An evaluation of information behaviour studies through the Scholarly Capital Model

Faramarz Soheili,1* Ali Akbar Khasseh,1 Afshin Mousavi Chelak,1 and Mohammad Tavakolizadeh-Ravari1,2

1Department of Library and Information Science, Payame Noor University, Tehran, Iran
2Department of Library and Information Science, Yazd University, Yazd, Iran

ORCID:
F. Soheili: 0000-0003-2581-7052
A.A. Khasseh: 0000-0001-5664-4671
A. Mousavi Chelak: 0000-0002-5614-1334
M. Tavakolizadeh-Ravari: 0000-0003-1649-671x

*Corresponding author: Faramarz Soheili
Email: f_soheili@pnu.ac.ir

Abstract
Influence and capital are two concepts used to evaluate scholarly outputs, and these can be measured using the Scholarly Capital Model as a modelling tool. The tool looks at the concepts of connectedness, venue representation, and ideational influence using centrality measures within a social network. This research used co-authorships and h-indices to investigate authors who have published papers in the field of information behaviour between 1980 and 2015 as extracted from Web of Science. The findings show a relationship between the authors’ connectedness and the venue (journal) representation. It could be seen that the venue (journal) influences the chance of citation, and equally, the prestige (centrality) of authors probably raises the citations of the journals. The research also shows a significant positive relationship between the venue representation and ideational influence. This means that a research work that is published in a highly cited journal will find more visibility and will receive more citations.

INTRODUCTION

The quality of scientific publications is of great interest to researchers in the field of scientometrics. Despite the availability of a wide variety of metrics, there remains no perfect method for measuring the quality of a publication. It is why Truex, Cuellar, and Takeda (2009) and Truex, Cuellar, Takeda, and Vidgen (2011), for example, propose to transform the question from ‘Is this scholar’s work of sufficient quality in his/her field?’ to ‘Is this scholar sufficiently influential in his/her field?’

The related literature recommends using a variety of metrics to evaluate the researcher’s impact and criticizes the use of only a single metric (Bornmann, Mutz, & Daniel, 2008; Mingers, Macri, & Petrovici, 2012). Scholarly influence and capital have been defined as two concepts in the evaluation of scholarly outputs with a support of theories and measures (Egghe, 2006; Freeman, 1979; Hirsch, 2005; Cuellar, Vidgen, Takeda, & Truex, 2016a). Cuellar et al. (2016a) worked on a multidimensional model called ‘Scholarly Capital Model’ in order to model the scholarly influence. They assert:

The Scholarly Capital Model is based on the idea that, when evaluating a scholar’s research capabilities as part of making hiring, promotion, and tenure decisions, organizations should consider three things: (1) the extent to which other scholars take up the scholar’s work (ideational influence), (2) who the scholar works with (connectedness), and (3) how well the scholar publishes in venues of the scholar’s field (venue representation).

These factors constitute the three dimensions of the Scholarly Capital Model – the dimensions that are necessary for scholars to impact their field and consequently provide prestige for their institutions.
Key points

- The concept of influence (and use of the Scholarly Capital Model) offers a theoretical background for evaluating scientific outputs.
- Authors with higher social interactions cooperate with more authors and increase the quality of their works.
- The popularity of a journal is influenced by the reputation of its authors, and together, this increases citation.
- The authors’ ability to communicate (connectedness) significantly increases the influence of their ideas (ideational influence).
- The popularity of the publication (venue) and the reputation of the authors increase interactively.

