The Use of Theory in Information Science Research

Karen E. Pettigrew
Assistant Professor, The Information School, University of Washington, Box 352930, Seattle, WA, 98195-3152. E-mail: kpettigr@u.washington.edu

Lynne (E.F.) McKechnie
Assistant Professor, Faculty of Information & Media Studies, The University of Western Ontario, Middlesex College, London, Ontario, N6A 5B7 Canada. E-mail: mckechnie@uwo.ca

We report on our findings regarding authors’ use of theory in 1,160 articles that appeared in six information science (IS) journals from 1993–1998. Our findings indicate that theory was discussed in 34.1% of the articles (0.93 theory incidents per article; 2.73 incidents per article when considering only those articles employing theory). The majority of these theories were from the social sciences (45.4%), followed by IS (29.9%), the sciences (19.3%), and humanities (5.4%). New IS theories were proposed by 71 authors. When compared with previous studies, our results suggest an increase in the use of theory within IS. However, clear discrepancies were evident in terms of how researchers working in different subfields define theory. Results from citation analysis indicate that IS theory is not heavily cited outside the field, except by IS authors publishing in other literatures. Suggestions for further research are discussed.

Background

“Having a theory” is today the mark of research seriousness and respectability. Theory is, of course, convenient, and helps to organize and communicate unwieldy data and simplify the terrible complexities of the social world, matters that may well be more important to the field than whether or not a given theory is true of false. (Van Maanen, 1998, p. xxix).

It is a well-known fact that IS lacks good theories. (Hjørland, 1998, p. 607)

[Theories] may be expressed or represented in written and graphical form. They may well inspire and guide practical achievements of a concrete form. Yet a theory remains a mental construct. A “good” theory is one that matches well our perception of whatever the theory is about. The closer the match, the better the theory is. (Buckland, 1991, p. 19)

Working with conceptual frameworks and empirical research has never been easy. (Chatman, 1996, p. 205)

According to the philosophy of science, the use of theory by scholars in their research is a hallmark of their discipline’s academic maturity (cf. Brookes, 1980; Hauser, 1988). Moreover, disciplines require theories that originate from within to attain recognition as an independent field of scientific inquiry. In other words, if fields such as information science (IS) are to delineate their disciplinary boundaries and build a central body of knowledge, then they require their own theoretical bases for framing research problems, building arguments, and interpreting empirical results. Even advocates of the opposing “philosophy of knowledge” view such as Wersig (1993), who argue that information science is a prototypical postmodern science that seeks to develop strategies for solving or dealing with problems that have been caused by classical sciences and technologies (as opposed to searching for complete understanding of how the world works), agree that a unified building of interconceptual underpinnings or a “conceptual navigation system” is, nonetheless, necessary for the field of IS.

But what is theory? Some basic definitions include: “a set of explanatory concepts” (Silverman, 1993, p. 1), “a statement or group of statements about how some part of the world works—frequently explaining relations among phenomena” (Vogt, 1993, p. 232), “an internally connected and logically consistent proposition about relationships among phenomena” (Odi, 1982, p. 313), “a systematic explanation for the observed facts and laws that relate to a particular aspect of life” (Babbie, 1992, p. 55), “a unified, systematic explanation of a diverse range of social phenomena” (Schwandt, 1997, p. 154), and “generalizations which seek to explain relationships among phenomena” (Grover & Glazier, 1986, p. 228). The Oxford English Dictionary defines “theory” as a “supposition or system of ideas explaining something, especially one based on general principles independent of the facts, phenomena, etc., to be explained (e.g., atomic theory, theory of gravitation, evolution); . . . exposi-
tion of the principles of a science, etc., (the theory of music); (Math.) collection of propositions to illustrate principles of a subject (probability theory, theory of equations)” (Sykes, 1982, p. 1109). In their 1985 ARIST chapter “Principles and Theories in Information Science,” Boyce and Kraft defined “principle” as “a single fundamental law, generally an empirically regularity based on continued observation” such that a “theory” “incorporates a body of such principles and suggests new principles that can be tested as hypotheses, both to increase knowledge and to invalidate or to strengthen the theory itself” (Boyce & Kraft, 1985, p. 154). More recently, Buckland (1991, p. 18) defined theory “in the broad sense of a description or explanation of the nature of things, not in the more restricted sense, used in some sciences, of denoting fundamental laws formally stated and falsifiable.” According to Hjørland (1998, p. 607), IS theory as “a theoretical explanation of information systems efficiency, of user behavior, of the function of different search agents such as descriptors, titles, and so on.” Regarding models in information behavior research, Wilson (1999, p. 250) said “a model may be described as a framework for thinking about a problem and may evolve into a statement of the relationship among theoretical propositions.”

The need for greater use of theory as a conceptual basis in IS research has been repeatedly voiced in the literature with respect to the discipline as a whole (e.g., Boyce & Kraft, 1985; Feehan, Gragg, Havener, & Kester, 1987; Grover & Glazier, 1986; Hjørland, 1998; Templeton, 1994) and its subfields such as information retrieval (e.g., Spink, 1997), human–computer interaction (e.g., Shackel, 1997), and information behavior (e.g., Vakkari, 1997; Zweizig & Dervin, 1977). But despite the many recent bibliometric studies and content analyses that have been conducted on the IS literature in terms of faculty productivity and citation and acknowledgment patterns (e.g., Atkins, 1988; Cronin & Overfelt, 1994; Davis & Cronin, 1993; Garland, 1990, 1991; Harter & Hooten, 1990; Harter, Nisonger, & Weng, 1993; Hayes, 1983; Kim, 1992; Kumpulainen, 1991; Mularski, 1991; Persson, 1994; Pettigrew & Nicholls, 1994; Varlejs & Dalrymple, 1986; White & McCain, 1998), few researchers have focused specifically on the role of theory in IS research, i.e., the importation, growth, and use of theory within the field. Instead, the prevalence of theory has been contained to a singular facet and not explored in depth. In their multivariable content analyses, researchers have tended to code articles simply according to whether or not a conceptual framework was employed with little attention to which theories from which fields were cited or how they were used therein.

