Information Processing and Management 52 (2016) 658–669

Contents lists available at ScienceDirect

Information Processing and Management

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

Influence of *human behavior and the principle of least effort* on library and information science research

Yu-Wei Chang*

Department of Library and Information Science, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 17 June 2015
Revised 17 December 2015
Accepted 17 December 2015
Available online 12 January 2016

Keywords:
Principle of least effort
Citation context
Library and information science

ABSTRACT

This study identified the influence of the main concepts contained in Zipf's classic 1949 book entitled *Human Behavior and the Principle of Least Effort* (*HBPLE*) on library and information science (LIS) research. The study analyzed LIS articles published between 1949 and 2013 that cited *HBPLE*. The results showed that *HBPLE* has a growing influence on LIS research. Of the 17 cited concepts that were identified, the concept of "Zipf's law" was cited most (64.8%), followed by "the principle of least effort" (24.5%). Although the concept of "the principle of least effort," the focus of *HBPLE*, was not most frequently observed, an increasing trend was evident regarding the influence of this concept. The concept of "the principle of least effort" has been cited mainly by researchers of information behavior and served to support the citing authors' claims. By contrast, the concept of "Zipf's law" received the most attention from bibliometrics research and was used mainly for comparisons with other informetrics laws or research results.

1. Introduction

The principle of least effort (PLE), a concept advanced by the American linguist George Kingsley Zipf, indicates that people complete tasks by choosing the way of least effort among various options (Zipf,1949). To prove that the PLE is an indication of human nature, Zipf analyzed numerous empirical data collected from various human activities including those related to language, geographic distribution of the population, economic activities, social status, and collaboration, and used mathematical formulae to explain his findings. Zipf explained the PLE in detail in his classic 1949 entitled *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (*HBPLE*).

Opinions on the PLE are controversial. A few researchers have criticized the PLE for its oversimplification of human behavior (Cummings, 1977; Rapoport, 1957), whereas other researchers have supported and further explained the concept. Gratch (1990) stated that the PLE is the result of human experience and that exerting the least amount of effort is not equivalent to laziness. Case (2005) reported that people are willing to acquire lower-quality or less information to reduce their expenses.

The PLE represents a common human behavior; it may thus be expected that the HBPLE has become visible in various fields and applied to various human activities. *HBPLE* was also compared with similar theories and was reconceptualized in the field of library and information science (LIS). For example, one study indicated that Mooers' law was derived from the PLE (Bierbaum, 1990). Mooers' law states that users would rather not have information if having that information is

* Tel.: +886 2 3366 1743; fax: +886 2 2363 2859.
E-mail address: yuweichang2013@ntu.edu.tw, yuweichang2007@gmail.com

http://dx.doi.org/10.1016/j.ipm.2015.12.011
0306-4573/

painful, which is true even for users who can efficiently access information by using information retrieval systems (Austin, 2001; Mooers, 1996). Saving readers' time, one of the five laws of library science, is also similar in meaning to the PLE (Gratch, 1990). The LIS publications on PLE have indicated that the concept of the PLE is connected to various topics, such as cataloguing (Chrzastowski, 1999; Wei, 2005), citation (White, 2001), information services (Chrzastowski, 1999; Duan, 2004), information seeking behavior (Bigdeli, 2007; Bronstein, 2008; Kebede, 2004; Xu, Tan, & Yang, 2006), information retrieval (Chen & Liang, 2000; Kim, 1982; Wang, 2000), and the location of resources (Chrzastowski, 1995; Wang, 2001). This shows that *HBPLE* has established its influence on various subfields within LIS research. To understand the relationship between *HBPLE* and LIS research, this study examined the influence of *HBPLE* on LIS articles.

Because the PLE is the core of *HBPLE*, I assumed that the PLE is the most frequently cited concept in *HBPLE*. Particularly, I assumed that the PLE is cited most by authors of information behavior articles because it has been regarded as an essential theory of information behavior according to the PLE being listed in the 2005 book, *Theories of Information Behavior*. Information behavior research is multidisciplinary and has been an essential subfield in LIS. Some theories of information behavior originating outside LIS have guided LIS researchers in advancing information behavior research, and the PLE is one example of these theories of information behavior. However, these assumptions required examination. I found that a few articles on the PLE did not cite the concept of the PLE in *HBPLE* and that certain publications not on the PLE cited the concept of the PLE in *HBPLE*. Furthermore, publications include numerous concepts. Apart from the PLE, *HBPLE* may consist of other influential concepts that have attracted LIS authors' attention.

To clarify whether the concept of the PLE has received the highest attention among all concepts in *HBPLE*, I examined which concept in *HBPLE* was most frequently cited by authors of LIS articles that were published between 1949 and 2013. This study also analyzed the citation functions that reveal the possible reasons why authors cite the concepts in *HBPLE*. Combining the analyses of cited concepts and citation functions can reveal the influence of *HBPLE* on LIS research in detail. In addition, I analyzed citation frequency trends and the research topics of citing articles to identify emerging trends in the influence of *HBPLE* on LIS research and to determine which topics in LIS research have involved applying the concepts in *HBPLE*. The assumption of whether information behavior articles cite the concept of the PLE most can be verified. The results contribute to the understanding how a classic book on linguistics has influenced LIS research from three aspects: the number of citations, cited concepts, and citation functions.

This study answers the following research questions:

(1) Does the concept of the PLE have the most influence on citing authors among all concepts in *HBPLE*?
(2) For the citing articles, what functions does *HBPLE* serve? What functions does the concept of the PLE serve?
(3) What research topics do the citing articles focus on? Is the concept of the PLE cited most by information behavior articles?

2. Literature review

Content analysis is useful for demonstrating the substantial influences on works. Citation context analysis, which is derived from the method of content analysis, can be used to identify the relationship between the contents of citing and cited works by examining partial texts surrounding a specific citation (Small, 1982). Citation context analysis relies on the researcher's judgment. Cited concepts can be confirmed by examining the content of cited work. However, analyzing citation functions is difficult because citation functions may not be apparent in texts. Some researchers have argued that authors' motivations for citing literature are complex and indicate that only authors know why they cite a specific work (Case, & Higgins, 2000; Shadish, Tolliver, Gray, & Gupta, 1995). In such a situation, interviewing or surveying citing authors is another common method for exploring authors' citing behavior (Bonzi & Synder, 1991; Harwood, 2009; White & Wang, 1997). However, some limitations still exist in interviews and surveys. Researchers have encountered difficulty in contacting all citing authors, especially when citing authors are deceased. This leads to a limited number of interviewees who can be interviewed. In addition, citing authors may forget why they cited a work a decade previously or may withhold the true reasons because they had negative intentions (Harwood, 2009). Because I conducted a longitudinal study to trace the changes in cited concepts and citation functions, I performed citation context analysis to explore the meaning of citations.

