Communications of the Association for Information Systems

Volume 38 Article 26

5-2016

Theory in North American Information Systems Research: A Culturomic Analysis

Daniel S. Soper
California State University, dsoper@fullerton.edu

Ofir Turel
California State University

Communications of the Association for Information Systems

Research Paper ISSN: 1529-3181

Theory in North American Information Systems Research: A Culturomic Analysis

Daniel S. Soper
Department of Information Systems and Decision Sciences, Mihaylo College of Business and Economics, California State University, Fullerton
dsoper@fullerton.edu

Ofir Turel
Department of Information Systems and Decision Sciences, Mihaylo College of Business and Economics, California State University, Fullerton

Abstract:

Since its inception, North American information systems (IS) research has relied on a broad and varied collection of theories. The core of this theoretical landscape is an important determinant of the IS research community's identity, and, as such, researchers have discussed it extensively in recent years. Nevertheless, we know few concrete facts about the composition, consistency, or evolution of this theoretical core over time. Using a set of 318 theories in conjunction with n-gram analyses, we address these issues empirically by computationally analyzing the complete text of every research paper published in three leading North American IS journals over a 24-year period. In examining these 2,215 papers and more than 3.54 billion n-gram records, we identify the theories that constitute the overall core of North American IS research and provide insights into the evolution of that core. We further identify and quantify the nature of theoretical pluralism in North American IS research and examine the evolution of the theoretical density of IS research studies over time. Finally, our results shed light on the patterns of theory co-occurrence in North American IS research studies and demonstrate how such information can facilitate increasingly imperative efforts aimed at theory consolidation and generalization.

Keywords: IS Research, Theoretical Core, N-gram Analysis, Culturomics, Theory Generalization.

This manuscript underwent peer review. It was received 22/12/2014 and was with the authors for 8 months for 2 revisions. Ping Zhang served as Associate Editor.

1 Introduction

One of the most interesting and distinguishing traits of information systems (IS) research is its reliance on a vast patchwork of theories drawn from virtually all realms of scientific and managerial inquiry (Benbasat & Zmud, 2003; Hassan, 2011; Hirschheim & Klein, 2012). This situation is perhaps not surprising, especially when one considers the immense variety of economic, behavioral, technological, and organizational phenomena that have come to be arrayed under the expansive and ever-growing banner of IS research. Although this theoretical diversity may in certain ways be advantageous for the field (e.g., by fostering creativity or innovative modes of thought among the IS research community (Agarwal & Lucas, 2005)), it also presents a serious challenge in that it makes the field's intellectual core and boundaries (i.e., the field's identity) difficult to describe (Benbasat & Zmud, 2003). Whereas a relatively small set of macro-theoretic paradigms define many scientific fields, the fragmentation and lack of cohesion among the myriad theories appearing in IS research makes defining the field quite difficult and confounds not only efforts aimed at understanding how the usage of theory in the field has evolved over time but also attempts to make informed predictions about the directions in which the field might move in the future (Benbasat & Zmud, 2003; Neufeld, Fang, & Huff, 2007; Somers, 2010).

In this paper, we shine an empirical light on the large and hitherto ill-defined ecosystem of theories that have appeared in IS research[1]. Scientific research in the IS field is, however, a vast and global enterprise, and, to begin to make inroads into this daunting area of inquiry, we intentionally constrained our study's scope to North American IS research (i.e., research published in leading North American IS journals). Although we limit our inquiry to this realm for practical reasons, we certainly do not discount the importance of other research regions (e.g., Asian-, African-, or European-centered IS research). As we show in Section 3, even focusing solely on North America and methodically analyzing the theories used therein requires managing billions of records. This high level of complexity is attributable to our reliance on advanced computational techniques, which we use to objectively quantify the ways in which leading North American IS research journals have used different theories and to determine how the usage of those theories has evolved over time. More specifically, innovative analyses drove our inquiry. With them, we examine and offer detailed answers to the following five research questions:

RQ1: Which theories comprise the theoretical core of North American IS research?

RQ2: How has the theoretical core of North American IS research evolved over time?

RQ3: What is the nature of theoretical pluralism within North American IS research studies?

RQ4: How has the theoretical density of North American IS research studies evolved over time?

RQ5: What is the nature of theory co-occurrence in North American IS research studies?

Many contemplative IS scholars have enumerated similar questions to those above, and the answers to questions such as these inarguably carry important implications for research into the IS community's history and evolving identity. Until recently, however, the computational tools and techniques required to objectively address such weighty questions did not exist, and, except for speculation and opinion, the IS literature have largely not answered these questions thus far. Fortunately, advancements in computational tools and techniques now allow us to bridge this divide and study these issues in an impartial, scientific manner.

This paper proceeds as follows. In Section 2, we first expand on the foundations of our research by exploring the philosophical and conceptual links that interconnect theory and identity. In Section 3, we describe our study's methodology and, in Section 4, present and discuss our quantitative results in the context of the study's five primary research questions. Finally, in Section 5, we conclude the paper.

2 Connections between Theory and Identity

Methodically analyzing and understanding the theories and theoretical patterns used in IS research is both important and necessary inasmuch as those theories and patterns characterize the field's identity and make it relevant in the broader scientific enterprise. We begin this endeavor by focusing on the theoretical nature of North American IS research outlets. Although the debates regarding the boundaries of IS

[1] In this paper, "theoretical ecosystem" means the collection of theories used in IS research and the relationships among those theories.

scholarship are still ongoing (and may never cease due to the field's evolving nature), research in this realm largely agrees that a key component of the IS research community's identity is its intellectual core, which particularly includes the primary theories that IS research studies use and advance (Hassan, 2011; Hirschheim & Klein, 2012; Neufeld, et al., 2007; Somers, 2010). Not only do the theories that a scientific field uses signal its boundaries (Benbasat & Zmud, 2003), but they also help to define the field's levels of cohesiveness and rigidity (Hassan, 2011), facilitate its impact on society (Walsham, 2012), determine the field's relevancy to practice (Ginzberg, 2012), and delimit the field's knowledge and academic domains (Somers, 2010). As such, many scholars have recently advocated establishing a more coherent identity for the IS community, which should, in part, include well-defined links to contributing fields (i.e., borrowed theories, phenomena, and methodologies) and which should effectively adapt over time to both contextual (Robey, 2003) and technological (Ginzberg, 2012) change.

Nearly every scientist has, by means of experience or study, formulated their own personal conceptualization of what theory means. Indeed, even a cursory review of the topic reveals a wide variety of definitions for theory that range from highly specific definitions rooted in mathematics and the rules of logic to broader, much more inclusive definitions rooted in philosophy and epistemology. To move forward, any study focused on the nature of theory in a scientific community must adopt a position regarding this contentious definitional issue. Therefore, we ask readers for a small degree of forbearance if the position adopted here is not perfectly aligned with their own perspectives. So, as to not be excessively exclusionary, we adopt Bacharach's (1989, p. 496) definition of theory:

> *A theory is a statement of relations among concepts within a set of boundary assumptions and constraints. It is no more than a linguistic device used to organize a complex empirical world… Therefore, the purpose of theoretical statements is twofold: to organize (parsimoniously) and to communicate (clearly).*

This definition is, we believe, sufficiently broad to encapsulate most scientific perspectives regarding what theory means, and it accommodates the notion that every theory lies along a spectrum that ranges from the macro to the micro and the notion that every theory's location on this spectrum will naturally change as it and its adjacent theories evolve over time.

Given theory's vital role in defining a scientific community's identity, one needs understand not only which theories and groups of theories the scientific community has used but also the extent to which it has used them and how their use patterns have evolved over time. Moreover, since researchers have alternately argued that consistency in the theoretical core and heterogeneity in the theoretical core are both a necessity and an impediment to IS research (Hassan, 2011; Hirschheim & Klein, 2012; Neufeld, et al., 2007; Somers, 2010), one needs to gain objective insights regarding the composition of the theoretical core of IS research in general and in narrower, more granular windows of time. Not only can addressing these issues provide insights that can contribute greatly to the study of the IS community's identity and history, but the results can also serve as objective and defensible scientific input for arguments regarding the community's intellectual core.

Certain studies have acknowledged these needs and have endeavored to use conceptual or empirical methods to identify various elements of IS research's theoretical core. Several studies, for example, have relied on experiential evidence (Agarwal & Lucas, 2005; Benbasat & Zmud, 2003), while others have employed citation analyses (Moody, Iacob, & Amrit, 2010), manual literature reviews (Hirschheim & Klein, 2012), or manual analyses of titles and abstracts (Neufeld, et al., 2007) to characterize various aspects of theory usage in IS research. Research has applied computational text-mining techniques but only in a limited fashion; for example, to analyze paper abstracts to identify intellectual communities in the IS field (Larsen, Monarchi, Hovorka, & Bailey, 2008). While such efforts have served as important preliminary steps in mapping the usage and evolution of theories in IS research over time, we believe that an approach that considers the full and complete text of thousands of North American IS research papers and that has proven successful in other social-scientific endeavors (Michel et al., 2011) has the potential to generate a much more comprehensive and accurate picture of the theoretical core of North American IS research. To that end, the findings of studies such as that described in this paper can serve as an objective, quantitative basis on which one can explore the identity of the IS research community.