The results from content analyses that were conducted periodically on the journal literature since 1950, however, overwhelming suggest that the vast majority of IS research is atheoretical. Peritz (1980, p. 252), for example, reported that through the years 1950 to 1975 only 14% of the articles sampled were theoretical in the analytical sense of employing a “mathematical, linguistic, logical, or other philosophical methodology.” She accredited the sudden increase in theoretical papers published between 1965 and 1975 to “the introduction of mathematical modelling, systems analysis, and linguistic analysis into the research literature” (p. 257).

Nour (1985) replicated Peritz’s study with the 1980 journal literature, and reported that while 21.2% of articles used theory in an analytical sense, less than 3% of the articles were about “information science theory.” In their investigation of the 1984 journal literature, Feehan, Gregg, Havened, and Keister (1987) reported that only 13% of the 123 research articles sampled from 91 IS journals either discussed/applied theory in the study design or “attempted to formulate theories or principles which can provide a theoretical basis” for IS. Although Feehan et al. coded these theories in terms of six IS categories (e.g., general, communication theory, information retrieval) and included the use of theory from other disciplines, they did not report the final breakdowns.

With respect to the 1985 literature, Järvelin and Vakkari (1990) found that theory was used in only 10% of the 449 empirical studies that they examined from 37 core journals. While these explanatory investigations were most frequently used in research on information seeking, professions, and scientific communication, Järvelin and Vakkari concluded that “there is little attempt in IS to discover the regularities prevailing in the research area [and that] this deficiency makes the formulation of theories more difficult” (p. 409). They further remarked:

LIS theories, as we know, are usually vague and conceptually unclear, and basic concepts have not been defined (cf. Poole, 1985; Schrader, 1986). Therefore the scarcity of conceptual analysis must be regarded as a grave deficiency. Although empirical studies and articles containing verbal argumentation and critique may contain conceptual analysis, they do not make up for the lack of more general analysis on basic concepts of the field. Conceptual analysis, too, should be more frequent, in order to clarify the unhoed rows in concepts in LIS (p. 415).

They pointed to the library and information services paradigm that they believe underlies most IS research (at least in 1985) as the reason for the lack of theory employed therein. According to Järvelin and Vakkari, “this paradigm typically has made little use of such traditional scientific approaches as foundations and conceptual analysis, or of scientific explanation and theory formulation. This may be due to the fact that the discipline was born out professional practice and is therefore intimately connected with its problems” (p. 415).

More recently, Julien (1996) focused on the information needs and uses literature published from 1990–1994. She reported an increase in theory use in that 28% of the 165 articles sampled were theoretically grounded, meaning they were “based on a coherent and explicit framework of assumptions, definitions, and propositions that, taken together, have some explanatory power” (p. 56). However, in a re-
luted study, Julien and Duggan (2000) reported that only 18.3% of the 300 research studies sampled from 1984–1989 and 1995–1994 were theoretically based and thus, their “great concern that such a small proportion of literature is based on theory.” But their results do suggest that theory use may be increasing, given Järvelin and Vakkari’s (1993) finding that only 6 to 8% of research articles on information-seeking sampled for the years 1965, 1975, and 1985 employed a theoretical framework. Julien (1996) also reported a significant relationship between author type and theoretical grounding where “researchers were more likely to theoretically ground their publications, whether these were research studies or not, than practitioners,” and also between journal type and theoretical grounding of the literature where “scholarly journals were more likely to publish theoretical articles than professional journals” (p. 59). She reported no significant relationship between author type and whether articles were typified as research in nature.

In short, while content analysts agree that the use of theory in IS research is consistently low (anywhere between 10 to 21% with a slight increase in studies that focus on information behavior) and members of the IS community in general recognize a need—albeit from different perspectives—for increased use of theory in our work, little is known about the current use and uses of theory in IS research. Although researchers have responded to the inherent issues surrounding theory in IS research in different ways [Pettingrew (1997), Vakkari (1998), and Vakkari & Kuokkanen (1997), e.g., advocated the use of structuralist tools from sociology for measuring theoretical growth and activity and for adapting theories from other disciplines in IS research, while Grover and Glazier (1986) proposed a model for theory building in IS, and various others, most notably Frohmann (1992), Hjøland and Albrechtsen (1995), and Schrader (1984, 1986), have debated the implications of different conceptualizations of the intellectual domain of information science], basic work is required on the role of theory in IS research and on how it is specifically being used.

**Current Study**

As a response to past criticism of the nonscientific nature of IS research, one may hypothesize that greater emphasis is being placed on conducting basic research that reflects the structure of a mature academic discipline with its own theoretical underpinnings. Thus, one may further hypothesize that theories are being born in IS and that a central core of knowledge, unique to the discipline, is taking shape. But what are these theories? Where did they come from? How are they being used in current research? And, given the diverse topical interests of researchers in IS, do the answers to these questions vary? The purpose of our study was to address these questions by analyzing how IS authors use theory in their published work, and by examining how these IS theories are used outside the field.

**Methodology**

In the current study, these questions were investigated by conducting a content analysis of 1,160 articles that appeared from 1993 to 1998 in six journals:

1. *Information Processing and Management* (IP&M; six issues per year)
3. *Journal of Documentation* (JDOC; quarterly)
4. *Journal of Education for Library and Information Science* (JELIS; quarterly)
5. *Library and Information Science Research* (LISR; quarterly)
6. *Library Quarterly* (LQ; quarterly)

These journals were chosen because they contain peer-reviewed articles that cover most areas of research interest within IS, namely: indexing, information retrieval, human–computer interaction, bibliometrics, information behavior, information policy, history, library services, management, and education. All articles, except for book reviews and news items such as conference reports, were coded for the authors’ use of theory.