The simultaneous use of the h-index, contemporary h-index, and g-index provides a profile of the scholars’ ideational influence and allows for the comparison of research in a certain field. Cuellar et al. (2016a) also consider the common indicators of co-authorship analysis to obtain the measure of connectedness, including degree, betweenness, and closeness centralities, which are parameters of social network analysis (SNA). Stringer (2009) argues that author centrality is an indication for research performance. Yan and Ding (2009) and Li, Liao, and Yen (2013) describe a relationship between the betweenness centrality and the number of citations. The closeness centrality shows how quickly an author is able to access other authors in the collaboration network since a node with a higher closeness centrality needs to take fewer steps on average to reach the other nodes of the network. This potentially indicates an association between the closeness centrality value of authors and the number of citations that they receive. The third element of the Scholarly Capital Model is venue representation, which increases through the publication of papers within the core journals of a given field. Together, these theories and indicators support the concept of scientific influence on a particular research area. In our opinion, although there may be some limitations in applying the Scholarly Capital Model to academic publishing (Crowston, 2016), it still provides a suitable tool for analysing influence and evaluating the scholarly capital of researchers in a specific field (Cuellar, Vidgen, Takeda, & Truex, 2016b).

The structural equation model is recognized as one of the methods for multivariate analysis and is the best suited for this research since this model inherently studies different main and sub-variables.

In this analysis, the Scholarly Capital Model is applied to the field of information behaviour. Information behaviour is a subfield of library and information science (LIS; Sugimoto, Li, Russell, Finlay, & Ding, 2011), whose history goes back five decades. Fisher and Julien (2009) believe that the research on information behaviour is maturing and expanding. This provided the initial stimulus for applying the Scholarly Capital Model to obtain a macroscopic picture of the key works.

Sugimoto et al. (2011) argue that different theories have emerged in the field of information behaviour over the past 50 years. Therefore, the associated issues of this field comprise a key part of discussions within LIS. Fisher and Julien (2009) also report the rising number of studies in this field over the years. This gradual development is an incentive for conducting macro and comprehensive research on this topic.

Information behaviour is inherently related to psychology and uses quantitative and qualitative methodologies, such as grounded theory and critical incident (Soheili & Khasseh, 2015). Jamali (2013) shows that theories of information behaviour are largely related to LIS and are also rooted in sociology, communications, psychology, management, education, and computer science. This consequently results in collaboration among the authors of different disciplines. Persson, Glänzel, and Danell (2004) found that the interdisciplinary and collaborative works receive more citations. The quality and quantity of research is greatly dependent on collaboration among individuals with high social interactions. High-quality studies presumably receive more citations, and in turn, the number of citations increases the h-family indices.

The current study aims to determine the scientific influence of authors in the field of information behaviour through the Scientific Capital Model because an influence model encloses the structural features of complex cooperation among the authors and identifies their roles in a co-authorship network.

The main question concerns the generalization of this model in an already developed area, apart from its origin. The research examines three hypotheses regarding this question:

- The more the social influence of the authors in the field of information behaviour, the more their ideational influence.
- The more the social influence of the authors in the field of information behaviour, the more their venue influence.
- The more the venue influence of the authors in the field of information behaviour, the more their ideational influence.

LITERATURE REVIEW

Julien (1996) conducted an early scientometric analysis of information behaviour, studying the bibliographic properties of papers published on information needs between the years 1990 and 1994. Key areas analysed include the paper type, source, and authors. This work was repeated for the years 1984–1989 and 1995–1998 by Julien and Duggan (2000) and for 1999–2008 by Julien, Pecoskie, and Reed (2011). The latter work concentrates on authorship, paper type, methodology, and journal type. A prominent finding of this work shows an increase in the interdisciplinarity of the associated literature.

McKechnie, Goodall, Lajoie-Paquette, and Julien (2005) studied 155 English-language human information behaviour (HIB) papers published between 1993 and 2000 in the six prominent
LIS journals. They found that the HIB literature was cited mostly (81.5%) by LIS authors and that these works cited a wide range of disciplines, including engineering, psychology, education, and medicine. They also found that 36.0% of the papers were cited for general purposes, 28.5% within the findings section, 25.3% for the theories, and a few within the methodology (6.0%). They concluded that ‘HIB literature has kept its significant impact on other disciplines’.