The method of bibliometric analysis is also used to measure the influence of a publication by counting its citation frequency and investigating the characteristics of citing papers (Carvalho, Diniz-Filho, & Bini, 2005; Chu, 2001; Furner, 2003). A high number of citations indicate greater influence. Although changes in the influence of a publication over time can be easily tracked, connections between the contents of cited and citing documents cannot be identified. This implies that the influence of a publication can be deeply understood by combining citation context analysis and bibliometric analysis. I used bibliometric analysis to observe the annual changes in the number of citations that *HBPLE* and specific concepts in *HBPLE* received.

Citation context analysis contains two types of content analysis: cited concepts and citation functions. The analysis of cited concepts is undertaken to identify content that has been cited by other researchers. A classic work has substantial influence on researchers in a specific field or across fields and usually becomes a primary subject of study. In the 1980s, Garfield (1980) identified 26 cited concepts from publications by Merton, a distinguished sociologist. Colemand and Salamon (1988) identified the main cited concepts embedded in Kuhn's *The Structure of Scientific Revolutions*. Since 2006, more related studies have appeared. Researchers have explored the influential concepts embedded in classic works on social psychology

(Anderson, 2006), software programs management (McCain & Salvucci, 2006), ecology (Richardson & Pysek, 2008; Thomaz, Michelan, Carvalho, & Bini, 2010), hypertext (Tsay, 2009), organizational memory (Anderson & Sun, 2010), information behavior (Chang, 2013b), and scientometrics (Chang, 2013a).

Citation function is a method used to reveal the reasons why the specific content of a publication was cited. Citation function studies have been paid attention since 1960s (Bornmann & Daniel, 2008; Case & Higgin, 2000; Chang, 2013a; Frost, 1979; Lipetz, 1965; McCain & Turner, 1989; Moravcsik & Murugesan, 1975; Oppenheim & Renn, 1978; Spiegel-Rosing, 1977; Tohidinasab & Jamali, 2013; White & Wang, 1997). Most researchers have developed classification schemes for citation functions in their studies and few researchers have modified previous classification schemes for citation functions. Some classification schemes consisted of more than 10 categories (Spiegel-Rosing, 1977; Tohidinasab & Jamali, 2013), whereas other classification schemes were simple and consisted of only four categories (Frost, 1979; Moravcsik & Murugesan, 1975). The structures of classification schemes depend on research purposes. Although Small (1982) claimed that most citation functions named with different terms among various studies were the same, I did not compare the results of citation functions because of differences in definition.

3. Methodology

3.1. Data collection

To collect bibliographic records of articles citing *HBPLE* published between 1949 and 2013, we searched the multidisciplinary citation index database Web of Science (WoS), which provides users access to a wide range of bibliographic and citation information from articles published in international journals over a long timespan. I obtained the full text of articles that WoS indicated had cited *HBPLE*. The citing articles analyzed in this study had to meet four requirements. First, they have to have been published in LIS journals. The LIS journal candidates had to be included in the subject category of "Information Science and Library Science" in the social science edition of the 2012 Journal Citation Reports (JCR) and the subject category of "Library and Information Science" in the database provided by Ulrichsweb.com. Each journal indexed by JCR and Ulrichsweb.com was categorized under one or more subject categories. Because a few journals belonging to the subject category of "Information Science and Library Science" in the JCR do not focus on LIS, their subject attributes were further examined using the database provided by Ulrichsweb.com. The common LIS journals indexed by the JCR and Ulrichweb.com databases were selected for this study. Second, the publication language of articles had to be English or Chinese due to our language limitations. Third, because texts vary in nature, only research articles—which are the main texts published in journals—were examined. Fourth, articles were required to have been published between 1949 and 2013.

Regarding the search strategy used for collecting the citing articles, search terms were first combined in two designated fields: the cited author field and publication year of the cited work. The title and author name of a cited work are not recorded in full in WoS. Because of the problem of inconsistent data regarding the recording of the same title and author name, the only possible search strategy was used: The surname of the cited author and the year of publication of the cited work were combined. Therefore, "Zipf" was designated in the field of the cited author. In addition, I learned that *HBPLE* was reprinted in 1956, 1972, and 2012, based on two databases: OCLC's WorldCat and Bowker's Global Books in Print. Four pairs of search terms combining the surname of the cited author and the specific publication year of the cited work were searched for separately.

Next, I chose the index list of the cited work including the author names and titles related to ZIpf's *HBPLE* generated from the search strategy. The variations in Zipf's name included "Zip," "Zipf, G," "Zipf, GK.," "Zipf George Kingsley," and "Zipf George K."; bibliographic title data included, for example, "HUMAN BEHAV," "HUMAN BEHAVIOR PRINC," and "BEHAV PRINCIPAL LEAS." Because the bibliographic records of a few citing articles included the wrong author's name and the correct title or the correct author's name and the wrong title, I further identified the correct citing articles by examining their full text. Thus, the initial search results were limited to a specific type of document (research articles), a specific publication year range (1949–2013), and a specific field (LIS). The combined search results returned 263 bibliographic records of LIS articles citing *HBPLE*.

3.2. Data analysis and processing

Before citation context analysis, I examined the in-text citations and reference lists of 263 English articles to ensure that the authors cited *HBPLE*. Each in-text citation was defined as an independent citation context. Of the 263 citing articles, three were excluded from the dataset because of citation errors existed between the in-text references and reference lists (two articles), or because full-text articles could not be obtained (one article). A citing article could have two or more citation contexts referring to *HBPLE*. Finally, I analyzed 260 citing articles including 310 citation contexts. Most citation contexts (91.9%) were collected from articles that cited the 1949 edition of *HBPLE*. Only 5.8% of citation contexts cited the 1965 edition of *HBPLE* and 2.3% cited the 1972 edition of *HBPLE*. In addition, each citing article had one to five citation contexts referring to *HBPLE*. Most of the citing articles (225 articles, 86.5%) had only one citation context referring to *HBPLE*. Twenty-six citing articles (10.0%) had two citation contexts referring to *HBPLE*. Four citing articles had three citation contexts referring to *HBPLE*, four citing articles had four citation contexts referring to *HBPLE*, and one citing article had five citation contexts referring to *HBPLE*.