Each article was initially coded in terms of the primary author’s affiliation (as listed in the article, i.e., private sector, government, academic department), broad subject matter, and type of article, while the theories cited therein were coded in terms of whether they originated from IS, the sciences, social sciences, or humanities, and where they were mentioned in the article (i.e., title, abstract, or main text). Each theory was counted only once per article. Our codebook is appended. If the author(s) discussed a theory that originated within IS or proposed such a theory, then the theory was tagged for additional analysis to determine its academic origins (Kuhlthau’s theory of the information search process, for example, can be traced to Kelly’s personal construct theory from the social sciences).

Before discussing our findings, it is necessary to explain how we defined “theory” in this study. Although varied

<table>
<thead>
<tr>
<th>Coding category</th>
<th>Intercoder agreement rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author affiliation</td>
<td>29/30 (96.7%)</td>
</tr>
<tr>
<td>Author background</td>
<td>30/30 (100.0%)</td>
</tr>
<tr>
<td>Article subject</td>
<td>25/30 (83.3%)</td>
</tr>
<tr>
<td>Article type</td>
<td>27/30 (90.0%)</td>
</tr>
<tr>
<td>Theory</td>
<td></td>
</tr>
<tr>
<td>a. Present: Agreement</td>
<td>40/41 (97.6%)</td>
</tr>
<tr>
<td>b. For articles with theory(ies)</td>
<td></td>
</tr>
<tr>
<td>i. Where mentioned:</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>10/10 (100.0%)</td>
</tr>
<tr>
<td>Abstract</td>
<td>9/10 (90.0%)</td>
</tr>
<tr>
<td>Text</td>
<td>10/10 (100.0%)</td>
</tr>
<tr>
<td>ii. Where arise from (sciences, social sciences, humanities)</td>
<td>10/10 (100.0%)</td>
</tr>
</tbody>
</table>
definitions appear in the literature, the results from our initial coding quickly indicated that there was no singular definition that would encompass all the varied uses of the term in the articles that we are examining. For example, articles about information retrieval differed vastly in its use of the term from those that focused on information behavior, history, or information policy. Moreover, heeding postmodernist writings by Budd (1995), Day (1996), Dick (1995), Hjørland (1998), and Wersig (1993), we wanted to avoid stigmatizing or contributing to the normalization of IS theory by operationalizing it solely along positivist lines. Therefore, for definitional coding purposes we adhered to the following rule:

Consider a “theory” as identified if the author(s) describes it as such in the article (applicable to established or proposed theories) or uses such key terms as “conceptual” (including its variations, e.g., conceptualization), “framework,” “grounded,” or “underpinnings” to describe an idea/view or approach as such.

If an author, for example, discussed Lotka’s Law (1926) using the term “theory” or in context of another term that was consistent with our rule, then Lotka’s Law was coded as theory. Conversely, if an author mentioned a theory such as Dervin’s sense-making (1992), which is frequently used as a conceptual framework in studies on information behavior, but the author did not discuss it in context of the rule, then its occurrence was not counted as theory. We made no attempt to distinguish among theories that could be categorized as paradigms, grand theory, formal theory, or substantive theory (as explained by Grover & Glazier, 1986). Our reason was that at this stage of our investigation we wanted to operationalize “theory” as broadly as possible so that we could capture its widest range of use in the literature. In short, we took an author perspective in our coding such that we identified theories on the bases of the authors’ descriptive context. Beyond enabling us to identify authors’ varied uses of the term, this approach was consistent with that used in studies of an exploratory nature.

Upon completing the content analysis of the six journals, we derived a list (n = 11) of the most prominent IS theories (i.e., most frequently cited as theory). The citing articles (from the content analysis stage) were examined to identify the bibliographic references that were used by the authors when citing these theories and to establish one or a group of “authority” references for each theory. Then, during November 1999, we searched the authority reference(s) for each theory using ISI’s three citation indices (Science Citation Index, Social Science Citation Index, and the Arts and Humanities Citation Index) to determine what impact particular IS theories have had on IS research and on research conducted in other disciplines. Following standard citation analysis procedures, we discounted all self-citations and duplicate records. ISI’s journal subject category was used to identify the primary discipline of each publication in which an IS theory was cited.

Intercoder Results

A sample of 30 articles, five randomly chosen from each of the six journals, was coded by an independent coder. The final rate of agreement for all coding decisions was 94.7% (see Table 1), suggesting that the coding scheme was reliable.

Findings

The findings from our content and citation analyses are presented in the following order:

1. Basic characteristics of the articles examined in terms of author affiliation, primary subject matter, and type of article.
2. Theory deployment: the frequency with which theories were identified in the articles in terms of journal title,
originating discipline, as well as the article’s primary subject matter, type of article, and article section.

3. Citation analysis results: the frequency that a subset of identified theories were cited both inside and outside IS.

**Basic Characteristics**

We analyzed the content of 1,160 articles (Table 2), the majority of which were from *JASIS* (40.9%) and *IP&M* (27.0%). Whereas the majority (54.7%) of articles listed the primary author as affiliated with an IS school or a library, there were clear indications that individuals from other disciplines as well as from the private sector have contributed to the literature (Table 3). In terms of primary subject matter, as shown in Table 4, the overall top six subjects were: information retrieval (32.7%), bibliometrics (10.9%), indexing, abstracting, cataloging and classification (9.3%), education and pedagogy (8.2%), human information behavior (8.2%), and library services (7.4%). Reports of empirical research were the most frequent type of article across all journals (59.3%) (Table 5).

**Theory Deployment**

Of the 1,160 articles examined, 34.1% (*n* = 396) incorporated theory in either the title, abstract, or text such that a total 1083 incidents of theory use were identified. Overall, each article contained an average of 0.93 theories (SD 2.06). This ranged from 0 (764 cases) to 29 (one case). However, if one considers only those 397 articles that used theory, then the average use increases to 2.73 (SD 2.73) (Table 5).