McKechnie et al. (2005) analysed the citations of 155 important papers published in the field of information behaviour between 1993 and 2000. They conducted a co-citation analysis on references of citing articles. The co-citation map reveals the existence of a central nucleus made up of the most commonly cited authors. This study also revealed a sub-discipline related to the Internet and electronic communication.

Podsakoff, MacKenzie, Podsakoff, and Bachrach (2008) studied the papers of 30 journals in the field of management that were published between the years 1981 and 2004. They recognized 100 highly cited universities and 150 highly cited researchers in the field. Their findings also show that few universities receive the most citations. This work considers the number of received citations as the scientific influence.

Chang (2011) compared the characteristics of research articles on information needs and information seeking published from 1962 to 2009. The papers included those that were indexed in the Web of Science. Their analysis shows that the majority of papers were published in medical journals, and the co-authorship is low (one and two authors on average).

Jamali (2013) visualized the works related to the field of information behaviour theories. He aimed to determine to what extent a set of 51 associated theories came from the fields outside LIS. The Web of Science database was searched for the bibliographic and citation information of these theories. He established a citation analysis and bibliographic coupling as two methods. The analysis and the visualization process were applied by applications such as Pajek, HistCite, Sci2 Tool, and VOS-Viewer. The results show that most of the theories come from LIS. The few others are from disciplines such as sociology, psychology, and computer science. Bibliographic coupling showed a cohered map within the network of theories. The results express the expansion of the field of LIS along with the HIB.

González-Teruel, González-Alcaide, Barrios, and Abad-García (2015) determined the contribution of researchers in creating the current body of knowledge on information behaviour. They found that a small number of authors are the most productive ones and publish regularly.

Soheili and Khasseh (2015) studied the historical origin of information behaviour literature through the reference publication years spectography (RPYS). They determined the yearly distribution of references among this field. The results show, respectively, three and six main shifts for this field during the 19th and 20th centuries. The results also show that this field is inspired by psychology, along with the quantitative and qualitative methods such as grounded theory and critical incident.

Other studies have also been performed on information behaviour, but they have not considered any scholarly capital, which is regarded as a social parameter. Li-Ping (2010) deals with SNA, which is a social quantitative method. He studied the co-authorship map of 1,889 papers on information behaviour that were indexed in Library and Information Science Abstracts. He found seven top authors: Spink A, Savolainen R, Nicolas D, Wilson TD, Ellis D, Kuhlthau GC, and Marchionini G.

The advance of the scholarly capital concept attracted more researchers: Truex, Cuellar, Vidgen, and Takeda (2011) focused their studies on scientific influence and scholarly capital. Scholarly influence research began in 2008 with Southern Association for Information Systems (SAIS), The Americas Conference on Information Systems (AMCIS), and International Conference on Information Systems (ICIS) papers, which showed how to measure the scientific influence by means of the Hirsch family indices (h-indices). The concept of connectedness was subsequently added to some studies (Truex et al., 2011; Takeda, 2012). Finally, the concept of venue representation was added by Cuellar et al. (2016a).

Cuellar et al. (2016a) introduced the Scholarly Capital Model for measuring scholarly influence/capital, which consists of three dimensions: ideational influence, connectedness, and venue representation. They established a list from Clark, Warren, and Au (2009) to delimit the information systems area and applied this model on the Google Scholar citations for this field (Cuellar et al., 2016a).

**RESEARCH METHODOLOGY**

The research is descriptive and studies the scholarly capital of authors using a model called the Scholarly Capital Model. The study was performed on the papers that are published in the field of information behaviour between 1980 and 2015. The Web of Science database was established as the source for extracting the data of these papers. The database was queried using the following search strategy, which has previously been employed by González-Teruel et al. (2015) and Soheili and Khasseh (2015):

- **TOPIC:** (information behavi* OR TOPIC: (information need*) OR TOPIC: (information seek*) OR TOPIC: (information us*) OR TOPIC: (information search* behavi*)
- **TOPIC:** (information shari* behavi*).