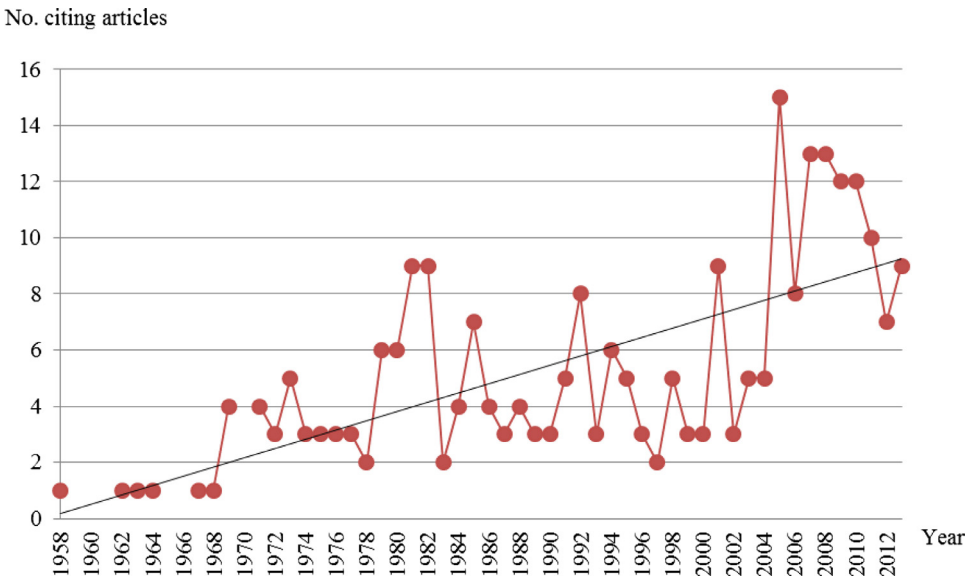

Fig. 1. Comparison of logarithms of *HBPLE*-citing articles in LIS and LIS articles by year.

The cited concept and citation function of each citation context were recorded after reading the text surrounding each citation context. During the analysis process, I reviewed *HBPLE* to ensure that the cited concepts were included in *HBPLE*; I then derived the keywords or statements used by citing authors to refer to the concepts in *HBPLE*. After analyzing all citation contexts, I examined and compared the notes for each citation context. Some different terms and statements that were identified to have the same meaning referring to the same concept in *HBPLE* were incorporated into a single category. One category represented one independently cited concept. Finally, several categories were formed.

The classification scheme of citation functions was developed based on a temporary classification scheme devised after reviewing previous studies and was modified during the analysis process. Regarding the main topic of each citing article, it was coded based on its title, keywords, abstracts, and full text. Similar topics were incorporated into topic that were the same but had broader coverage.

All analyses were conducted by one researcher. To improve the consistency of the classification schemes of cited concepts, citation functions, and topics of citing articles, all citing articles were analyzed twice or thrice at various times. When the first two results of classification were inconsistent, a third analysis for classification was conducted. In addition, citing articles in the same topic category were reexamined to guarantee that consistent criteria were applied during the analysis process.

4. Results

4.1. Trends in the number of LIS citing articles

As shown in Fig. 1, comparing the logarithms of the annual totals of *HBPLE*-citing articles in LIS from 1958 to 2013 showed a clear increasing trend. Comparing the logarithms of the annual totals of LIS articles during the same period also presented a clear increasing trend (Fig. 1). The continually increasing number of articles published in LIS journals may support the fact that an increasing number of LIS articles cite *HBPLE*. However, the point is that, although *HBPLE* was published 65 years ago, the influence of *HBPLE* has not decreased over time. *HBPLE* has been proven as a classic work in the field of LIS. *HBPLE*, however, was not cited in LIS articles until the tenth year after its publication. During this study period, 260 LIS articles cited *HBPLE*. The annual number of citations that *HBPLE* received was less than 10 before 2005. During the period of 2005–2013, *HBPLE* received 7 to 15 citations per year. The largest growth in the annual number of citing articles was observed from 2004 to 2005. The number of citing articles in 2005 was 10 articles more than that in 2004.

4.2. Journal titles of the citing articles

Table 1 shows the 39 LIS journals that cited *HBPLE*. Among the 39 LIS journals, 17 (43.6%) focus on library science and 22 (56.4%) on information science. Seventeen journals (35.4% of 39 journals) cited *HBPLE* only once. Regarding the number of the citing articles, only 19.6% of the citing articles were published in library science journals. Over half of the LIS citing articles (56.2%) were published in three journals: *Journal of the American Society for Information Science and Technology* (25.0%), *Information Processing & Management* (16.5%), and *Scientometrics* (14.6%).

Table 1
Journals with articles that cited *HBPLE*.

No.	Journal title	No. of articles	Percentage (%)
1	*Journal of the American Society for Information Science and Technology**	65	25.0
2	*Information Processing & Management**	43	16.5
3	*Scientometrics**	38	14.6
4	*Journal of Documentation**	16	6.2
5	*Journal of Information Science**	14	5.4
6	*Information Research: An International Electronic Journal**	7	2.7
7	*Journal of Informetrics**	7	2.7
8	*Library Trends*	7	2.7
9	*Proceedings of the American Society for Information Science**	7	2.7
10	*Knowledge Organization*	6	2.3
11	*College & Research Libraries*	4	1.5
12	*Library & Information Science Research*	4	1.5
13	*Libri*	4	1.5
14	*Canadian Journal of Information and Library Science*	3	1.2
15	*Library Quarterly*	3	1.2
16	*Special Libraries*	3	1.2
17	*Annual Review of Information Science and Technology**	2	0.8
18	*Journal of Library Automation*	2	0.8
19	*Library Resources & Technical Services*	2	0.8
20	*Online Information Review**	2	0.8
21	*Reference & User Services Quarterly*	2	0.8
22	*Research Evaluation**	2	0.8
23	*American Archivist*	1	0.4
24	*Bulletin of the Medical Library Association*	1	0.4
25	*Government Information Quarterly*	1	0.4
26	*Health Information and Libraries Journal*	1	0.4
27	*Information & Management**	1	0.4
28	*International Forum on Information and Documentation**	1	0.4
29	*International Journal of Information Management**	1	0.4
30	*Journal of Librarianship and Information Science*	1	0.4
31	*Journal of Research Communication Studies**	1	0.4
32	*Journal of Strategic Information Systems**	1	0.4
33	*Journal of the American Medical Informatics Association**	1	0.4
34	*Law Library Journal*	1	0.4
35	*Library Hi Tech*	1	0.4
36	*Portal: Libraries and the Academy*	1	0.4
37	*Program: Electronic Library and Information Systems*	1	0.4
38	*Publishing Research Quarterly*	1	0.4
39	*Serials Librarian*	1	0.4

Note: * information science journals.