**Table 4. Articles by subject (%).**

<table>
<thead>
<tr>
<th>Primary subject</th>
<th>IP&amp;M</th>
<th>JASIS</th>
<th>JDOC</th>
<th>JELIS</th>
<th>LISR</th>
<th>LQ</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 314</td>
<td>n = 347</td>
<td>n = 105</td>
<td>n = 96</td>
<td>n = 95</td>
<td>n = 76</td>
<td>n = 1160</td>
<td></td>
</tr>
<tr>
<td>Bibliometrics</td>
<td>6.1</td>
<td>14.8</td>
<td>19.0</td>
<td>2.1</td>
<td>10.5</td>
<td>7.9</td>
<td>10.9</td>
</tr>
<tr>
<td>Education</td>
<td>1.3</td>
<td>3.2</td>
<td>—</td>
<td>84.4</td>
<td>—</td>
<td>9.2</td>
<td>9.2</td>
</tr>
<tr>
<td>HCI</td>
<td>1.6</td>
<td>3.0</td>
<td>3.8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2.0</td>
</tr>
<tr>
<td>History</td>
<td>1.3</td>
<td>3.2</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>13.2</td>
<td>2.8</td>
</tr>
<tr>
<td>Human info. behavior</td>
<td>3.5</td>
<td>7.2</td>
<td>10.5</td>
<td>3.1</td>
<td>24.2</td>
<td>17.1</td>
<td>8.2</td>
</tr>
<tr>
<td>Indexing</td>
<td>15.3</td>
<td>11.0</td>
<td>6.7</td>
<td>—</td>
<td>1.1</td>
<td>—</td>
<td>9.3</td>
</tr>
<tr>
<td>Info. policy</td>
<td>0.3</td>
<td>3.2</td>
<td>2.9</td>
<td>1.0</td>
<td>4.2</td>
<td>10.5</td>
<td>2.8</td>
</tr>
<tr>
<td>Info. retrieval</td>
<td>58.0</td>
<td>30.2</td>
<td>31.4</td>
<td>—</td>
<td>16.8</td>
<td>6.6</td>
<td>32.7</td>
</tr>
<tr>
<td>Info. technology</td>
<td>3.8</td>
<td>7.2</td>
<td>5.7</td>
<td>—</td>
<td>1.1</td>
<td>—</td>
<td>4.6</td>
</tr>
<tr>
<td>Library services</td>
<td>1.6</td>
<td>7.4</td>
<td>8.6</td>
<td>2.1</td>
<td>24.2</td>
<td>15.8</td>
<td>7.4</td>
</tr>
<tr>
<td>Management</td>
<td>—</td>
<td>1.3</td>
<td>1.0</td>
<td>—</td>
<td>7.4</td>
<td>3.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Sch. communication</td>
<td>2.3</td>
<td>5.3</td>
<td>6.7</td>
<td>—</td>
<td>2.1</td>
<td>2.6</td>
<td>3.7</td>
</tr>
<tr>
<td>General IS/other</td>
<td>5.1</td>
<td>3.4</td>
<td>2.9</td>
<td>6.3</td>
<td>7.4</td>
<td>13.2</td>
<td>5.0</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Table 5. Articles by type (%).**

<table>
<thead>
<tr>
<th>Article type</th>
<th>IP&amp;M</th>
<th>JASIS</th>
<th>JDOC</th>
<th>JELIS</th>
<th>LISR</th>
<th>LQ</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 314</td>
<td>n = 347</td>
<td>n = 105</td>
<td>n = 96</td>
<td>n = 95</td>
<td>n = 76</td>
<td>n = 1160</td>
<td></td>
</tr>
<tr>
<td>Descriptive paper</td>
<td>13.1</td>
<td>17.5</td>
<td>19.0</td>
<td>18.8</td>
<td>1.1</td>
<td>3.9</td>
<td>14.3</td>
</tr>
<tr>
<td>Discourse analysis</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1.0</td>
<td>—</td>
<td>3.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Empirical research</td>
<td>64.3</td>
<td>57.2</td>
<td>49.5</td>
<td>49.0</td>
<td>82.1</td>
<td>50.0</td>
<td>59.3</td>
</tr>
<tr>
<td>Historical paper</td>
<td>1.9</td>
<td>4.9</td>
<td>5.7</td>
<td>2.1</td>
<td>1.1</td>
<td>18.4</td>
<td>4.5</td>
</tr>
<tr>
<td>Literature review</td>
<td>2.2</td>
<td>3.6</td>
<td>4.8</td>
<td>2.1</td>
<td>6.3</td>
<td>—</td>
<td>3.2</td>
</tr>
<tr>
<td>Mathematical modeling</td>
<td>19.6</td>
<td>3.8</td>
<td>4.8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>4.6</td>
</tr>
<tr>
<td>Method paper</td>
<td>1.3</td>
<td>2.7</td>
<td>1.9</td>
<td>4.2</td>
<td>7.4</td>
<td>5.3</td>
<td>2.9</td>
</tr>
<tr>
<td>Theory paper</td>
<td>2.2</td>
<td>3.8</td>
<td>3.8</td>
<td>4.2</td>
<td>1.1</td>
<td>5.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Verbal Argumentation</td>
<td>4.8</td>
<td>6.3</td>
<td>10.5</td>
<td>18.8</td>
<td>1.1</td>
<td>13.2</td>
<td>7.3</td>
</tr>
<tr>
<td>Other</td>
<td>0.6</td>
<td>0.2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.3</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
were well represented in JASIS, IP&M, and JDOC articles, few theories were drawn from the humanities by authors in any of the journals.

Authors who wrote articles dealing with IS in general (which were sometimes entirely conceptual in nature), overwhelmingly tended to use the most theories per article (2.59), followed by authors writing about information behavior (1.99) and information policy (1.94) (Table 8). When only those articles that cited are considered, papers about information policy rank highest (5.17), followed by general IS (4.17), human information behavior (3.37), and information technology (3.27) (Table 9).