3 Method

As our general strategy for inquiring into the extent to which the North American IS community has used different theories, we first identified a large set of theories used in North American IS research and

assessed the frequency with which those theories have appeared in the North American IS literature over time. The logic of this approach follows from the social science culturomic perspective[2] and builds on the theoretical notion that the importance of concepts (theories in the current case) is reflected in how frequently and consistently those concepts are mentioned over time in a contextually relevant corpus of text (Bohannon, 2011). Here, we measure theory frequency and stability by applying a well-established computational text-mining technique known as n-gram analysis.

As used in natural language processing, an n-gram is a sequence of words of length n that is extracted from a larger sequence of words (Manning & Schütze, 1999). As such, a theory mentioned in a large body of text constitutes an n-gram whose frequency of appearance can be computationally determined via analyzing the complete text. Consider, for example, the phrase "we used TAM". We can subdivide this phrase into three 1-grams ("we", "used", and "TAM"), two 2-grams ("we used" and "used TAM"), and one 3-gram ("we used TAM"). The central theoretical tenant underlying culturomic n-gram analysis is that the frequency with which a corpus of text mentions a concept mirrors the relative importance of that concept in its parent domain at the time when the text was written (Michel et al., 2011). For example, if the n-gram "smartphone" appears 500 percent more frequently than the n-gram "fax machine" in a community's corpus of text during a particular year, one can reasonably infer that smartphones garnered more interest or was more important or influential in the community during that year than fax machines.

Extending this notion, if one were to identify all of the unique concepts appearing in a corpus of text during a particular year and count the frequency with which each of those concepts appeared in the corpus, one could gain a great deal of insight into what the community that generated the corpus found to be most interesting or culturally absorbing during that particular year. The true power of n-gram analysis, however, becomes evident only when one considers trends over time. By standardizing each n-gram frequency according to the total quantity of text published during a given year, one can identify trends and patterns that reveal powerful insights into the underlying domain of interest. Consider, for example, Figure 1 below, which we generated by means of an n-gram analysis of the text contained in millions of books written between 1800 and 2000 (Google, 2015). The figure clearly shows (and quantifies) the declining centrality of religion in English-speaking culture during the 19th and 20th centuries and reveals an accompanying rise in the culture's interest in science. The ability of n-gram analyses to yield insights such as these have led researchers in many different fields to adopt the technique as the basis of culturomic and scientometric inquiry (Bohannon, 2011; Michel et al., 2011), including the IS field (Soper & Turel, 2012).

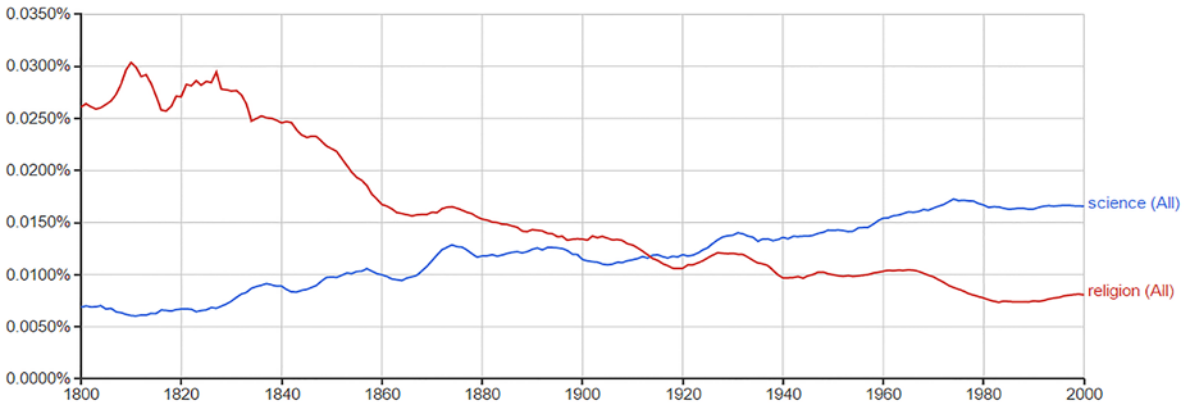

Figure 1. An N-gram Analysis of Science and Religion in the 19th and 20th Centuries

Given the focus of the current investigation on theory use in the IS field, we first needed to identify a large set of theory n-grams that could serve as input for our analyses. For this purpose, we adopted Soper and Turel's (2015) computationally identified set of 318 unique theories that have verifiably appeared in three leading North American IS research journals over time. The field believes this source to be the most complete list of theories that North American IS researchers have used, and it substantially surpasses the 88 theories that the Association for Information Systems (AIS)-affiliated "theories used in IS research" website listed at the time of writing (Larsen, 2014). As with most scientific fields, IS researchers frequently

[2] Culturomics refers to the study of cultural or behavioral trends by computationally analyzing large collections of digitized text (Bohannon, 2011).

use acronyms and multiple names to refer to the same underlying theory (e.g., theory of reasoned action and TRA). If not handled properly, this linguistic phenomenon—known as the coreference problem (Crystal, 1997)—can cause issues with respect to the validity and reliability of results obtained from n-gram analyses. To ensure that our analytic results would be as accurate as possible, we incorporated the collection of alternate names and acronyms that researchers have used for each theory. Soper and Turel (2015) also include this information, and, as such, we also adopted their list of alternate theory names and acronyms for our study. We each independently reviewed the final list for accuracy before proceeding with the analysis.

Having identified a set of theories, we next constructed a corpus containing the complete text of every research paper published in *Information Systems Research* (ISR), the *Journal of Management Information Systems* (JMIS), and *Management Information Systems Quarterly* (MISQ) during the 24-year period between 1990 and 2013 (inclusive). Note that we constrained the content of the corpus to research papers; we intentionally excluded other types of papers such as editorial introductions, book reviews, and so on from the corpus. We chose *ISR*, *JMIS*, and *MISQ* because the IS field widely regards them as the leading IS research journals published in North America (Ferratt, Gorman, Kanet, & Salisbury, 2007; Peffers & Ya, 2003; Rainer & Miller, 2005). Since these three journals are all North American scholarly outlets focused largely on IS-related phenomena, one should note that they are an imperfect proxy for global IS research as a whole. Therefore, readers should view the results we report here as particularly illuminating to North American IS research rather than to the IS field as a whole. Again, we constrained our focus to North American IS research for feasibility purposes, and one should not misconstrue this focus as our discounting the importance or contributions of other perspectives. We chose the timespan for the papers we included in the corpus (i.e., 1990 through 2013) because 1990 was the first year in which all three journals we selected concurrently existed and 2013 was the last year for which a complete set of papers was available at the time when we constructed the corpus. In total, the corpus spanned 24 years and contained 2,215 research papers that, together, comprised nearly 30 million words. To put the size of the corpus in perspective, consider that if one were to spend 40 hours per week reading at the average adult rate of 250 words per minute, one would need nearly a full year to read every research papers published by these three journals between 1990 and 2013.

With our corpus complete, we next constructed a custom software system to tokenize the text of each paper into a series of n-grams. The number of n-grams that one can theoretically extract from a large corpus of text greatly exceeds the number of words in the corpus itself, which presents serious scaling and performance implications for a corpus containing millions of words. Although past research has imposed an upper limit of no more than five words for the length of a single n-gram (Michel, et al., 2011; Soper & Turel, 2012), we constrained our analysis to include n-grams with a maximum length of eight words so we could capture of even very lengthy theory names (e.g., Venkatesh, Morris, Davis, & Davis's (2003) unified theory of acceptance and use of technology). Except for acronyms, we converted all words in the corpus to lowercase prior to tokenization to eliminate any problems that might otherwise arise due to capitalization. Using this strategy, the system would view "Theory of Reasoned Action" as equivalent to "theory of reasoned action" and an acronym such as "IS" as distinct from the word "is", which ensured the accuracy of the results. Following these preliminary tasks, we transformed the complete text of each paper into a large database of n-grams, with the source year, paper, and journal metadata being retained for each n-gram.