Table 2
Distribution of citing article topics.

Topics	No. of articles	Percentage
Bibliometrics	121	46.5
Information retrieval	76	29.2
Information behavior	26	10.0
Information service	7	2.7
Collection development	7	2.7
Information science	7	2.7
Knowledge organization	7	2.7
Management	6	2.3
Scholarly communication	3	1.2
Total	260	100.0

4.3. Topics of citing articles

Table 2 shows that the 260 citing articles represent 9 topics. Bibliometrics articles accounted for the largest share (46.5%), followed by information retrieval (29.2%). Each of the 7 remaining topics accounted for between 1.2% and 10% of the citing articles. Bibliometric articles include numerous studies that have used bibliometric analysis. Information retrieval articles focus on indexing and information retrieval systems. Information behavior articles address information acquisition, information seeking, information use, and information overload. Information service articles mainly refer to reference services and evaluations of reference services. Most knowledge organization articles focus on classification, especially folksonomy. Information science articles cover discussions on information theory, the nature of information science, and information science

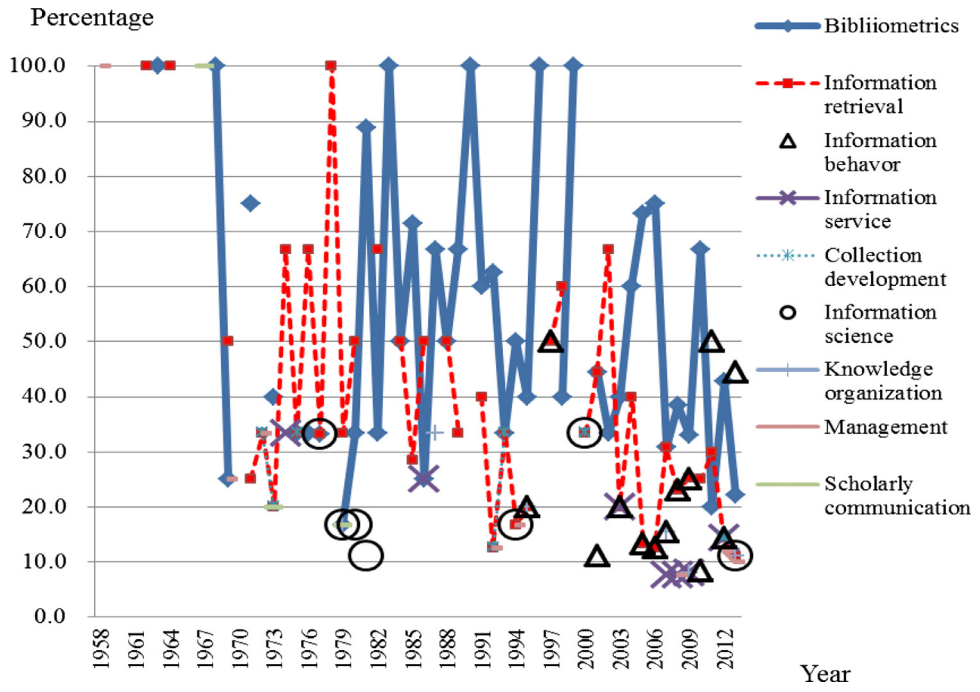

Fig. 2. Distribution of citing article topics by year.

curricula. Management articles refer to library management. Scholarly communication articles involve knowledge communication. This result showed that *HBPLE* is more associated with bibliometrics and information retrieval research than are other research topics.

Fig. 2 shows annual changes in the proportions of citing article topics. Except for 5 years (1959, 1960, 1961, 1965, and 1966), all citing articles represent only a specific topic per year before 1968. However, the citing articles representing three to six topics are observed for the period 2005–2013. The tendency for citing articles to cover diverse topics indicates that *HBPLE* tends to have a higher visibility in certain areas in LIS research. This tendency may also explain why the percentage of the top two topics clearly decreased after 2007. Although obvious fluctuations were observed in the annual proportions of bibliometrics articles and information retrieval articles, bibliometrics research and information retrieval articles dominated in most years. Authors of information behavior articles did not cite *HBPLE* until 1997. Information behavior dominated citing articles after 2011, with the exception of 2012. An increasing trend was evident in the proportion of information behavior articles after 2000.

4.4. Cited concepts

Table 3 shows the 17 concepts extracted from 310 citation contexts. The citing articles mainly cited two concepts. "Zipf's law" is the most cited concept (64.8%), which involves various terms or statements describing the concept. Among 201 citation contexts referring to the concept of "Zipf's law," 52.2% used the term "Zipf's law," 28.4% used other terms, such as "Zipfian distribution," "power law," "hypobolic distribution," and "rank-size law," and 19.4% contained a statement to describe or imply the concept of "Zipf's law." Although Zipf's law is a well-known informetrics law, not all authors have used the formal term "Zipf's law" to refer to the law emphasizing the relationship between word rank and word frequency.

The second most cited concept was the PLE (24.5%), which is derived from Zipf's law. Although the concept of the PLE is the focus of *HBPLE*, the number of citation contexts referring to the PLE was lower than that referring to "Zipf's law." This result ran counter to our assumption that the number of citation contexts referring to the concept of the PLE would be highest. Regarding to the terms used for the concept of the PLE, 84.2% of citation contexts referring to the concept of the PLE used the term "principle of least effort." The remaining citation contexts used a definition or a brief statement to refer to this concept.

The remaining 15 concepts accounted for less than 10.7% of the concepts cited. "HBPLE" referred to Zipf's 1949 publication. "Word distribution" referred to the analysis or discussion of word distribution in human language. "Human behavior" included all types of information-seeking behavior. "Information cycle" referred to the cycle of scholarly communication. "Publication productivity" was equivalent to author productivity. "Rank" referred to the general concept of rank. "Sample size" referred to the quantity of data analyzed in Zipf's studies. "Information nonuse" indicated that the nonuse of information may result from certain situations. "Language analysis" referred to the statistical process of word distribution related to

Table 3
Distribution of cited concepts.