Not surprisingly, under “type of article,” theory papers had the highest incidence of theory use (5.47). Literature reviews (1.65), and papers that employed verbal argumentation (1.52) or discourse analysis (1.50) ranked second, third, and fourth, respectively (Table 10). When only those articles that contain theory are considered, these rankings vary slightly: theory papers (5.78), literature reviews (4.36), method papers (3.64), and verbal argumentation (3.49). Most noticeably, method papers rose in rank to third place (Table 11).

Although most theories (99.2%) were mentioned in the text, only 9.4% were mentioned in the title and 19.9% in the abstract (Table 12). Surprisingly, a few authors mentioned a theory in the title or abstract without any reference in the text, while others referred to a theory in two or three of these sections. Of further interest are that 95 authors did not provide bibliographic references for 151 or 13.9% of the theories mentioned.

Citation Analysis Results

As revealed from our content analysis work, the most frequently cited theories (in alphabetical order) were as follows: Bates’ Berry Picking, Belkin et al.’s Anomalous States of Knowledge, Dervin’s Sense Making, Ellis’s Information Seeking, Harter’s Psychological Relevance, Ingwerson’s Cognitive IR Theory, Kuhlthau’s Information Search Process, Salton’s Vector Space Model, Schamber et al.’s Theory of Relevance, Taylor’s Information Needs and Negotiation, and Wilson’s Situation Relevance.

For the citation analysis, the authority reference(s) for each of these 11 IS theories was searched using ISI’s citation indices. As shown in Table 13, the searches resulted in 2,098 citations, 79.9% of which were in IS publications. For the remaining 20.1% of the citations, the majority were from the sciences, followed by the social sciences. Thus, it appears that approximately one-quarter of the works to which the IS theories were attributed are being cited outside the field. However, a closer examination reveals that these data are heavily skewed: the aggregate number of citations for Dervin and Salton is 918, thus accounting for 43.8% of the total citations. When the citations for these outliers are removed, only 8.9% of the remaining citations occur outside IS, and many of these citations were written by IS authors. In short, the IS theories most frequently cited outside IS are actually the work of two individual authors whose primary affiliations are not IS, and who publish broadly across multiple fields. Although one may argue that these two theories are not IS because of their creators’ non-IS affiliation/background, one may also argue that they are indeed IS theories because they were published in the IS literature and focus on IS problems.

Discussion

The results of our content analysis of the IS journal literature suggest an increase in the amount of theory used by authors. However, the discrepancy between our finding

<table>
<thead>
<tr>
<th>Discipline</th>
<th>IP&amp;M (n = 252)</th>
<th>JASIS (n = 497)</th>
<th>JDOC (n = 90)</th>
<th>JELIS (n = 62)</th>
<th>LISR (n = 92)</th>
<th>LQ (n = 90)</th>
<th>Total (n = 1083)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS</td>
<td>36.4 (92)</td>
<td>30.6 (152)</td>
<td>40.0 (36)</td>
<td>4.8 (3)</td>
<td>25.0 (23)</td>
<td>20.0 (18)</td>
<td>29.9 (324)</td>
</tr>
<tr>
<td>Humanities</td>
<td>2.8 (7)</td>
<td>6.2 (31)</td>
<td>3.3 (3)</td>
<td>—</td>
<td>5.4 (5)</td>
<td>13.3 (12)</td>
<td>5.4 (58)</td>
</tr>
<tr>
<td>Sciences</td>
<td>31.8 (80)</td>
<td>20.3 (101)</td>
<td>17.8 (16)</td>
<td>8.1 (5)</td>
<td>3.3 (3)</td>
<td>4.5 (4)</td>
<td>19.3 (209)</td>
</tr>
<tr>
<td>Social sciences</td>
<td>29.0 (73)</td>
<td>42.9 (213)</td>
<td>38.9 (35)</td>
<td>87.1 (54)</td>
<td>66.3 (61)</td>
<td>62.2 (56)</td>
<td>45.4 (492)</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
that 34.1% of articles included theory with results reported in earlier studies (anywhere from 10–21%) may lie in differences with the six journals we chose for examination (largely research-oriented journals), our sampling time frame, and in how the term “theory” was operationalized. In other words, our approach of identifying theories according to authors’ use of the term may have led to the identification of more incidents than if we had solely used definitions commonly found in textbooks on research methods. However, as others such as Buckland (1991) have suggested, a singular definition of theory may not be adequate for discussing the varied ways in which theory is used within the subfields of IS. This observation was confirmed in our study, as we discovered that some colleagues include bibliometric laws and mathematical distributions in their conceptions of theory [as Triolo & Bao (1993, p. 158), remark: “Bradford rankings and other informetric projections appeal to the theorist”], while others emphasize the need for theory to be able to explain or predict social phenomena. Still others stress the criterion that theories should be “scientific” in the sense of having “quantitative definitions, measures and relationships” (Yovitz & Kleye, 1993, p. 353). Differences in how members of the IS community perceive theory is also readily evident from how journal editors classify articles in their tables of contents. For example, in the September 1997 issue of JASIS (volume 48, issue 9) the editors classify the following articles under the heading “theory”:

- What is a “document”? (Buckland)
- Relevance: The whole story (Mizzaro)
- Bradford’s: From the classical bibliometric “law” to the more general stochastic models (Oluic-Vukovic)

Although it is possible that conceptual differences regarding the nature of theory may be due to the different disciplinary backgrounds of researchers in IS, other themes emerged from our data that suggest a general confusion exists about theory even within subfields. Numerous examples came to light during our analysis in which an author would simultaneously refer to something as a theory and a method, or as a theory and a model, or as a theory and a reported finding. In other words, it seems as though authors, themselves, are sometimes unsure about what constitutes theory. Questions even arose regarding whether the author to whom a theory was credited would him or herself consider his or her work as theory. Some authors, for example, credited Marchionini’s (1992) and Marchionini and Schneiderman’s (1988) work on hypertext and information retrieval