As we note previously, we predicated our analysis on the theory-driven, social-scientific notion that the relative frequency with which a given theory appears in the IS research community's leading journals over time reflects the degree of influence of that theory in IS research. As such, we needed to compute the frequencies with which each n-gram in the corpus appeared in *ISR*, *JMIS*, and *MISQ* during each year of the analysis. Directly comparing raw frequency counts for the same n-gram across journals or across time would be misleading, however, since doing so would ignore differences in the number of words published by each journal from year to year. Therefore, we had to calculate the relative frequencies for each n-gram by dividing their respective raw frequency counts in each journal for a given year by the total number of words published by those journals during the year in question. This approach is consistent with past n-gram research and yielded a standardized measure of frequency that would allow valid comparisons to be made between n-grams across journals and across time (Michel et al., 2011; Soper & Turel, 2012). Thus, the standardized frequency values resulting from this process indicated how often a particular n-gram appeared in a particular journal during a specific year relative to the total quantity of text published in that journal during that year.

As a result of these data-extraction and processing tasks, we produced a large database containing more than 49 million unique n-grams. Since we computed a standardized frequency measure for each of these n-grams for each of the three journals and for each of the 24 years of the analysis, the final dataset contained more than 3.5 billion n-gram frequency records. With the final database complete, we could use database queries to identify the relative frequency with which any n-gram of length ($1 \leq n \leq 8$) appeared during a particular year. Further, we also computed the average of the relative frequencies across all three journals for each combination of n-gram and year so to have a metric that could serve as a mean-stabilized proxy for the relative frequency with which a given n-gram appeared in the corpus of North American IS literature during a particular year.

Next, we constructed a custom Web-based information system that enabled us to query, visualize, and explore our n-grams database. With this system, we could not only simultaneously plot and analyze multiple n-grams but also combine related n-grams into a single result, which resolved the coreference problem we describe previously. For example, the search phrase "theory of reasoned action + TRA" would produce output representing the combined frequencies of the n-grams "theory of reasoned action" and "TRA" over time. We used this approach to combine the relative frequencies of all of the labels associated with each theory (i.e., each theory's name, alternative names, and acronyms) into a single result that accurately represented the overall relative frequency with which a particular theory appeared in the three leading North American IS journals during a specific year.

Finally, prior to presenting our results, we must also consider the difference between a paper's *mentioning* a theory and *using* a theory. Clearly, papers sometimes mention theories tangentially or for purposes of completeness in a manuscript without using it in a substantive or significant way. Indeed, authors regularly mention theories in literature reviews or at the suggestion of editors or anonymous referees without relying on those theories as the basis for their own research. Given that we focus on the theoretical core of the IS field, we needed to identify and exclude theories that papers merely mentioned (rather than used) to avoid artificially inflating the frequency counts for each theory. For this reason, we conducted a paper-level theory frequency analysis to quantify not just *whether* but also *how frequently* each theory appeared in each research paper. From the perspective of linguistic theory, the conceptual notion underlying this analysis is that the names of theories that papers actually use will, on average, appear much more frequently in that paper than the names of theories that the paper mentions in passing or for the sake of completeness. Figure 2 shows the results of this paper-theory frequency analysis.

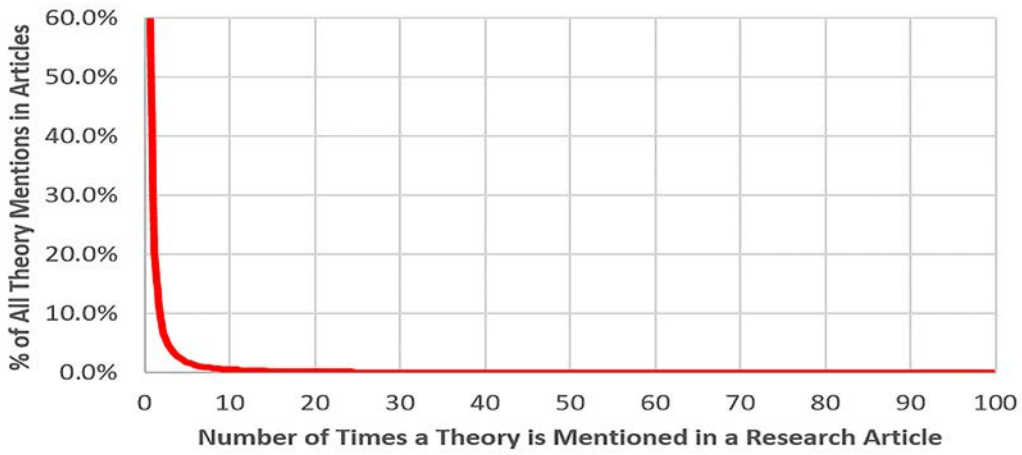

Figure 2. Distribution of Theory Appearances in IS Research Papers

As Figure 2 shows, in the vast majority of situations where a theory appears in an IS research paper, the paper mentioned the theory only a few times. More specifically, when an IS research paper mentioned a theory, 59.6 percent of the time the paper mentioned it just once and 75.7 percent of the time once or twice. If we are to determine the composition of the IS field's theoretical core based on the theories that IS scholars actually use in their research (as opposed to theories that they merely mention), then these results clearly suggest that we need a method by which to categorize the theories appearing in each IS research paper as core or peripheral according to the frequency with which each paper mentions each theory.

In light of the considerations above, we adopted a statistical approach to identifying peripheral theories that relied on binomial probability distributions. The null hypothesis in this approach was that a paper mentioned all theories in it with equal frequency, which signaled that the paper viewed the theories as equally important. Thus, by identifying the number of unique theories appearing in a paper and the total number of times the paper mentioned theories, the expected frequency that any of the theories would appear under the null hypothesis would be equal to the ratio of the collective sum of theory mentions to the total number of unique theories appearing in the paper. If the observed appearance frequency for any theory was statistically less than its expected frequency, then one would label that theory as peripheral to the research paper and would, for the research paper in question's purposes, omit it from considerations relating to the IS field's core. As a practical example, imagine a scenario in which three unique theories appeared in a research paper and that the paper mentioned these theories 24 times in total. Using the method described above, one would identify a theory as unimportant to a paper if it appeared in the paper three or fewer times since its observed frequency of appearance would be statistically less than expected at the $p < 0.05$ level. For the special case in which a paper mentioned only a single theory, one would consider the theory core to the paper only if its observed frequency fell in the fourth quartile of all paper-theory mentions in the dataset. Guided by this method, we could focus our inquiry exclusively on those theories that IS research papers actually used in a significant way. Except where otherwise indicated, readers should interpret the results in Section 4 as applying only to such substantive instances of theory usage.

4 Results

Before detailing the results, which speak directly to the study's five primary research questions, we briefly examine some of the structural changes that have occurred in leading North American IS journals during the past 24 years. As Table 1 shows, the overall total number of research papers published per year by *ISR*, *JMIS*, and *MISQ* increased dramatically between 1990 and 2013—an increase that partially explains an even larger proportional increase in the total number of words published per year. In consequence of these concomitant increases, the overall average length of each research papers published in leading North American IS journals has nearly doubled from 8,286 words per paper in 1990 to 15,383 words per papers in 2013 (excluding references).

Table 1. Publication Statistics for Research Papers Appearing in North American IS Journals

Publication statistic	Publication year						
	1990	1994	1998	2002	2006	2010	2013
Total papers	70	72	74	90	103	130	156
Total words	580,052	711,914	883,106	944,900	1,252,780	1,618,105	2,399,755
Words per paper	8,286	9,888	11,934	10,499	12,163	12,447	15,383

Collectively, we believe the trends in Table 1 indicate 1) that editors of leading North American IS journals have heeded Dennis, Valacich, Fuller, and Schneider's (2006), among others', calls to substantially increase the volume of high-quality papers published in elite IS journals; and 2) that North American IS researchers may be investigating increasingly complex phenomena that require both multi-theoretic foundations and a lengthier exposition. Projecting this trend a decade into the future reveals that, ceteris paribus, an average research papers published in one of these journals in the year 2026 will exceed 17,300 words in length (or approximately 70 double-spaced pages) (excluding references). It seems clear that this trajectory will not be sustainable in the long term.

4.1 Research Question 1: The Theoretical Core of North American IS Research

Our first research question concerns identifying the overall theoretical core of North American IS research. As a broad overview, we adopt the notion that the strongest and most judicious candidates for membership in the theoretical core are those theories that have exerted both a significant and a consistent impact on North American IS research over time. Thus, in identifying the theoretical core, we focused on isolating theories that meet these two criteria. As such, we begin by presenting our primary findings with Table 2 below, which lists the top 10 percent of all of the theories considered in the analysis ranked according to their usage in North American IS research as measured by the average overall relative frequency with which each theory appeared in leading North American IS journals between 1990 and 2013. To aid in interpretation, the table presents the frequency and standard deviation values as

percentages of all of the text that appeared in the research papers published in the three leading North American IS journals during the timeframe in question. Note that these values reflect only substantive theory usage in IS research papers as we describe in Section 3—these results exclude any theories that papers merely mentioned (as to opposed to used). Additionally, the table lists only specific theories; we removed high-level families of theories (e.g., "organizational theory", "economic theory", and "social theory") from the results. The rightmost column in the table shows each theory's standardized moment (or signal-to-noise ratio), which measures the extent to which each theory's usage in North American IS journals was stable over time; higher values indicate greater levels of stability (Box, 1988). We discuss the stability of the various theories in this section after discussing theory frequencies.