Concept	No. citation contexts	Percentage
Zipf's law	201	64.8
Principle of least effort	76	24.5
HBPLE	9	2.9
Word distribution	6	1.9
Human behavior	2	0.6
Information cycle	2	0.6
Publication productivity	2	0.6
Rank	2	0.6
Sample size	2	0.6
Information nonuse	1	0.3
Language analysis	1	0.3
Lotka's law	1	0.3
Richer effect	1	0.3
R. Y. Chao	1	0.3
Signal information theory	1	0.3
Social physics	1	0.3
Optimization problem	1	0.3
Total	310	100.0

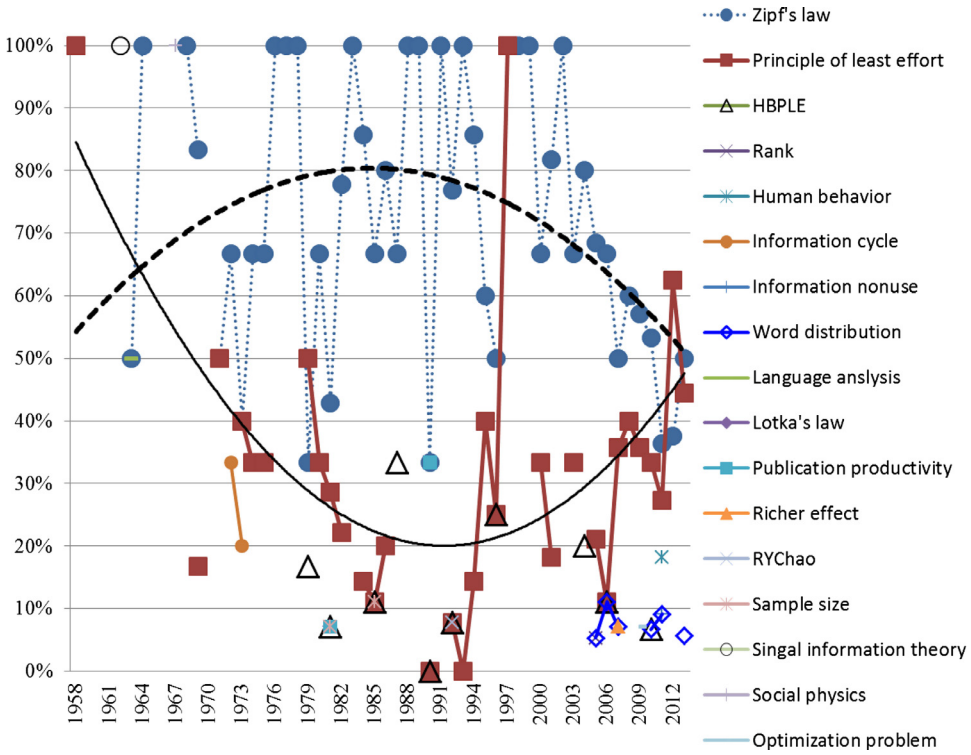

Fig. 3. Changes in the percentage of cited concepts by year.

the types of language analyzed. "Lotka's law" was an informetrics law related to author productivity. "Richer effect" referred to the fact that the rich have a tendency to become richer. "R. Y. Chao" was one of Zipf's friends. "Signal information theory" referred to signal processing and communication. "Social physics" was coined by A. Comte and referred to sociology. "Optimization problem" was used to describe the presentation of possible solutions.

The 17 cited concepts were examined by year. Fig. 3 shows large fluctuations for the two concepts of "Zipf's law" and the PLE; opposing trends appear. A "falling after rising" trend (dashed line) was observed in the concept of "Zipf's law" whereas a "rising after falling" trend (solid line) was evident for the concept of the PLE. No trends were observed for the remaining 15 cited concepts because each cited concept appeared only in specific periods.

Table 4 shows the distribution of the 17 cited concepts in the 12 citing article topics. Researchers of bibliometrics and information retrieval have mainly cited the concept of Zipf's law, followed by the PLE. Authors of bibliometrics research and authors of information retrieval research have comprised approximately 60.7% and 26.4% of citation contexts referring to

Table 4
Distribution of cited concepts based on the topics of citing articles.

Cited concepts	Citing article topics									
	Bib	IR	IB	Iser	CD	Isci	KO	M	SC	Total
Zipf's law	122	61	4		2	4	4	4		201
Principle of least effort	13	18	25	7	4	2	4	2	1	76
HBPLE	8					1				9
Word distribution	3	2	1							6
Human behavior		2								2
Rank	2									2
Information cycle					1				1	2
Publication productivity	2									2
Sample size	1	1								2
Information nonuse			1							1
Language analysis	1									1
Lotka's law	1									1
Richer effect		1								1
R. Y. Chao	1									1
Signal information theory		1								1
Social physics									1	1
Optimization problem	1									1
Total	155	86	31	7	7	7	8	6	3	310

Note: (1) Bib: Bibliometrics. (2) IR: Information retrieval. (3) IB: Information behavior. (4) Iser: Information service. (5) CD: Collection development. (6) Isci: Information science. (7) KO: Knowledge organization. (8) M: Management. (9) SC: Scholarly communication.

Table 5
Distribution of citation functions.

Citation functions	No. of citation contexts	Percentage
Evidence	53	17.1
Comparison	52	16.8
Related studies	45	14.5
History	37	11.9
Relationship	34	11.0
Definitions	30	9.7
Examples	21	6.8
Further reading	14	4.5
Explanation	13	4.2
Terms	9	2.9
Method	2	0.6
Total	310	100.0

the concept of Zipf's law, respectively. The second most cited concept of the PLE has been cited by authors of articles representing 12 topics on information behavior research, followed by bibliometrics research and information retrieval research. Although authors of only nine articles have cited the concept of *HBPLE*; most of the citations were from bibliometrics researchers.

4.5. Citation functions

Table 5 shows the distribution of 11 citation functions based on 310 citation contexts. The top five citation functions were close in percentage and ranged from 11.0% to 17.1%. "Evidence" supporting citing authors' claims was dominant (17.1%), followed by "comparison" (16.8%), which referred to comparing the findings or claims of *HBPLE* and results of other studies. "Related studies" referred to the basic part of a scientific paper presenting the findings of previous studies. "History" referred to the tracing of pioneers or earlier works with the aim of offering background information. "Relationship" referred to the foundation of Zipf's claims; Zipf's claims then became the foundation of other claims. "Definitions" referred to those of concepts that originated from *HBPLE*. "Examples" referred to *HBPLE* or to Zipf's claims being given as examples to clarify the author's ideas. "Further reading" referred to suggestions that readers read *HBPLE* to obtain a deep understanding of certain concepts. Finally, "Method" referred to the provision of detailed information regarding the methodology used in Zipf's studies.