---

**TABLE 8. Incidents of theory use per article by subject (n = 1160).**

<table>
<thead>
<tr>
<th>Primary subject</th>
<th>Incidents per article</th>
</tr>
</thead>
<tbody>
<tr>
<td>General IS/other (n = 58)</td>
<td>2.59</td>
</tr>
<tr>
<td>Human info. behavior (n = 95)</td>
<td>1.99</td>
</tr>
<tr>
<td>Info. policy (n = 32)</td>
<td>1.94</td>
</tr>
<tr>
<td>Info. technology (n = 53)</td>
<td>.92</td>
</tr>
<tr>
<td>Info. retrieval (n = 379)</td>
<td>.88</td>
</tr>
<tr>
<td>Management (n = 17)</td>
<td>.76</td>
</tr>
<tr>
<td>Library services (n = 86)</td>
<td>.65</td>
</tr>
<tr>
<td>Sch. communication (n = 43)</td>
<td>.58</td>
</tr>
<tr>
<td>Education (n = 107)</td>
<td>.55</td>
</tr>
<tr>
<td>Bibliometrics (n = 127)</td>
<td>.55</td>
</tr>
<tr>
<td>Indexing (n = 108)</td>
<td>.52</td>
</tr>
<tr>
<td>HCI (n = 23)</td>
<td>.48</td>
</tr>
<tr>
<td>History (n = 32)</td>
<td>.28</td>
</tr>
<tr>
<td>Overall</td>
<td>.93</td>
</tr>
</tbody>
</table>

---

**TABLE 9. Incidents of theory use per article employing theory by subject (n = 397).**

<table>
<thead>
<tr>
<th>Primary subject</th>
<th>Incidents per article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info. policy (n = 12)</td>
<td>5.17</td>
</tr>
<tr>
<td>General IS/other (n = 36)</td>
<td>4.17</td>
</tr>
<tr>
<td>Human info. behavior (n = 56)</td>
<td>3.37</td>
</tr>
<tr>
<td>Info. technology (n = 15)</td>
<td>3.27</td>
</tr>
<tr>
<td>Sch. communication (n = 9)</td>
<td>2.78</td>
</tr>
<tr>
<td>Education (n = 22)</td>
<td>2.68</td>
</tr>
<tr>
<td>Info. retrieval (n = 135)</td>
<td>2.47</td>
</tr>
<tr>
<td>HCI (n = 5)</td>
<td>2.20</td>
</tr>
<tr>
<td>Management (n = 6)</td>
<td>2.17</td>
</tr>
<tr>
<td>Library services (n = 26)</td>
<td>2.15</td>
</tr>
<tr>
<td>Bibliometrics (n = 36)</td>
<td>1.94</td>
</tr>
<tr>
<td>Indexing (n = 33)</td>
<td>1.70</td>
</tr>
<tr>
<td>History (n = 6)</td>
<td>1.50</td>
</tr>
<tr>
<td>Overall</td>
<td>2.73</td>
</tr>
</tbody>
</table>

---

**TABLE 10. Incidents of theory use per article by type (n = 1160).**

<table>
<thead>
<tr>
<th>Type of article</th>
<th>Incidents per article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory paper (n = 38)</td>
<td>5.47</td>
</tr>
<tr>
<td>Literature review (n = 37)</td>
<td>1.65</td>
</tr>
<tr>
<td>Verbal argumentation (n = 85)</td>
<td>1.52</td>
</tr>
<tr>
<td>Discourse analysis (n = 4)</td>
<td>1.50</td>
</tr>
<tr>
<td>Method paper (n = 34)</td>
<td>1.18</td>
</tr>
<tr>
<td>Mathematical modeling (n = 53)</td>
<td>.87</td>
</tr>
<tr>
<td>Empirical research (n = 688)</td>
<td>.75</td>
</tr>
<tr>
<td>Descriptive paper (n = 166)</td>
<td>.37</td>
</tr>
<tr>
<td>Historical paper (n = 52)</td>
<td>.35</td>
</tr>
<tr>
<td>Other (n = 3)</td>
<td>.33</td>
</tr>
<tr>
<td>Overall</td>
<td>.93</td>
</tr>
</tbody>
</table>

---

**TABLE 11. Incidents of theory use per article employing theory by type (n = 397).**

<table>
<thead>
<tr>
<th>Type of article</th>
<th>Incidents per article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory paper (n = 36)</td>
<td>5.78</td>
</tr>
<tr>
<td>Literature review (n = 14)</td>
<td>4.36</td>
</tr>
<tr>
<td>Method paper (n = 11)</td>
<td>3.64</td>
</tr>
<tr>
<td>Verbal argumentation (n = 37)</td>
<td>3.49</td>
</tr>
<tr>
<td>Discourse analysis (n = 3)</td>
<td>2.0</td>
</tr>
<tr>
<td>Mathematical modeling (n = 25)</td>
<td>1.84</td>
</tr>
<tr>
<td>Empirical research (n = 234)</td>
<td>2.19</td>
</tr>
<tr>
<td>Descriptive paper (n = 27)</td>
<td>2.26</td>
</tr>
<tr>
<td>Historical paper (n = 9)</td>
<td>2.0</td>
</tr>
<tr>
<td>Other (n = 1)</td>
<td>1.0</td>
</tr>
<tr>
<td>Overall</td>
<td>2.73</td>
</tr>
</tbody>
</table>
retrieval as “theories of information seeking,” which may not be consistent with the tradition or theoretical body of knowledge developed within the subfield of information behavior.