Table 2. Most Commonly Used Theories in Leading North American IS Research Journals

Rank	Theory name	Overall average frequency	Overall standard deviation	Standardized moment (signal-to-noise ratio)
1	Technology acceptance model	0.010302%	0.007593%	1.357
2	Theory of planned behavior	0.004296%	0.005020%	0.856
3	Structuration theory	0.003752%	0.003944%	0.951
4	Theory of reasoned action	0.001834%	0.001828%	1.003
5	Agency theory	0.001312%	0.001444%	0.909
6	Media richness theory	0.001006%	0.002136%	0.471
7	Social cognitive theory	0.000796%	0.001383%	0.576
8	Exchange theory	0.000774%	0.001034%	0.749
9	Expectancy theory	0.000769%	0.002177%	0.353
10	IS success model	0.000720%	0.001261%	0.571
11	Contingency theory	0.000717%	0.000661%	1.085
12	Dissonance theory	0.000667%	0.001880%	0.355
13	Game theory	0.000648%	0.000607%	1.067
14	Role theory	0.000601%	0.002202%	0.273
15	Search theory	0.000586%	0.000438%	1.337
16	Activity theory	0.000503%	0.001595%	0.316
17	Theory of the firm	0.000496%	0.000490%	1.012
18	Institutional theory	0.000492%	0.000755%	0.652
19	Prospect theory	0.000479%	0.000757%	0.633
20	Network theory	0.000407%	0.000572%	0.711
21	Deterrence theory	0.000367%	0.000621%	0.591
22	Detection theory	0.000354%	0.001289%	0.274
23	Control theory	0.000336%	0.000496%	0.678
24	Bass diffusion model	0.000326%	0.000785%	0.415
25	Information processing theory	0.000313%	0.000552%	0.567
26	Actor-network theory	0.000307%	0.000927%	0.331
27	Facet theory	0.000261%	0.001279%	0.204
28	Resource dependence theory	0.000260%	0.000641%	0.406
29	Contract theory	0.000260%	0.000660%	0.394
30	Diffusion of innovations theory	0.000246%	0.000544%	0.453
31	Capability maturity model	0.000244%	0.000308%	0.791
32	Action theory	0.000243%	0.000286%	0.850

As the table shows, the technology acceptance model (TAM) (Davis, 1989) was the most frequently used theory overall in leading North American IS journals between 1990 and 2013. The theory of planned behavior (Ajzen, 1991), and structuration theory (Giddens, 1984) were the second and third most

frequently used theories, respectively. When considering the table, one can see that North American IS research has relied on a highly diverse theoretical foundation. Consistent with the perspectives that Agarwal and Lucas (2005) espouse, it seems clear that the North American IS research community has not focused exclusively on user interactions with IT artifacts but has also given primacy to the IS function and IS's organizational impact (as the commonality of the agent-centric and organization-centric theories suggest). Indeed, North American IS research does not seem to have a single unifying theoretical theme but rather appears to comprise a loose patchwork of perspectives on the uses and impacts of IS, an observation consistent with the views of many IS scholars (Agarwal & Lucas, 2005; Robey, 2003; Walsham, 2012).

Note that nearly two orders of magnitude exist between the frequencies of the highest- and lowest-ranked theories in the table, which indicates that IS scholars do not just use the lowest-ranked theories less frequently by IS scholars but exponentially less frequently than the highest-ranked theories. Figure 3 below, which depicts the average overall relative frequency of each theory according to its ordinal impact ranking in Table 2, illustrates this phenomenon. Figure 3 also includes a trend line that shows how the comparative impacts of the various theories decline according to an exponential decay function (R^2 = 0.976). From this finding, we can conclude that only a handful of theories have exerted a substantial influence on North American IS research. The degree of influence of most theories appearing in leading North American IS journals has been, by contrast, comparatively marginal.

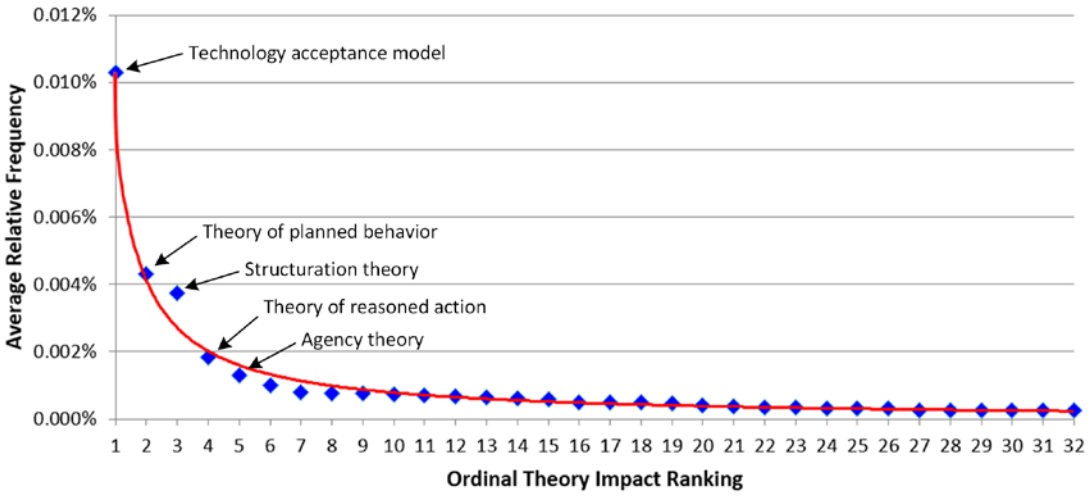

Figure 3. Comparative Overall Impact of Theories by Ordinal Ranking

Aside from the overall impact that a theory has exerted over time, considering the relative stability and consistency of that theory's impact over time is also a crucial determinant of whether one can reasonably argue that it resides in the theoretical core of North American IS research. As a means of providing a methodologically sound basis through which we could compare the relative stability of the various theories, we computed the overall standardized moment (or signal-to-noise ratio) for each theory (Box, 1988). Table 3 lists the top 10 percent of all of the theories considered in the analysis ranked according to the stability of their usage in leading North American IS journals over time as measured by each theory's standardized moment.

In contrast to theoretical "fads", theories that have appeared both frequently in leading IS journals and in a manner that is comparatively stable and consistent over time are the strongest and most judicious candidates for membership in the overall theoretical core. Tables 2 and 3 list the most common and most stable theories (respectively) that have appeared in leading North American IS journals between 1990 and 2013, and a total of 17 unique theories lie at the intersection of those two tables. Put differently, among all of the theories used in these three leading IS journals, a total of 17 theories appeared in the top 10 percent in terms of both frequency and consistency of usage over time. Figure 4 shows this set of 17 theories as a Venn diagram. To answer our first research question, the data and subsequent relative frequency and stability analyses suggest that one can reasonably argue the theories appearing in the figure to constitute the overall theoretical core of North American IS research between 1990 and 2013. To put the impact of this theoretical core into perspective, consider that approximately 43 percent of all of the

research papers published in *ISR*, *JMIS*, and *MISQ* between 1990 and 2013 relied on one or more of these core theories.

Table 3. Most Stable Theories in Leading North American IS Research Journals

Rank	Theory name	Overall average frequency	Overall standard deviation	Standardized moment (signal-to-noise ratio)
1	Technology acceptance model	0.010302%	0.007593%	1.357
2	Search theory	0.000586%	0.000438%	1.337
3	Contingency theory	0.000717%	0.000661%	1.085
4	Game theory	0.000648%	0.000607%	1.067
5	Theory of the firm	0.000496%	0.000490%	1.012
6	Theory of reasoned action	0.001834%	0.001828%	1.003
7	Theory of groups	0.000073%	0.000076%	0.959
8	Structuration theory	0.003752%	0.003944%	0.951
9	Agency theory	0.001312%	0.001444%	0.909
10	Theory of planned behavior	0.004296%	0.005020%	0.856
11	Action theory	0.000243%	0.000286%	0.850
12	Social presence theory	0.000216%	0.000267%	0.811
13	Attribution theory	0.000226%	0.000282%	0.803
14	Communication theory	0.000148%	0.000185%	0.802
15	Capability maturity model	0.000244%	0.000308%	0.791
16	Behavioral theory of the firm	0.000060%	0.000078%	0.766
17	Exchange theory	0.000774%	0.001034%	0.749
18	Network theory	0.000407%	0.000572%	0.711
19	Process theory	0.000230%	0.000332%	0.692
20	Complexity theory	0.000107%	0.000155%	0.691
21	Queuing theory	0.000171%	0.000252%	0.680
22	Interaction theory	0.000142%	0.000208%	0.679
23	Control theory	0.000336%	0.000496%	0.678
24	Institutional theory	0.000492%	0.000755%	0.652
25	Prospect theory	0.000479%	0.000757%	0.633
26	Capital asset pricing model	0.000055%	0.000087%	0.627
27	Social network theory	0.000103%	0.000165%	0.622
28	Information theory	0.000085%	0.000141%	0.600
29	Deterrence theory	0.000367%	0.000621%	0.591
30	Graph theory	0.000052%	0.000089%	0.589
31	Five forces model	0.000100%	0.000172%	0.585
32	Social learning theory	0.000112%	0.000195%	0.576