Table 6 shows the distribution of 17 cited concepts in 11 citation functions. The most frequently cited concept was Zipf's law and was mainly used for comparison with other bibliometric laws, whereas the second-most cited concept, the PLE, was mainly used as evidence. The same concepts could have numerous citation functions. In addition, this reflected the complications of citing publications (Bonzi & Snyder, 1991; Liu, 1993).

Table 6
Distribution of cited concepts according to citation functions.

Cited concepts	Citation functions											
	E	C	RS	H	R	D	E	F	Exp	T	M	Total
Zipf's law	29	38	30	27	21	22	17	7	4	5	1	201
Principle of least effort	15	13	8	6	11	7	1	4	8	3		76
HBPLE	2		2	2	2		1					9
Word distribution	3	1	1				1					6
Human behavior			2									2
Information cycle	2											2
Publication productivity			1	1								2
Rank				1								1
Sample size							1		1		1	3
Information nonuse			1									1
Language analysis	1											1
Lotka's law						1						1
Richer effect										1		1
R.Y. Chao	1											1
Signal information theory								1				1
Social physics								1				1
Optimization problem								1				1
Total	53	52	45	37	34	30	21	14	13	9	2	310

Note: (1)E: Evidence. (2)C: Comparison. (3)RS: Related studies. (4)H: History. (5) R: Relationship (6)D: Definitions. (7)E: Examples. (8)F: Further reading. (9)Exp: Explanations. (10)T: Terms. (11) M: Methods.

5. Discussion

According to the growing trend in the annual number of citations that *HBPLE* received from the authors of LIS articles during the study period of 1949–2013, *HBPLE* was identified as a classic work in the field of LIS. The initial results of a search in WoS showed that articles in numerous disciplines, in addition to LIS, have cited *HBPLE*. Computer science articles, followed by physics and LIS articles, cited it most frequently. *HBPLE* was proved to exert an interdisciplinary influence.

The authors of LIS articles did not cite *HBPLE* until the tenth year after it was published, however; this may indicate that classic works may not emerge as essential in the field of LIS immediately after publication. This was consistent with the findings of previous studies that reported that certain highly cited articles did not receive a high number of citations within a few years after publication (Aversa, 1985). Although the rapid increase in the number of publications may have contributed to the increase in the number of citations that *HBPLE* has received, it also leads to a specific concept appearing in various publications. This means that authors can cite a specific concept from various publications, indicating that not all authors cite the same concept from the original work. Therefore, literature growth does not guarantee an increase in the number of citations that a publication receives. An increase in the annual number of citations that a publication receives during a long period is still a simple and useful indicator for identifying a classic work.

According to the results, the concept of Zipf's law was the most frequently cited concept, not the PLE, which is the focus of *HBPLE*. The difference in the number of citations that these two concepts received further demonstrates that the influence of Zipf's law is much stronger than that of the PLE. Although the concept of the PLE is derived from Zipf's law and is also referred to as Zipf's law (Case, 2005; Power, 1998), the two terms differ in meaning when cited. The concept of Zipf's law emphasizes the word "distribution," which is one of the three bibliometric laws. By contrast, the concept of the PLE highlights that people tend to exert the least effort in conducting particular activities. The concept of Zipf's law originated from Zipf's book entitled *The Psychobiology of Language* (PL). PL was published in 1935 and reprinted in 1965, 1968, and 1999. Although a concept that is similar to the PLE appeared in Zipf's 1935 book, it was not fully developed (Chao, 1950). Zipf explained the concept of the PLE in his 1949 book.

Based on the relationship between Zipf's law and the PLE, it may be understood why the concept of Zipf's law is contained in *HBPLE* and why authors have referred to the concept in *HBPLE*. Nevertheless, one compelling finding is that the number of citations PL received is much lower compared with *HBPLE* based on the WoS data. I found that only some authors have cited original publications when referring to a specific concept. For instance, Naranan (1971) cited the concept of Zipf's law in *HBPLE* and the concept of the PLE in PL. In addition, researchers might cite a specific concept from various sources including the original work, or mention a specific concept without citing any related publications. For example, Sun, Shaw, and David (1999) cited the concept of Zipf's law in Zipf's 1935 and 1949 books. Kebede (2004) and White (2001) have mentioned the concept of the PLE without citing any publications containing the concept. This reveals that the influence of a specific concept may be underestimated because authors have not cited the original work.

Citing concepts from original publications shows respect to the authors who have advanced those concepts, however, several reasons may have affected authors' citation of concepts from other studies, including availability or the accessibility of documents (Kim & Sin, 2011; Quigley, Peck, Rutter, & Williams, 2002), language barriers (Murphy, 2003), and other reasons (Bonzi & Snyder, 1991). Furthermore, a few cited concepts were not extracted from the specific content of *HBPLE*.

"HBPLE" was used to confirm the existence of a publication. "Sample size" referred to the sample size used by Zipf in word analysis. A few cited concepts were not submitted by Zipf, such as "Lotka's law," "signal information theory," and "social physics." This implies that literature citing is a complex behavior.

The findings also reveal that *HBPLE* has great influence on authors of bibliometric research, information retrieval, and information behavior. *HBPLE* has the strongest influence on authors of bibliometrics research. Most bibliometrics researchers have cited the concept of Zipf's law in *HBPLE* because they discussed informetrics laws, including Zipf's law. The second-largest group of authors who have cited *HBPLE* was information retrieval researchers. They also most frequently cited the concept of Zipf's law because they focused on index terms. Index terms are related to Zipf's study of word distribution. Information behavior researchers, the third-largest group of authors who have cited *HBPLE*, cited the concept of the PLE most. Only a few information behavior researchers have cited the concept of Zipf's law. Information behavior researchers were identified as the main group citing the core concept originating from *HBPLE*.