Authors’ citation practices have added to this general confusion in several ways. First, authors are not consistent in their use of theory throughout an article. One would expect, for example, for authors to discuss a theory in the article’s text if they had mentioned it in the title or abstract. Conversely, including the name of the theory in the article’s title or abstract would aid researchers when retrieving these articles in databases such as LISA (Library and Information Science Abstracts). Second, authors frequently do not include bibliographic references for the theories that they mention, and seem to assume that all readers are familiar with such IS concepts as “citation theory,” “berry picking,” “the information search process (ISP),” and “the theory of human information seeking and information retrieval,” and with ideas from outside the field such as “graph theory,” “chaos theory,” and “equiavailability theory.” A third problem with authors’ citation practices is that they sometimes refer the reader to introductory textbooks, review articles, or reports of other research that discuss the theory instead of citing the original source(s). In addition to providing full and accurate bibliographic references, it would be helpful if authors included some information in their text that identifies the discipline in which the theory originated. Another problem that we identified was the use of variant names for the same theory, which can make it difficult for the reader to understand exactly what the author intended. Ingwersen’s work, for example, was alternatively called his: theory of knowledge structures, theory of the interaction IR process, cognitive theory of IR, cognitive viewpoint of IS, theory of cognitive space, and cognitive theory of polyrepresentation. One way to support increased use of theory within IS research would be to adopt good citation practices. Clearly identifying a theory, naming it in the abstract, and providing one or more references to primary sources for the theory would be helpful for other IS scholars with little or no knowledge of the theory.

Despite the confusion caused by authors’ use and citation of theory, our results to date suggest that theory may be playing a stronger role than previously observed in the IS research literature. Moreover, through our content analysis we recorded references to well over 100 distinct theories that we recognized as having been born in IS. The frequently repeated citing of Kuhlthau’s information search process, Bates’s berry-picking, Belkin, Oddy, and Brooks’s

TABLE 12. Where theory mentioned (%).

<table>
<thead>
<tr>
<th>Where mentioned</th>
<th>IP&amp;M n = 252</th>
<th>JASIS n = 497</th>
<th>JDOC n = 90</th>
<th>JELIS n = 62</th>
<th>LISR n = 92</th>
<th>LQ n = 90</th>
<th>Total n = 1083</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>10.3 (26)</td>
<td>5.0 (25)</td>
<td>1.1 (9)</td>
<td>27.4 (17)</td>
<td>15.2 (14)</td>
<td>12.2 (11)</td>
<td>9.4 (102)</td>
</tr>
<tr>
<td>Abstract</td>
<td>17.6 (45)</td>
<td>13.9 (68)</td>
<td>22.2 (20)</td>
<td>41.9 (26)</td>
<td>32.6 (30)</td>
<td>30.0 (27)</td>
<td>19.9 (216)</td>
</tr>
<tr>
<td>Text</td>
<td>100.0 (252)</td>
<td>98.4 (489)</td>
<td>100.0 (90)</td>
<td>100.0 (62)</td>
<td>98.9 (91)</td>
<td>100.0 (90)</td>
<td>99.2 (1074)</td>
</tr>
</tbody>
</table>

n = number theories cited in journal across all articles.

TABLE 13. Citation analysis findings.

<table>
<thead>
<tr>
<th>Theory</th>
<th>Total</th>
<th>IS</th>
<th>SCI</th>
<th>SOC SCI</th>
<th>HUM</th>
<th>MED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bates’ berry picking</td>
<td>210</td>
<td>181</td>
<td>23</td>
<td>4</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>Belkin’s ASK</td>
<td>243</td>
<td>218</td>
<td>16</td>
<td>7</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>Dervin’s sense making</td>
<td>304</td>
<td>274</td>
<td>5</td>
<td>24</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>Ellis’s information seeking</td>
<td>48</td>
<td>47</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Harter’s psychological relevance</td>
<td>64</td>
<td>61</td>
<td>2</td>
<td>1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ingwersen’s cognitive IR model</td>
<td>144</td>
<td>130</td>
<td>—</td>
<td>14</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Kuhlthau’s ISP</td>
<td>164</td>
<td>154</td>
<td>3</td>
<td>6</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>Salton’s vector space model</td>
<td>614</td>
<td>328</td>
<td>262</td>
<td>18</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Schamber et al.’s relevance</td>
<td>80</td>
<td>71</td>
<td>6</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Taylor’s information need</td>
<td>178</td>
<td>167</td>
<td>8</td>
<td>1</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>Wilson’s situational relevance</td>
<td>49</td>
<td>46</td>
<td>3</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Total</td>
<td>2098</td>
<td>1677 (79.9%)</td>
<td>329 (15.7%)</td>
<td>78 (3.7%)</td>
<td>3 (.14%)</td>
<td>11 (.52%)</td>
</tr>
</tbody>
</table>

IS—information science
SCI—sciences
SOC SCI—social sciences
HUM—humanities
MED—medicine and health sciences.
ASK, and Harter’s psychological relevance as theory suggests that a unique theoretical core is developing within IS. The identification of 71 new theories that were proposed in the papers that we examined reinforces the notion of a growing body of theory unique to IS. However, the results from the citation analysis suggest that IS theory is not well cited outside the field, except by IS authors who publish in non-IS journals. Although Dervin’s sense making and Salton’s vector space model account for 91% of the citations outside IS, their scores are not surprising considering that their backgrounds, and primary affiliations are not in IS. In social networking terms, these authors are connectors: by conducting IS research and publishing it in different fields, they increase outsiders’ awareness of IS theory and research.

Although the results reported in this paper represent a first step towards understanding the use of theory in IS research, substantial work remains to be undertaken. In addition to conducting similar analyses on other key IS journals and examining other time periods (it will be interesting to track, e.g., how some of the 71 proposed theories are used by IS authors in future articles), further work is needed to determine how different subgroups of the IS community define and use theory in their research, how theories from other disciplines have contributed to the growth of IS theory, and on how theory from IS is shaping work afield. In-depth examination of how authors are using theory also is needed: while many authors referred to multiple theories in their work, it was beyond the scope of the current study to determine if these theories were used in substantive ways or if they were simply part of a literature review. In other words, citing several theoretical works does not necessarily indicate that theory comprises a substantive element of a reported study.