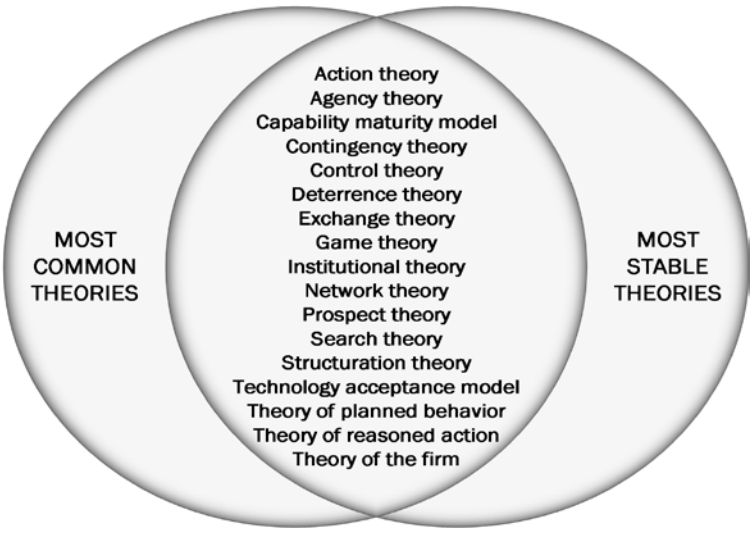

Figure 4. Overall Theoretical Core of North American IS Research from 1990 through 2013

4.2 Research Question 2: The Evolution of the North American IS Theoretical Core

Our second research question concerns the ways in which the theoretical core of North American IS research has evolved over time. One useful way of gaining insights into this question is to compare and contrast the varying impacts that the theories in the overall theoretical core exerted on North American IS research from year to year. For this purpose, we divided the relative frequency values for each year in the analysis into quintiles, which we subsequently used to represent the comparative impact of the theories in the theoretical core over time. In this way, we could readily compare the relative magnitude of each core theory's impact during a particular year to the relative impacts of the other core theories at any point in the analytic timeframe. Figure 5 provides the results of these activities. Since we divided the comparative impacts into quintiles, a cell in the figure shaded as "low" indicates that a theory's relative impact in the core was in the bottom 20 percent at that point in time, while a cell shaded as "high" indicates that a theory's a relative impact in the core was in the top 20 percent at that point in time.

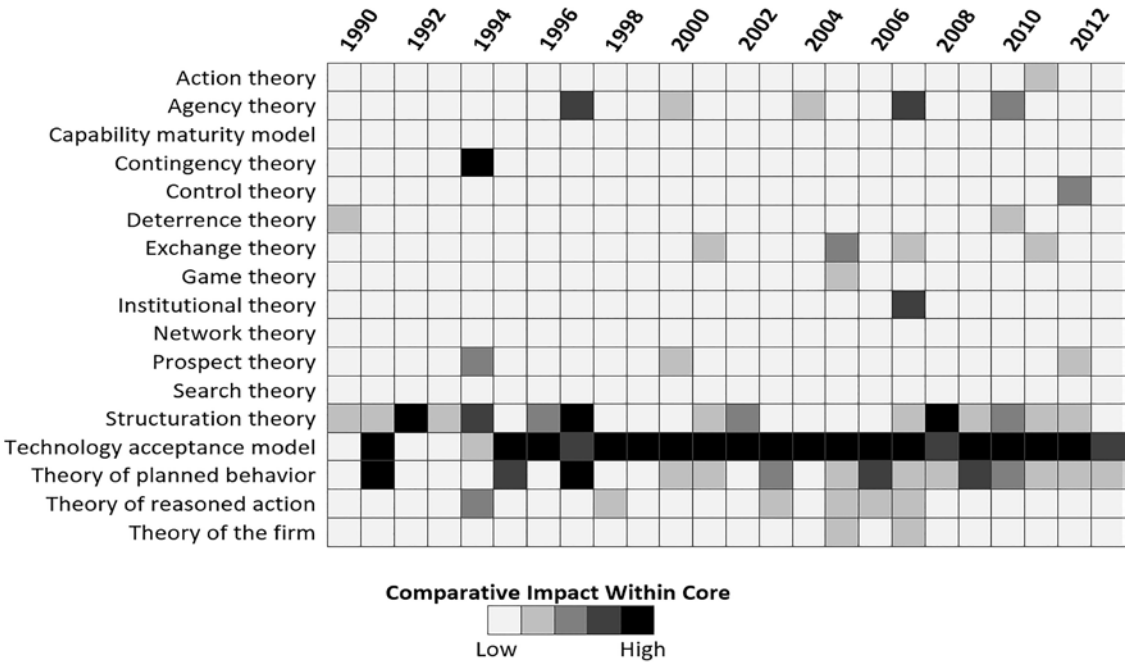

Figure 5. Comparative Impact of Core North American IS Theories from 1990 through 2013

We can glean several interesting insights regarding the evolution of the North American IS theoretical core from Figure 5. To begin, when considering the impact of the sundry core theories over time, one can see that few theories have tended to be particularly dominant during any given year. This is not to say that the other core theories are not important to North American IS research as a whole during a given year; indeed, all of the core theories are highly influential with respect to the complete set of 318 theories considered in our analyses. This observation simply reveals that, when focusing exclusively on the set of 17 core theories, a small number of those theories tend to predominate over the other members of the core during each particular year.

When considering the impact of individual core theories, we can see that each theory has exhibited a unique pattern of influence and dormancy over time. Three specific examples (agency theory, structuration theory, and the technology acceptance model) illustrate this situation. Agency theory represents the first pattern. As Figure 5 shows, agency theory's comparative impact has waxed and waned in regular three- to four-year cycles over time and exhibits a pattern similar to that of a sine wave, which one can interpret as brief periods of intense interest among IS researchers followed by lengthier periods of comparative dormancy. Several possible explanations for this phenomenon exist, including journal publication cycles, agent-centric special issues, and so on. Structuration theory represents the second pattern. Structuration theory has enjoyed an extended period of intense interest among IS researchers in the early- to mid-1990s followed by an extended period of comparative dormancy. In more recent years, however, IS researchers again exhibited an intense and extended interest in structuration theory. Like agency theory, one can also envisage this pattern as a sine wave, albeit one with a much greater wavelength. Possible explanations for the pattern might include the discovery of Anthony Giddens' work by a new generation of IS scholars, or, given the nature of structuration theory, a longer term cycle of interest among IS scholars in the connections between the individual and the group. Together, these two sinusoidal patterns may provide some support for the theoretical "fashion waves" hypothesis that Baskerville and Myers (2009) has espoused and that posits that relatively transitory bursts of interest in particular topics characterizes IS research. Finally, the pattern of influence exhibited by the technology acceptance model deserves a few comments. Notwithstanding several inconsistencies in the early 1990s, the technology acceptance model has been a powerful and highly influential force in IS research for more than a quarter of a century. Indeed, for an extended period spanning the late 1990s to mid-2000s, the technology acceptance model was the *only* highly influential theory in the North American IS theoretical core. Clearly, this theory deserves a place of special prominence in the annals of IS research history. It remains to be seen whether any future theory will ever have such a prolonged and pronounced impact on the field.

As a means of providing further insights into the evolution of the theoretical core of North American IS research over time, we applied the same methodological strategy used to address our first research question to four six-year "windows" evenly divided among the 24 years' of IS research papers contained in our corpus. Table 4 presents the results of these analysis. Since inclusion in the core required theories to be both influential and stable during the timeframe in question, the number of theories comprising the core naturally varied from one window of time to the next. Reflecting on the composition of the theoretical core during these more temporally granular windows of time reveals several additional insights regarding the evolution of North American IS research between 1990 and 2013. Note also that the rightmost column in the table (i.e., the column entitled core theories 2008 to 2013) contains a list of the theories that might reasonably be considered to comprise the current theoretical core of North American IS research.