Regarding the concept of the PLE, it has been mainly cited by three groups of researchers: information behavior (32.9%), information retrieval (23.7%), and bibliometric research (17.1%). This is consistent with the assumption that information behavior researchers have most frequently cited the concept of the PLE. Information behavior researchers have mainly applied the concept of the PLE to explain the reasons for the selection and use of information. Even information behavior researchers did not pay attention to the concept of the PLE until 1997; the increasing trend in the number of citations may be understood as indicating the improved visibility of *HBPLE* in the area of information behavior. Information retrieval researchers have also mainly used the concept of the PLE to emphasize that users' information seeking behaviors tend to follow the concept of the PLE. Bibliometric researchers have applied the concept of the PLE to various contexts, including word distribution, journal distribution, language use, comparing the PLE with other laws, and information seeking behavior.

Although the number of citation contexts containing the concept of Zipf's law was much higher than the number of citation contexts containing the concept of the PLE, opposing trends appeared in the annual proportions of the two cited concepts. A "falling after rising" trend was identified in the annual percentage of the concept of Zipf's law, whereas a "rising after falling" trend was observed in the annual percentage of the concept of the PLE. These opposing trends have resulted in a decreased difference in the annual percentage between the top two cited concepts since 2006. The changes in influence of the two concepts of Zipf's law and the PLE are worth monitoring. Regarding other cited concepts, their changes in influence were neglected because of a limited number of citations. No trends were formed.

Regarding the citation functions of *HBPLE*, the concept of Zip's law has the highest number of citation functions (11), followed by the concept of the PLE (10). The numbers of citation functions of the two concepts were similar. This demonstrated that the same concepts may play various roles in citation functions. The reasons for authors' citing publications can affect the type of citation function. The concept of Zipf's law has been used mainly for comparison with other informetrics laws or statistical distributions that have resulted from other studies; however, the difference in the number of citation functions among "comparison," "related studies," "evidence," and "history" was small. The main citation function of the concept of the PLE has been used as evidence to support authors' claims. Most authors have emphasized that users tend to expend minimal effort to obtain necessary information. The marginal difference in the numbers of citation functions was identified among "evidence," "comparison," and "relationship." In short, "evidence" and "comparison" were the two main citation functions for the two concepts of Zipf's law and the PLE.

6. Conclusion

HBPLE is a classic work on linguistics that has been cited by authors representing various disciplines, including LIS. This study tracked the influence of *HBPLE* on LIS research. An increasing trend in the annual number of citations indicated that the influence of *HBPLE* on LIS research has not yet diminished. Based on the findings, the influential concepts embedded in *HBPLE* and which topics in LIS research have been affected by *HBPLE* were identified. In particular, I focused on the visibility of the concept of the PLE, which comprises the core of *HBPLE*. Citation context analysis helped to clarify the relationship in terms of content between *HBPLE* and LIS research and to explore the citation functions of *HBPLE*. The most substantial discovery was the fact that the concept of Zipf's law has greater visibility than does the concept of the PLE. A concept highlighted in a publication may not receive substantial attention. This implies that citing behavior is complicated and that various motivations for citing publications also affect the visibility of cited publications. In addition, I observed that the influence of *HBPLE* is visible in numerous areas in LIS research. Information behavior researchers exhibited the most interest in the concept of the PLE whereas bibliometrics researchers paid the most attention to the concept of Zipf's law. Although a close relationship exists between the PLE and Zipf's law, they exert an evidently different influence. Because *HBPLE* has had an interdisciplinary influence and because citing behavior varies among disciplines, I suggest that future studies aim to identify the influence of *HBPLE* on various disciplines. Based on the WoS data, I determined that although *HBPLE* is a classic work on linguistics, it substantially influences numerous fields outside linguistics.

Acknowledgments

This research was supported by a grant from the Ministry of Science and Technology of Taiwan (MOST 103-2410-H-002-172).

References

Anderson, M. H. (2006). How can we know what we think until we see what we said?: a citation and citation context analysis of Karl Weick's *The Social Psychology of Organizing*. *Organization Studies, 27*(11), 1675–1692.

Anderson, M. H., & Sun, P. Y. T. (2010). What have scholars retrieved from Walsh and Ungson (1991)? A citation context analysis. *Management Learning, 41*(2), 131–145.

Austin, B. (2001). Mooers' law: in and out of context. *Journal of the American Society for Information Science and Technology, 52*(8), 607–609.

Aversa, E. S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: a study of the working literature. *Scientometrics, 7*(3-6), 383–389.

Bierbaum, E. G. (1990). A paradigm for the '90 s. *American Libraries, 21*(1), 18–19.

Bigdeli, Z. (2007). Iranian engineers' information needs and seeking habits: an agro-industry company experience. *Information Research, 12*(2). http://www.informationr.net/ir/12-2/paper290.html.

Bonzi, S., & Snyder, H. W. (1991). Motivations for citation: a comparison of self-citation and citation to others. *Scientometrics, 21*(2), 245–254.

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45–80.

Bronstein, J. (2008). The application of cost-benefit and least effort theories in studies of information seeking behavior of humanities scholars: the case of Jewish studies scholars in Israel. *Journal of Information Science, 34*(2), 131–144.

Carvalho, P., Diniz-Filho, J. S. F., & Bini, L. M. (2005). The impact of Felsentein's "Phylogenies and the comparative method" on evolutionary biology. *Scientometrics, 62*(1), 53–66.

Case, D. O. (2005). Principle of least effort. In K. E. Fisher, S. Erdelez, & L. E. F. McKechnie (Eds.), *Theories of information behavior* (pp. 289–292). Medford, N.J.: Information Today.

Case, D. O., & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science, 51*(7), 635–645.

Chang, Y. W. (2013a). A comparison of citation contexts between natural sciences and social sciences and humanities. *Scientometrics, 96*(2), 535–553.

Chang, Y. W. (2013b). The influence of Taylor's paper, Question-negotiation and information-seeking in libraries. *Information Processing & Management, 49*(5), 983–994.

Chao, Y. R. (1950). Review of the book Human behavior and the principle of least effort: an introduction to human ecology. *Language, 26*(3), 394–401.

Chen, Y. P., & Liang, G. (2000). Behavior of library user and law of save effort. *Journal of Human Urban Construction College, 9*, 75–77.

Chrzastowski, T. E. (1995). Do workstations work too well? An investigation into library workstation popularity and the 'principle of least effort'. *Journal of the American Society for Information Science, 46*(8), 638–641.

Chrzastowski, T. E. (1999). E-journal access: the online catalog (856 field), Web lists, and 'The principle of least effort'. *Library Computing, 18*(4), 317–322.

Chu, H. (2001). Intellectual activities and influences of Belver C. Griffith: a citation perspective. *Scientometrics, 51*(3), 481–488.

