect the motivations of evaluators to publicize their work. In a sense, they can be seen as a true reflection of what the field in general is doing and sees as interesting. Special issues reflect the elite (and narrow) opinions of the editors and the editorial board. It is worth noting that articles that come from open submissions from the field are cited more frequently than articles in special issues. This finding presents a challenge to editorial policy. On the one hand, EPP must reflect the interests and inclinations of the field. On the other hand, another function of EPP is to be more than just reactive to its readers. Rather, the journal is committed to using its resources to make intelligent guesses as to where the field should be going, and to publicize those possible directions. Or as another way of stating

this, we see EPP as having a responsibility to influence the course of development of the field, and not just passively reflect it.

5.3. What is evaluated?

All programs of any kind are evaluated in the sense that value judgments are made about them by stakeholders and interested parties. But only some programs are evaluated in the sense intended by the group of people who have coalesced into what we call the field of evaluation. As reflected in the pages of EPP, the field of evaluation consists mostly of work dealing with social work and community, public health, and organizational management and performance. This is certainly a function of the authors that EPP has attracted over the years, rather than an unbiased reflection of work done by people who call themselves evaluators. We encourage people to do similar analyses of other journals whose public image attracts work by other evaluators.

5.4. Methodology

The predominant methods used by EPP's authors are mixed methods and secondary data. This finding reflects three related aspects of the programs that are evaluated by contributors to EPP. The first is the challenge of examining programs that are set up primarily to produce some social impact, rather than set up primarily as a test of some particular program or intervention. As a result, opportunity is limited for evaluators to design and implement targeted measures. The second is the reality that any change dropped into a real-world setting is likely to exhibit elaborate and varied connections among many elements that reside both inside and outside of whatever is being evaluated. Thus even if some special measures could be implemented by the evaluators, it is unlikely that special measures could be designed and implemented for all the program activity that may be of interest. The third is that anything that is evaluated will generate records of its operations that are either internal to its operations, or are indications of how the environment reacts to those operations. Further, some of this data will be the most relevant information needed for the evaluation. Because of the dearth of specially developed measures, the variety of issues that must be assessed, and the high relevance of existing information, evaluators are naturally attracted to searching for any and all data that may be available in records. Because the data are so varied, so too must be the methodologies used to analyze them.

Acknowledgement

The authors would like to thank Said M.S. for his tremendous assistance for developing the database for this study.

References

Baumgartner, H., & Pieters, R. (2003). The structural influence of marketing journals: a citation analysis of the discipline and its subareas over time. *Journal of Marketing*, *67*(2), 123–139.

Brennan, R., & Ankers, P. (2004). In search of relevance: is there an academic-practitioner divide in business-to-business marketing? *Marketing Intelligence & Planning*, *22*(5), 511–519.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21.

Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

Fink, A. (2003). *A survey handbook*. Thousand Oaks, CA: Sage Publications.

Harzing, A. W., & Van der Wal, R. (2007). Google scholar: the democratization of citation analysis? *Ethics in Science and Environmental Politics*, *8*(1), 61–73.

Hussein, M. E., Hirst, S., Salyers, V., & Osuji, J. (2014). Using grounded theory as a method of inquiry: advantages and disadvantages. *The Qualitative Report*, *19*(27), 1–15.

Marsh, S. J., & Hunt, C. S. (2006). Not quite as simple as ABC reflections on one department's experiences with publication ranking. *Journal of Management Inquiry*, *15*(3), 301–315.

Meho, L. I. (2006). The rise and rise of citation analysis. arXiv preprint physics/0701012.

Morell, J. (2013). *Using an evolutionary biology view to connect the intellectual development of evaluation and the development of the evaluation community*. . http://evaluationuncertainty.com/2013/06/27/using-an-evolutionary-biology-view-to-connect-the-intellectual-development-of-evaluation-and-the-development-of-the-evaluation-community/.

Nerur, S. P., Rasheed, A. A., & Natarajan, V. (2008). The intellectual structure of the strategic management field: an author co-citation analysis. *Strategic Management Journal*, *29*(3), 319–336.

Nkomo, S. M. (2009). The seductive power of academic journal rankings: challenges of searching for the otherwise. *Academy of Management Learning & Education*, *8*(1), 106–121.

Powell, R. R. (2006). Evaluation research: an overview. *Library Trends*, *55*(1), 102–120.

Scriven, M. (1991). *Evaluation thesaurus*. Thousand Oaks, CA: Sage Publications.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference*. Independence, KY: Wadsworth Cengage Learning.

Stern, E. (2005). *Evaluation research methods*. London, UK: Sage Publications.

Steward, A., & Cotton, J. (2013). Making sense of entrepreneurship journals: journal rankings and strategy choices. *International Journal of Entrepreneurial Behaviour & Research*, *19*(3), 303–323.

Trochim, W. M., & Donnelly, J. P. (2001). *The research methods knowledge base*. Cincinnati, OH: Atomic Dog Publishing.

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: an author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, *49*(4), 327–355.

Williams, D. D. (2000). Evaluation of learning objects and instruction using learning objects. *The instructional use of learning objects*. . http://www.reusability.org/read/chapters/williams.doc.

Yin, R. K. (2009). *Case study research: design and methods*. Thousand Oaks, CA: Sage Publications.

Zhao, D. (2010). Characteristics and impact of grant-funded research: a case study of the library and information science field. *Scientometrics*, *84*(2), 293–306.

Abu H. Ayob is a senior lecturer at the Faculty of Economics and Management, National University of Malaysia. He holds a MSc in Management from Toulouse Business School and a PhD in Management Science from Universite Toulouse 1 Capitole, France. He is currently a visiting post-doctoral researcher at Department of Strategic Management and Globalization, Copenhagen Business School. His primary research field is the internationalization of firms in emerging economies.

Jonathan A. Morell PhD (Jonny) has extensive experience in a wide range of evaluation settings. Lately he has been evaluating industrial safety, and R&D programs. He is a leader in developing applications from complexity science to traditional program evaluation. Jonny is Director of Evaluation at Syntek Technologies, Editor-in-Chief of the journal Evaluation and Program Planning, and a Principal at 4.669 Evaluation and Planning, LLC. He blogs at www.evaluationuncertainty.com.