Of the 23 unique theories appearing in Table 4, only three (13.0%) appear consistently in all four timeframes (i.e., agency theory, the technology acceptance model, and the theory of reasoned action). By contrast, five of the theories (21.7%) appear in three of the four timeframes (i.e., contingency theory, game theory, search theory, the theory of planned behavior, and the theory of the firm). Interestingly, all eight of the abovementioned theories are also members of the overall theoretical core (see Figure 4). Also, Table 4 contains only two native information systems theories: the technology acceptance model and the IS success model. Whereas TAM almost immediately became a part of the theoretical core of North American IS research following its introduction in 1989, the IS success model did not achieve maximal prominence in the field until some ten years after its introduction. The remaining 15 theories listed in Table 4 appear in the core during only one or two timeframes.

Table 4. The Evolution of the Theoretical Core of North American IS Research Over Time

Core theories 1990 to 1995	Core theories 1996 to 2001	Core theories 2002 to 2007	Core theories 2008 to 2013
• Action theory • Agency theory • Blackboard model • Contingency theory • Game theory • Interaction theory • Queuing theory • Search theory • Structuration theory • Technology acceptance model • Theory of planned behavior • Theory of reasoned action	• Agency theory • Contingency theory • Deterrence theory • Dissonance theory • Structuration theory • Technology acceptance model • Theory of planned behavior • Theory of reasoned action • Theory of the firm	• Agency theory • Capability maturity model • Contingency theory • Control theory • Exchange theory • Game theory • Institutional theory • IS success model • Network theory • Search theory • Social cognitive theory • Technology acceptance model • Theory of reasoned action • Theory of the firm	• Agency theory • Control theory • Exchange theory • Game theory • Institutional theory • Network theory • Prospect theory • Search theory • Social cognitive theory • Technology acceptance model • Theory of planned behavior • Theory of reasoned action • Theory of the firm

Collectively, the observations we note above reveal that, when viewed through narrower windows of time, the North American IS theoretical core has evolved substantially over the past 24 years. One useful way of quantifying the rate of this evolution is to examine the extent to which the composition of the theoretical core changed from one timeframe to the next. For this purpose, we computed the percentage of theories appearing in the core during one timeframe that did not appear in the core during the previous timeframe. The results indicated that 33.3 percent of the core theories during the 1996 to 2001 timeframe were not present during the 1990 to 1995 timeframe. By contrast, 64.3 percent of the core theories in the 2002 to 2007 timeframe were not present during the previous 1996 to 2001 timeframe, while only 15.4 percent of the theories appearing in the 2008 to 2013 timeframe were not present in the 2002 to 2007 timeframe. Together, these results imply that the rate of evolution of the North American IS theoretical core has been widely variable over time, with a period of relative theoretical stability in the 1990s being supplanted by a period of rapid theoretical turnover and evolution of the core during the early to mid-2000s. In more recent years, however, the field has once again returned to a period of comparative theoretical stability. It is fascinating to consider that this observed pattern of evolution in the IS theoretical core (i.e., long periods of slow change followed by short bursts of rapid change) is fully consistent with Niles Eldredge and Stephen Jay Gould's (1972) landmark evolutionary theory of punctuated equilibrium, which postulates that long periods of slow change that are occasionally interrupted by brief periods of rapid change characterize biological evolution.

4.3 Research Question 3: The Nature of Theoretical Pluralism in North American IS Research Studies

Although knowledge of how specific theories have shaped North American IS research over time is valuable, one can derive deeper insights into the nature and evolution of IS research by examining theory frequencies at the level of individual research papers. As such, our third research question concerns the nature of theoretical pluralism in individual IS research studies. To address this question, we computationally examined the theory frequency data to determine which (if any) of the 318 theories under scrutiny were used (as opposed to merely mentioned) in each of the 2,215 IS research papers in our corpus. In this way, we could quantify the number of unique theories that each research paper used during the timeframe of our inquiry—a metric we refer to as a paper's "theoretical density". Table 5 presents the results of these activities.

Table 5. Number of Unique Theories Used per Paper in Leading North American IS Journals from 1990 through 2013

		Theories per paper				
Source	**n**	**Median**	**Mean**	**Std dev**	**Min**	**Max**
ISR	645	1	1.355	1.880	0	14
JMIS	865	1	1.636	2.174	0	11

Table 5. Number of Unique Theories Used per Paper in Leading North American IS Journals from 1990 through 2013

MISQ	705	2	2.586	2.982	0	20
All journals	2,215	1	1.856	2.442	0	20

As Table 5 shows, North American IS research papers published between 1990 and 2013 relied on an average of approximately 1.86 unique theories per paper. The mean and standard deviation values reported in the table, however, reveal a highly positive skew in the distribution of theories per paper. Therefore, prior to conducting any statistical tests, we applied a natural log transformation to the theory frequency data to ensure that the normality assumption would not be violated. One-way analyses of variance applied to the transformed theory frequencies revealed that, on average, papers published in *MISQ* between 1990 and 2013 exhibited a greater degree of theoretical pluralism than papers published in *JMIS* ($F_{1,1568} = 54.076$, $p < 0.001$) or *ISR* ($F_{1,1348} = 79.671$, $p < 0.001$) during the same timeframe, while papers published in *JMIS* exhibited a greater degree of theoretical pluralism than papers published in ISR ($F_{1,1508} = 4.266$, $p < 0.05$). Comparatively speaking, these observations imply that, on average, research papers published in *MISQ* between 1990 and 2013 relied on the most elaborate theoretical frameworks, while the research papers published in *ISR* and *JMIS* tended to be more theoretically parsimonious. Examining each journal's median number of unique theories per paper as Table 5 shows also supports this conclusion. One may attribute the differences across journals to divergent editorial preferences for either theoretically complex or theoretically parsimonious research, differences in the nature of the phenomena the various journals address (e.g., individual vs. organizational phenomena, or cross-domain vs. intra-domain phenomena), or both. Interestingly, observationally speaking, at the time of writing, a strong positive correlation ($r = 0.948$) existed between the mean number of theories used in each journal's research papers and the journal's most recent Thomson Reuters impact factor[3].

4.4 Research Question 4: The Evolving Theoretical Density of North American IS Research Studies

While insight into the overall theoretical nature of North American IS research studies is valuable of its own accord, one can gain a deeper understanding of IS research by considering the way in which the theoretical density of such studies has changed over time. Accordingly, our fourth research question concerns the evolving theoretical density of North American IS research studies between 1990 and 2013. To begin addressing this question, we first computed the average overall number of theories (i.e., the average theoretical density) used in each IS research study over time. Figure 6 visually summarizes the results.

Figure 6 reveals a strong upward trend in the theoretical density of North American IS research studies over time. Notably, the average number of theories used in each North American IS research study has more than quadrupled from 0.57 theories per paper in 1990 to 2.40 theories per paper in 2013. One possible explanation for this remarkable growth is that, when North American IS research was in a more nascent stage of its development, North American IS researchers studied phenomena in a comparatively isolated or insulated manner. As North American IS research has evolved and matured, the knowledge it has produced and the expansion of its scope of inquiry have led researchers to consider increasingly complex and nuanced phenomena, which necessitated adopting a larger and more diverse theoretical ecosystem. In the absence of concerted efforts aimed at consolidating and generalizing theory, Figure 6 would depict the natural result. Although the growth of the theoretical density of IS research papers appears to be slowing, if we project a linear trendline a decade into the future, the data suggest that, by the year 2026, each research papers appearing in leading North American IS journals will rely on a theoretical framework that integrates an average of four distinct theories. As with the rapid growth in the average number of words per IS research paper, this concomitant rapid growth in the average number of theories used per papers is clearly not sustainable in the long run. If not reconciled via consolidating and generalizing theories, such complex theoretical frameworks may ultimately precipitate a scientific crisis in IS research (Kuhn, 1962).

[3] Thomson Reuters 2014 impact factors: *MISQ* = 5.311, *ISR* = 2.436, *JMIS* = 2.062.