This query was limited to searching the papers categorized under ‘information science and library science’ in the Social Science Citation Index (SSCI) and the Conference Proceedings Citation Index-Social Science and Humanities (CPCI-SSH). The other limitation was on the document type. Three types of English documents were considered: journal articles, review papers, and conference proceedings. The search returned 3,493 records, which were then parsed. Some extracted authors’ names were written differently. These names were standardized, and their count of contributions to the dataset was calculated. They were divided
into two groups of prolific and non-prolific contributors based on the Pareto’s 20/80 method. Applying this method showed that the authors with three and more contributions should be considered to be in the upper level. In total, 481 authors fulfilled this condition. A symmetric matrix was created from the number of authors’ co-occurrence in publishing papers. This matrix is a necessity for determining the three centralities of authors through SNA. In addition, the \( h \)-family indices \((h, hc, \text{ and } g)\) were provided for each author. Authors’ centralities were determined through the UCINET computer programme (a software package for the analysis of social network data) and the \( h \)-family indices through the Bibexcel. The matrix was copied to the Excel sheet of UCINET and was converted to the correct format for the programme. The three centralities (degree, betweenness, and closeness) were calculated separately to the correct format for the programme. The three centralities were determined through the UCINET computer programme and the \( h \)-family indices through the Bibexcel. The matrix was copied to the Excel sheet of UCINET and was converted to the correct format for the programme. The three centralities (degree, betweenness, and closeness) were calculated separately through the ‘measures and centralities’ submenu of that programme. Finally, the records were imported into the Bibexcel to determine the authors’ \( h \)-index. The other two indices are also rooted in this index. \( G \)-index is based on Egghe’s (2006) formula:

\[
A \text{ set of papers has a } g \text{-index } g \text{ if } g \text{ is the highest rank such that the top } g \text{ papers have, together, at least } g^2 \text{ citations. This also means that the top } g + 1 \text{ papers have less than } (g + 1)^2 \text{ papers (Egghe, 2006).}
\]

The \( hc \)-index is based on the formula that was suggested by Sidiropoulos et al. (2007):

\[
Sc (i) = \gamma^s (Y \text{ (now)} - Y (i) + 1) - \delta^s |C(i)|
\]

where \( Y (i) \) is the publication year of article \( i \) and \( C (i) \) are the articles citing the article \( i \). If we set \( \delta = 1 \), then \( Sc (i) \) is the number of citations that the article \( i \) has received, divided by the ‘age’ of the article. Following Sidiropoulos, Katsaros, and Manolopoulos (2007), the value for the coefficient \( \gamma \) was set as 4.\(^1\)

Venue influence reflects the source in which a work is published. A researcher increases this influence by publishing his or her works in focal sources of a given field. The more focal a source, the more influential the researcher is. There is consequently a direct relationship between the level of visibility of a scientific work and its venue influence (Cuellar et al., 2016a).

The determination of influences is based on proximity matrices. The networks that are the result of such matrices allow for the calculation of degree, closeness, and betweenness centralities through some computer programmes like UCINET. The venue and the social influences need the same matrices for calculation of centralities, but the contents of their rows and columns differ. The social influence focuses on the authors and the relationship between them, but the determination of the venue influence is possible through an affiliation network, which is twofold: actors and events. In such a network, the links are not between the actors but between the affiliation and the events. The actors of an affiliation network are tied because they belong in the same group (Sasson, 2008). These networks are the results of a matrix, which consists of authors as rows and journals as columns. The matrix is then analysed to determine the authors who are closer to the focal source publications of a given field.

Calculations of these two current indices were performed using the functions of Microsoft Excel, SPSS and the LISREL were used for the correlation test and the structured equation modelling. The validity of data was confirmed through convergent and divergent validity criteria (Table 1).

Convergent analysis is the degree of confidence that a trait is well measured by its indicators. According to Fornell and Larcker’s (1981) criterion, the convergent validity of the measurement model can be assessed by the average variance extracted (AVE). AVE measures the level of variance captured by a construct.

### Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Average variance extracted</th>
<th>Factors loading</th>
<th>Construct reliability ( \rho_c &gt; 0.7 )</th>
<th>Cronbach’s ( \alpha ) reliability coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectedness</td>
<td>0.52</td>
<td>—</td>
<td>0.76</td>
<td>0.75</td>
</tr>
<tr>
<td>Betweenness</td>
<td>—</td>
<td>0.63</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Closeness</td>
<td>—</td>
<td>0.63</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Degree</td>
<td>—</td>
<td>0.89</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Venue representation</td>
<td>0.70</td>
<td>—</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>Venue betweenness</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Venue closeness</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Venue degree</td>
<td>—</td>
<td>0.91</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ideational influence</td>
<td>0.93</td>
<td>—</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>( h )-index</td>
<td>—</td>
<td>0.98</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( g )-index</td>
<td>—</td>
<td>0.96</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( hc )-index</td>
<td>—</td>
<td>0.94</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

\(^1\)[Correction added on 16 October 2017 after first online publication: The twelve lines of text in the left-hand column immediately above Table 1 were originally published in the incorrect order. This has now been rectified.]
versus the level due to measurement error; values above 0.7 are considered very good, whereas the level of 0.5 is acceptable.

Divergent analysis helps to establish construct validity by demonstrating that the given construct is different from the other constructs in the study. For each reflective factor, AVE square root must be more than that factor’s correlation with the other factor of the model (Choua & Chen, 2009).

The reliability was determined by Cronbach’s alpha coefficient along with the composite reliability coefficient (Fornell & Larcker, 1981). The coefficient of Cronbach’s alpha implicitly assumes that components have the same weight, whereas the coefficient of composite reliability references the actual factor loadings. The resulting coefficients for Cronbach’s $\alpha$ were over 0.7, which is the minimum value for confirming reliability. Tables 1 and 2 show the results for the validity and the reliability of research.

RESULTS

The structural equation model was applied to find out the linear relations between the hidden (unobservable) and the obvious (observable) variables. This model is a combination of the measuring model (confirmation factor analysis) and the structural model (regression or path analysis). SMART-PLS was used for data analysis. It is a tool for multiple variables analysis, measuring direct, indirect, and interactive effects and testing moderating effects (Bagozzi & Fornell, 1982). The path model of PLS estimates the score of each hidden variable and evaluates the moderating role of hidden variables in the pass model (Esposito-Vinzi, Chin, Henseler, & Wang, 2010). The programme tests the moderating effects in the evaluation step.

Figures 1 and 2 show the results of analysis by SMART-PLS.

Figure 1 shows the coefficients of the structural equation model for social and venue influences regarding degree, closeness, and betweenness centralities and for ideational influence regarding $h$, $g$, and $hc$-indices. The lambdas of hidden and external variable for ideational influence are, respectively, 0.982, 0.968, and 0.948 for $h$, $g$, and $hc$-index. The same $\lambda$s for venue influence are, respectively, 0.694, 0.635, and 0.651 for degree, closeness, and betweenness. These $\lambda$s for venue influence are, respectively, 0.91, 0.731, and 0.876 for degree, closeness, and betweenness centralities.

The model reveals that the coefficients for direct and indirect effects of social influence on ideational influence, respectively, are 0.204 and 0.401. In other words, 60% of ideational influence changes are impacted by social influence. The coefficients also show that the direct effect of venue influence on ideational influence is 0.683, which means that 68% of ideational changes are impacted by venue influence. The result also shows that social influence impacts the venue influence with a coefficient equal to 0.588. It means that 59% of the venue influence changes are impacted by the social one.

The $\chi^2$ test is used to determine the fitting model for covariance-based structural equation modelling. Cv-communality and cv-redundancy address another method for this objective. Cv-communality measures the quality of the model for calculating each block. Its positive result guarantees the quality of the measuring model (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005). The positive result displayed in Figure 3 confirms the quality of the model for the current research.