In future work we plan to trace the academic ancestral origins or influences of several theories for the purpose of identifying which disciplines IS researchers are using as a basis for building their own theory. As part of this work we also will examine the original IS work and survey the authors to determine whether they themselves considered or proposed their work as “theory.” Other research might examine the fit between the theories we identified and the three substratas or “Big Questions” of information science recently described by Bates (1999, p. 1048). Can the theories ascribed by the authors in our study, in other words, be grouped according to whether they address or complement (1) the physical question: What are the features of the universe? (2) The social question: How do people relate to, seek and use information? or (3) The design question: How can access to recorded information be made most rapid and effective? Further thought also should be given to Chatman’s (1996, p. 193) statement that “we are currently focused on the application of conceptual frameworks rather than on the generation of specific theories.” Answers to such questions will facilitate IS efforts at theory building and understanding its theoretical core.

Acknowledgments

This research received the Association for Library and Information Science Education’s 1999 Research Award. We are grateful to ALISE for sponsoring this study. Dr. Pettigrew and Dr. MeKechnie equally share first authorship of this article.

APPENDIX I—CONTENT ANALYSIS CODE BOOK

NOTE: “Primary affiliation of the first author” was coded using information provided within the article itself or in another part of the journal issue such as an “About our Contributors” section. “Type of article” codes were developed to answer the question “What kind of article is this?” or “What approach is used in writing this article?” Subject codes describe main content areas of IS. When an article covered two or more subject areas, the principal subject (that receiving the most coverage) was coded. Articles that dealt with IS in general or IS research were included in the “other” category.

Affiliation of First Author
- Private sector
- Government
- IS/LIS university/practice
- Sciences university
- Social sciences university
- Humanities university
- Unknown

Primary Subject of Article
- Indexing/abstracting/cataloging and classification
- Information retrieval
- Information technology (including www, cd-rom, gis, systems)
- HCl/interface design
- Bibliometrics
- Information policy
- Library services (design and delivery of services and programs)
- Management (human resources, fiscal, planning)
- Scholarly communication and publishing
- History
- Human information behavior
- Education and pedagogy
- Other (including general IS and IS research)

Type of Article
- Report of empirical research
- Descriptive paper
- Verbal argumentation
- Mathematical modeling/algorithm development
- Discourse analysis
- Historical paper
- Literature review
- Theory paper
- Method paper
- Other

APPENDIX II—THEORIES ATTRIBUTED TO IS AUTHORS (partial listing).

Agosti’s two-level retrieval model
Albrechtsen’s domain analysis theory
Bates’s berry picking
Bates’s “hit-side-of-the barn” principle
Belkin, Oddy & Brook’s Anomalous States of Knowledge
Briet’s document
Brook’s equation for information science
Chatman’s insider-outsider effect theory
APPENDIX III—THEORIES ATTRIBUTED TO AUTHORS IN THE SCIENCES (partial listing)

Bayesian regression theory
Bradford’s law
Chaos theory
Chen’s conceptual model for storage and retrieval
Dempster–Shafer theory of evidence
Fuzzy set theory
Graph theory
Greek stemmer algorithm
Information exploration paradigm
Lotka’s law
Markov chain theory
Newtonian physics theory
Probability theory
Relational database model/theory
Rhetorical structure theory
Rough set theory
Shannon’s mathematical theory of communication
Simon–Yule distribution
Theoretical aspects of multivariate data
Wiener’s cybernetics
Zipf’s law

APPENDIX IV—THEORIES ATTRIBUTED TO AUTHORS IN THE SOCIAL SCIENCES (partial listing)

Activity theory
Alienation theory
Attitude theory
Attribution theory
Behaviorist learning theory
Berger & Luckman’s social construction
Boisot’s approach to the political economy of information
Bourdieu’s sociological theory based on his conception of “fields”
Bruner’s theory of instruction
Cameron’s organizational effectiveness models
Cannibalism theory
Clifford’s cultural relativity theory
Communication theory
Congruity theory
Constructivist approach to learning
Critical theory
Dale’s cone of experience
Decision theory
de Solla Price’s invisible colleges
Deutsch’s cycles of diffusion theory
Dewey’s theory of pragmatics
Disconfirmation theory
Ecological approach
Economics theory
Education evaluation theory
Experiential approach to learning
Faibisoff and Ely’s principle of least effort
Feminist theory
Fishbein & Ajzen’s theory of reasoned action
Freud’s theory of psychoanalysis
Functionalism
Game theory
Garfinkle’s ethnomethodology
General theory for management information systems
Gidden’s structuration theory
Ginsberg & Venkatraman’s organizational structure theory
Glaser & Strauss’s grounded theory
Habermas’s communicative action theory
Interest group behavior theory
Izard’s evolutionary-development view of emotion
Johnson–Laird’s mental models theory
Kelly’s personal construct theory
Kuhn’s theory about scientific paradigm shifts
Learning theory
Lincoln & Guba’s theory of truth
Loose coupling theory
Marx’s value–labor theory
Media richness theory
Merton’s insiders outsiders theory
Minsky’s frame theory
Multiple constituencies model of organizational effectiveness
Neo-marxism
Organization theory
Ortega hypothesis
Phase model of burnout
Piaget’s child development theory
Pragmatics
Roger’s diffusion theory
Schemata theory
Simon’s theory of intelligence, design, and choice
Social impact theory
Social network theory
Speech-act theory
Statistics anxiety
Stress and coping theory
Symbolic interactionism
Time series analysis
Uncertainty theory
Uses and gratification theory
Vygotsky’s cognition and sociocultural development theory
Weick’s sense-making theory

APPENDIX V—THEORIES ATTRIBUTED TO AUTHORS IN THE HUMANITIES (partial listing)

Chomsky’s syntactic rules
de Saussure’s semiotics
Discourse theory
Foucault’s notion of power and knowledge
Foucault’s social constructivism
Habermas’s discourse or dialogue democracy
Hermeneutics (Ricoeur)
Iser’s reader response theory
Postmodernism
Wittgenstein’s central role of language

References