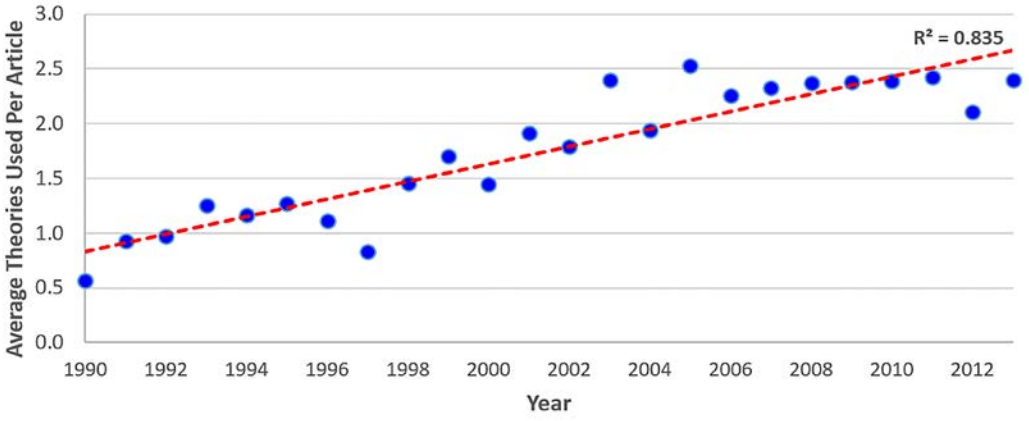

Figure 6. Average Number of Theories Used in Each North American IS Research Paper over Time

4.5 Research Question 5: The Nature of Theory Co-occurrence in North American IS Research

In light of the rapidly rising number of unique theories being used in North American IS research studies, gaining insights into the theories that are commonly used together by IS researchers is particularly relevant. As such, directly addressing this issue, our fifth and final research question concerns the nature of theory co-occurrence in North American IS research studies. To address this question, we conducted an affinity (or "market-basket") analysis (Rushing, Ranganath, Hinke, & Graves, 2001) of our paper-level theory frequency data to identify the most common theory dyads and triads that North American IS research studies used between 1990 and 2013. Given the 318 unique theories under consideration, a standard combinatorial analysis indicated that there were a total of 50,403 possible theory dyads (i.e., co-occurring pairs of theories) and 5,309,116 possible theory triads (i.e., co-occurring groups of three theories) that the corpus's papers could have used. Therefore, we examined the complete text of all 2,215 papers to ascertain which (if any) of the 5,359,519 possible theory dyads and triads each paper used. Tables and 6 present the results of these analyses sorted according to their levels of statistical support (i.e., the percentage of all North American IS research studies in which the theory dyads or triads appeared). Again, we emphasize that these calculations refer only to theories that papers actually used in (as opposed to mentioned).

As one might reasonably expect in light of the previous discussion, Tables 6 and 7 once again reveal the overwhelmingly large influence that the technology acceptance model (TAM) has exerted on IS research since the early 1990s. This single theory has, in point of fact, exerted so much influence that the tables almost entirely lack inter-theory relationships not involving TAM. For this reason, we also include Tables 8 and 9 below, which respectively list the most common theory dyads and triads used in IS research that do not involve TAM.

Returning to Tables 6 and 7, one can see that TAM, the theory of planned behavior (TPB), the theory of reasoned action (TRA), social cognitive theory, search theory, structuration theory, and the unified theory of acceptance and use of technology have commonly appeared in some combination in North American IS research papers. This finding suggests that each of these theories is but a facet of a larger, more general theory that may subsume and extend the predictive and explanatory power of its sub-theories. Indeed, one of the rare efforts to consolidate and generalize theory in IS research—namely, Venkatesh et al.'s (2003) unified theory of acceptance and use of technology (UTAUT)—seeks to achieve this very objective, but even UTAUT does not subsume all of the various constructs that characterize the theories listed above. Further, despite its introduction more than a decade ago, the IS community has adopted UTAUT rather slowly. For example, of the 1,162 studies in our corpus that appeared after the introduction of UTAUT (i.e., studies published between 2004 and 2013), only 26 studies (2.2%) mentioned UTAUT, while 232 other studies (19.9%) used TAM, TPB, and/or TRA without relying on or even mentioning UTAUT. This behavior perhaps indicates a certain degree of reluctance among IS researchers to exchange the simplicity of theories such as TAM, TBP, and TRA for more complex theories such as UTAUT despite the latter's advantages in terms of predictive and explanatory power. One can see a historical parallel to this behavior in the ongoing use of Newtonian mechanics even more than 100 years

after the introduction of Albert Einstein's more comprehensive and powerful—but more complex—general theory of relativity.

Table 6. Most Common Theory Dyads in North American IS Research Papers from 1990 through 2013

Rank	Theory dyad	Support
1	Technology acceptance model, theory of planned behavior	5.688%
2	Technology acceptance model, theory of reasoned action	4.740%
3	Theory of planned behavior, theory of reasoned action	4.018%
4	Technology acceptance model, social cognitive theory	2.212%
5	Theory of planned behavior, social cognitive theory	1.896%
6	Technology acceptance model, search theory	1.625%
7	Social cognitive theory, theory of reasoned action	1.580%
8	Technology acceptance model, structuration theory	1.535%
9	Technology acceptance model, unified theory of acceptance and use of technology	1.084%
10	Technology acceptance model, exchange theory	1.084%

Table 7. Most Common Theory Triads in North American IS Research Papers from 1990 through 2013

Rank	Theory triad	Support
1	Technology acceptance model, theory of planned behavior, theory of reasoned action	3.521%
2	Technology acceptance model, theory of planned behavior, social cognitive theory	1.625%
3	Theory of planned behavior, social cognitive theory, theory of reasoned action	1.400%
4	Technology acceptance model, social cognitive theory, theory of reasoned action	1.264%
5	Technology acceptance model, theory of planned behavior, unified theory of acceptance and use of technology	0.813%
6	Technology acceptance model, structuration theory, theory of planned behavior	0.587%
7	Technology acceptance model, theory of planned behavior, search theory	0.587%
8	Technology acceptance model, search theory, theory of reasoned action	0.542%
9	Technology acceptance model, theory of reasoned action, unified theory of acceptance and use of technology	0.497%
10	Technology acceptance model, structuration theory, theory of reasoned action	0.451%

Table 8. Most Common Theory Dyads in North American IS Research Papers (Excluding TAM)

Rank	Theory dyad	Support
1	Theory of planned behavior, theory of reasoned action	4.018%
2	Theory of planned behavior, social cognitive theory	1.896%
3	Social cognitive theory, theory of reasoned action	1.580%
4	Network theory, social network theory	1.084%
5	Structuration theory, theory of planned behavior	0.858%
6	Theory of planned behavior, unified theory of acceptance and use of technology	0.858%
7	Identity theory, social identity theory	0.858%
8	Network theory, structuration theory	0.813%
9	Systems theory, structuration theory	0.767%
10	Theory of the firm, behavioral theory of the firm	0.722%

Table 9. Most Common Theory Triads in North American IS Research Papers (Excluding TAM)

Rank	Theory triad	Support
1	Theory of planned behavior, social cognitive theory, theory of reasoned action	1.400%
2	Theory of planned behavior, search theory, theory of reasoned action	0.451%
3	Theory of planned behavior, theory of reasoned action, unified theory of acceptance and use of technology	0.406%
4	Theory of planned behavior, social cognitive theory, unified theory of acceptance and use of technology	0.406%
5	Structuration theory, theory of planned behavior, theory of reasoned action	0.406%
6	Theory of planned behavior, motivation theory, theory of reasoned action	0.361%
7	Social cognitive theory, motivation theory, theory of reasoned action	0.316%
8	Theory of planned behavior, social cognitive theory, motivation theory	0.316%
9	Theory of planned behavior, exchange theory, theory of reasoned action	0.271%
10	Structuration theory, media richness theory, channel expansion theory	0.271%

Notwithstanding the IS community's seeming reluctance to rapidly assimilate UTAUT into its mainstream theoretical ecosystem, the field clearly needs to develop and adopt higher-order theories. As the field evolves and the scope of inquiry of IS research expands, the current trajectory in which researchers attempt to address their phenomena of interest by incorporating elements from more and more theories on an ad-hoc basis is, as we note above, simply not sustainable in the long run. On the contrary, consolidating and generalizing lower-order theories into more powerful, higher-order theories is a core activity in the progress of normal science and will become increasingly critical if the IS field hopes to avoid fragmentation and theoretical crisis (Kuhn, 1962).

Forging higher-order theories is, of course, a difficult task, but IS researchers need not approach this problem unarmed and poorly equipped. Inasmuch as knowledge of both the interrelationships among lower-order theories and the ways in which IS research has applied those interrelated theories are prerequisites for higher-order theory development, quantitatively derived information such as that reported in Tables 6, 7, 8, and 9 can serve as a cornerstone and guidepost for concerted efforts aimed at consolidating and generalizing theory. As an embarkation point in this process, we conducted a multidimensional scaling analysis of the co-occurrence rates for the 17 core IS theories (Borg & Groenen, 2005), the results of which Figure 7 presents.