Table 3 summarizes the results of the hypotheses tests. The table shows a positive effect of connectedness on the venue representation ($r = 0.588$, $t = 10.738$). The venue representation significantly affects the ideational influence ($r = 0.683$, $t = 15.760$). It shows the positive effect of connectedness on the ideational influence ($r = 0.242$, $t = 4.576$).

Shrout and Bolger (2002) believe that ‘Mediation is said to occur when a causal effect of some variable X on an outcome Y is explained by some intervening variable M. The authors recommend that with small to moderate samples, bootstrap methods be used to assess mediation. Bootstrap tests are powerful because they detect that the sampling distribution of the mediated effect is skewed away from 0’. The mediation analysis was performed in the way suggested by Shrout and Bolger (2002). The direct, indirect, and total effects of intra-genic variables were also measured (Table 4). The results confirm that the ‘venue representation’ plays a mediation role between the connectedness and the ideational influence since the connectedness directly affects the ‘venue representation’, and this variable substantially affects the ideational influence.

The goodness of fit was tested for outer and inner models. The outer one is determined by the communality mean, which represents the variance of references (in percent) through the corresponding constructs. The other one is based on the $R^2$ value, which is common for AMOS (a statistical software which stands for analysis of a moment structures), EQS (one of the most widely used Structural Equation Modeling programs), and LISREL (a proprietary statistical software package used in structural equation modeling) to test the structural models. Table 5 shows that the average sharing values are above 0.05, which complies with the minimum requirement of confirmation (Lee, Park, Baek, & Lee, 2008). It reveals that the model is also able to describe the construct through venue representation ($R^2 = 0.346$) and ideational influence ($R^2 = 0.718$).

DISCUSSION AND CONCLUSION

This research studied papers in the field of information behaviour through a model called the Scholarly Capital Model. The findings show correlations among the variables in the model. As a result, social influence positively affects the ideational influence. This means that the authors with higher social interactions cooperate with more authors and increase the quality of their work. They consequently find a better position in a co-authorship network and receive more centrality.

The review of literature shows a positive correlation between the centralities and receiving citations (Li et al., 2013; Yan & Ding,
FIGURE 2  The results of t-value analysis. Note. If $t > 1.96$, then the effect is significantly positive; if $-1.96 < t < 1.96$, then there is no significant effect; if $t < -1.96$, then the effect is significantly negative (Chin, 2003).

FIGURE 1  The structural coefficients of the structural equation model.
The popularity of a journal is influenced by the reputation of its authors. This reputation increases the chances of a journal receiving more citations. Journals and authors interactively increase each other’s position. In other words, an author with higher centrality is more recognized, and his or her contribution in a journal increases its centrality within the network of journals.

The highly centralized authors are interested in cooperation with the other focal authors in the field who sometimes are the editors, editorial boards, or referees. This potentially increases their chance for publishing in popular journals. As such, social influence impacts publication venue, as confirmed by Cuellar et al. (2016a).

The further finding of the current research shows the positive impact of the venue on the ideational influence. This concerns the relationship between a study’s visibility and its citation potential, demonstrating that the increased visibility of papers in popular journals increases their likelihood of being cited (Cuellar et al., 2016a). Truex et al. (2009) similarly found that publishing in central journals of a field increases the ideational influence.

This research also shows that influential authors (defined by high connectedness and/or h-indices) tend to write in journals that have similar high connectedness within the journals network. This means that papers published in high-impact journals are more likely to be cited, and journals that publish the papers of prestigious authors, in turn, receive more citations.

The current research shows that the model for scientific influence is applicable to information behaviour, including LIS areas. The suitability of this model for other fields may provide a model for assessing the authors’ influence. This is an area for subsequent study.

ACKNOWLEDGEMENTS

This research was funded by the Payame Noor University (PNU).
REFERENCES


---

Registration opens 3 April
**ALPSP Annual Conference and Awards 2018**
Beaumont Estate, Old Windsor, UK • 12-14 September

[www.alpsp.org/Conference](http://www.alpsp.org/Conference)  #alpsp18