Coleman, S. R., & Salamon, R. (1988). Kuhn's structure of scientific revolutions in the psychological journal literature, 1969–1983: a descriptive study. *The Journal of Mind and Behavior, 9*(4), 415–446.

Cummings, L. D. (1977). Voluntary strategies in environmental movement: recycling as cooptation. *Journal of Voluntary Action Research, 6*(3/4), 153–160.

Duan, X. B. (2004). Principle of minimum effort in a university library. *Journal of Mianyang Normal University, 6*, 60–62.

Frost, C. (1979). The use of citations in literary research: a preliminary classification of citation functions. *Library Quarterly, 49*, 399–414.

Furner, J. (2003). Little book, big book: before and after little science, big science: a review article, part II. *Journal of Librarianship and Information Science, 35*(3), 189–201.

Garfield, E. (1980). Citation measures of the influence of Robert K. Merton. *Transactions New York Academy of Sciences, 39*(1), 61–74.

Gratch, B. G. (1990). Exploring the principle of least effort and its value to research. *College and Research Libraries News, 51*(8), 727–728.

Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics, 41*(3), 497–518.

Kebede, G. (2004). The information needs of end-users of Sub-Saharan Africa in the digital information environment. *International Information and Library Review, 36*(3), 273–279.

Kim, C. (1982). Retrieval language of social sciences and natural sciences: a statistical investigation. *Journal of the American Society for Information Science, 33*(1), 1–7.

Kim, K. S., & Sin, S. C. J. (2011). Selecting quality sources: bridging the gap between the perception and use of information sources. *Journal of Information Science, 37*(2), 178–188.

Lipetz, B. A. (1965). Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators. *American Documentation, 16*(2), 81–90.

Liu, M. (1993). Progress in documentation the complexities of citation practice: a review of citation studies. *Journal of Documentation, 49*(4), 370–408.

McCain, K. W., & Salvucci, L. J. (2006). How influential is Brooks' law? A longitudinal citation context analysis of Frederick Brooks' The Mythical Man-Month. *Journal of Information Science, 32*(3), 277–295.

McCain, K. W., & Turner, K. (1989). Citation context analysis and aging patterns of journal articles in molecular genetics. *Scientometrics, 17*(1), 127–163.

Mooers, C. N. (1996). Mooers' law or, why some retrieval systems are used and others are not. *Journal of the American Society for Information Science, 23*(1), 22–23.

Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science, 5*(1), 86–92.

Murphy, J. (2003). Information-seeking habits of environmental scientists: a study of interdisciplinary scientists at the environmental protection agency in Research Triangle Park, North Carolina. *Issues in Science and Technology Librarianship, 38*. Retrieved from http://www.istl.org/03-summer/refereed.html.

Naranan, S. (1971). Power law relations in science bibliography: self-consistent interpretation. *Journal of Documentation, 27*(2), 83–97.

Oppenheim, C., & Renn, S. P. (1978). Highly cited old papers and the reasons why they continue to be cited. *Journal of the American Society for Information Science, 29*(5), 225–231.

Power, D. M. W. (1998). Applications and explanations of Zipf's law. In D. M. W. Powers (Ed.), *NeMLaP3/CoNLL98: new methods in language processing and computational natural language learning* (pp. 151–160). Stroudsburg, PA: Association for Computational Linguistics. Retrieved from http://acl.ldc.upenn.edu/W/W98/W98-1218.pdf.

Quigley, J., Peck, D. R., Rutter, S., & Williams, E. M. (2002). Making choices: factors in the selection of information resources among science faculty at the University of Michigan. Results of a survey conducted July–September, 2000. *Issues in Science and Technology Librarianship, 34*. Retrieved from http://www.istl.org/02-spring/refereed.html.

Rapoport, A. (1957). Comments: the stochastic and the 'teleological' rationales of certain distributions and the so-called principle of least effort. *Behavioral Science, 2*(2), 147–161.

Richard, D. M., & Pysek, P. (2008). Fifty years of invasion ecology: the legacy of Charles Elton. *Diversity of Distributions, 14*(2), 161–168.

Shadish, W. R., Tolliver, D., Gray, M., & Gupta, S. K. S. (1995). Author judgments about works they cite: three studies from psychology journals. *Social Studies of Science, 25*(3), 477–498.

Small, H. (1982). Citation context analysis. In B. J. Dervin, & M. J. Voight (Eds.), *Progress in communication science* (pp. 287–310). Norwood, NJ: Ablex Pub.

Spiegel-Rosing, I. (1977). Science studies: bibliometric and content analysis. *Social Studies of Science, 7*(1), 97–113.

Sun, Q., Shaw, D., & David, C. H. (1999). A model for estimating the occurrence of same-frequency words and the boundary between high- and low-frequency words in texts. *Journal of the American Society for Information Science, 50*(3), 280–286.

Thomaz, S. M., Michelan, T. S., Carvalho, P., & Bini, L. M. (2010). The influence of "Homage to Santa Rosalia" on aquatic ecology: a scientometric approach. *Hydrobiologia, 653*, 7–13.

Tohidinasab, F., & Jamali, H. R. (2013). Why and where wikipedia is cited in journal articles? *Journal of Scientometric Research, 2*(3), 231–238. Retrieved from http://jscires.org/temp/JSciRes23231-3309863_091138.pdf.
Tsay, M. Y. (2009). Citation analysis of Ted Nelson's works and his influence on hypertext concept. *Scientometrics, 79*(3), 451–472.
Wang, X. N. (2000). Accessibility of minimum power-saving principle to information retrieval system. *Information Science, 18*, 135–136.
Wang, Z. J. (2001). The application of minimum power-saving principle to library services. *Library and Information Services, 6*, 79–81.
Wei, H. (2005). Application of principle of least effort in cataloguing. *Journal of the Library Sciencein Jiangxi, 4*, 21–22.
White, H. (2001). Authors as citers over time. *Journal of the American Society for Information Science and Technology, 52*(2), 87–108.
White, M. D., & Wang, P. (1997). A qualitative study of citing behavior: contribution, criteria, and metalevel documentation concerns. *Library Quarterly, 67*(2), 122–154.
Xu, Y. C., Tan, B. C. Y., & Yang, L. (2006). Who will you ask? An empirical study of interpersonal task information seeking. *Journal of the American Society for Information Science and Technology, 57*(12), 1666–1677.
Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, Mass.: Addison-Wesley Press.