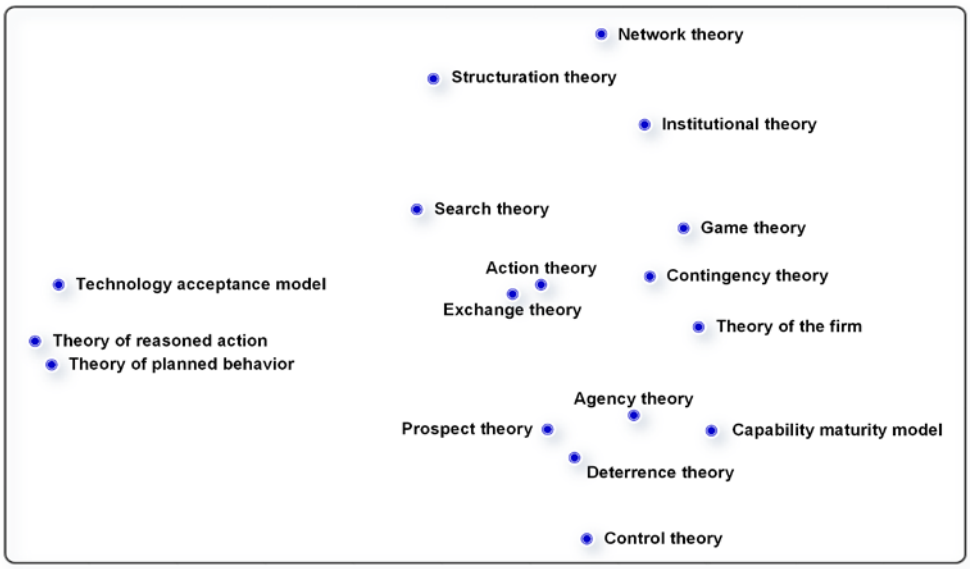

Figure 7. Relationships among Core Theories in North American IS Research

Figure 7 depicts the nature of the associations among the 17 core theories (as the North American IS community has used). The distances between each pair of theories indicates the overall likelihood of those two theories appearing together in an IS research paper. Thus, one can interpret Figure 7 as a graphical representation of the theory co-occurrence data presented in Tables 6-9 above (e.g., the most commonly co-occurring triad of theories—the technology acceptance model, the theory of planned behavior, and the theory of reasoned action—are geometrically closest to each other in the figure). Although scholars have already begun integrating the most closely related of these theories (c.f., Venkatesh et al.'s (2003) formulation of UTAUT), many additional and potentially valuable opportunities for theory consolidation remain, particularly when one considers that Figure 7 depicts the relationships among only 17 of the 318 theories we examined.

5 Summary and Discussion

In this paper, we analyzed the theoretical core of North American IS research as reflected by the theories that have been most commonly and consistently used in leading North American IS journals over a 24-year period. To gain systematic and scientific insights into this core, we relied on an established set of 318 theories and a computational text-mining technique known as n-gram analysis to evaluate the complete text of every research paper published in *ISR*, *JMIS,* and *MISQ* between 1990 and 2013 (2,215 papers in total). By following and analyzing a series of five research questions, we 1) identify the theories that comprise the theoretical core of North American IS research, 2) identify how the theoretical core has evolved over time, 3) identify the nature of theoretical pluralism in North American IS research studies, 4) identify how the theoretical density of those studies has evolved over time, and 5) identify and describe the nature of theory co-occurrence in North American IS research. The results of our analyses paint a vivid and colorful portrait of how the IS community has used theory in its efforts to examine a diverse and rapidly growing set of economic, behavioral, technological, and organizational phenomena.

In summary, we identified a small set of 17 theories that have together exerted a large and consistent influence on IS research over time. In deference to the unusual magnitude and stability of these theories' impacts on the IS community over several decades, we argue that they are the strongest and most judicious candidates for membership in the IS theoretical core. Our analyses also reveal that each theory in the core has exhibited a unique pattern of influence and dormancy over time, many of which are consistent with the fashion waves hypothesis that Baskerville and Myers (2009) advance. Further analyses revealed that the North American IS theoretical core has evolved substantially over the past 24 years. Specifically, when viewed through narrower, more temporally granular windows of time, we observed the composition of the theoretical core of North American IS research to be evolving according the tenets of Eldredge and Gould's (1972) theory of punctuated equilibrium in which brief periods of rapid change interrupt periods of relative stability.

In addition to specifics regarding the composition and evolution of the North American IS theoretical core, our findings also reveal several noteworthy insights regarding the nature of theoretical pluralism in IS research studies and the rapidly growing and fragmenting theoretical ecosystem that the North American IS community has used—topics which have of late been the focus of many researchers' intense speculation (Agarwal & Lucas, 2005; Benbasat & Zmud, 2003; Ginzberg, 2012; Grover, 2012; Hirschheim, 2007; Hirschheim & Klein, 2012; Todd, 2012; Walsham, 2012; Wastell, 2010). In the aggregate, our results point to a substantial degree of theoretical pluralism in North American IS research studies between 1990 and 2013, with researchers needing to rely on an average of approximately two unique theories per study to adequately address their phenomena of interest. Our analyses also reveal that the average number of unique theories used in each IS research study has grown by more than 400 percent since 1990. This pattern of growth has been steady and consistent over the past 24 years, and, if the pattern of linear growth continues, North American IS research studies published in the year 2026 will need to integrate elements from an average of four distinct theories to address their phenomena of interest. Such rapid growth in the complexity of the theoretical frameworks underlying IS research is clearly not sustainable in the long run and, if left unchecked, may precipitate a theoretical crisis in the North American IS community.

One tenable way of avoiding such a crisis—and indeed, the approach forwarded by Thomas Kuhn in his highly influential theory of scientific progress (Kuhn, 1962)—is to consolidate and generalize lower-order theories into more powerful, higher-order theories that encapsulate and extend the predictive and

explanatory power of the subordinate theories from which they were derived. Our analysis and results demonstrate that we can use computational techniques to guide efforts to consolidate and generalize theory by identifying theories that have been used together in IS research studies over time and quantifying the strength of those inter-theory relationships. To wit, our analyses reveal that specific theories have commonly and repeatedly co-occurred in North American IS research studies, which indicates that they share a common associative bond in the context of IS-related research phenomena and that the field can possibly integrate them into higher-order theories. Knowledge of these inter-theory relationships can undoubtedly serve as a useful guidepost for concerted efforts aimed at developing higher-order theory in IS research. As we note above, such efforts will become increasingly important—and necessary—in the coming years.

Beyond providing detailed answers to the study's five primary research questions, further major contributions of this work include the information presented in Tables 2 and 3, which respectively provide a ranked list of the most common and most stable theories appearing in North American IS research over the 24-year span of our analyses. Adding to these insights, Tables 6-9 and Figure 7 demonstrate that IS research has rarely used theories in isolation, which, we believe, not only shows the inadequacy of extant lower-order theories to independently address the phenomena of interest in contemporary IS research but also stands as a major opportunity to influence the future development and evolution of the field by means of higher-order theory development. Recalling our discussing RQ5 in Section 4.5, the seeming reluctance of the IS community to rapidly adopt the unified theory of acceptance and use of technology (Venkatesh, et al., 2003) serves as a clear warning to IS scholars seeking to formulate higher-order theories: namely, that such higher-order theories must be both powerful and parsimonious if they hope to gain widespread acceptance (Bacharach, 1989; Kuhn, 1962).

Given that much of the past commentary on the IS field's identity has been based on opinion, experience, or limited textual analyses (Agarwal & Lucas, 2005; Benbasat & Zmud, 2003; Ginzberg, 2012; Grover, 2012; Hirschheim, 2007; Hirschheim & Klein, 2012; Todd, 2012; Walsham, 2012; Wastell, 2010), we note that our findings provide objective and scientific reinforcement to some of the theory-related arguments that past research has advanced. Also note that our findings shed new light on several deep and profound theory-related questions that many IS scholars have surely pondered over the past several decades. As such, this study contributes valuable insights to our collective understanding of the past, present, and future of IS research. Finally, our using n-gram methods borrowed from computational linguistics (Suen, 1979) and culturomic ideas borrowed from social science research (Bohannon, 2011) serves as an example of the ongoing expansion of the IS field's theoretical ecosystem and its growing ties to many other fields of research. This expansion not only demonstrates the field's plasticity (King & Lyytinen, 2004; Lyytinen & King, 2006) but also points to emerging opportunities to both influence and be influenced by other fields.

5.1 Limitations and Directions for Future Research

As with all research, our work has several limitations that merit acknowledgement. First, our measurement model rests on the notion that the relative frequency with which a theory appears in a field's research literature reflects the level of interest in or influence of that theory to the field at the point in time when the research was published. Although both theoretical and applied work in the areas of computational linguistics (Bohannon, 2011) and social science (Michel et al., 2011) support this culturomic notion, there may be other scientometric methods of measuring the impact of a theory on a scientific field that future research can use.

Second, we relied on a predefined set of 318 theories as input for our analyses. Although this set was, at the time of writing, the most complete collection of theories known to have been used in IS research (Soper & Turel, 2015) and although adopting this set of theories was both expedient and methodologically justifiable, our strategy was nevertheless imperfect. To wit, the possibility remains that one or more theories appearing in our corpus of North American IS research papers did not appear in the set of theories used in the study, which makes the results potentially incomplete.

Finally, our work is limited insofar as we included papers from only three North American IS journals in the corpus. We acknowledge that the papers published in these three journals are an imperfect proxy for North American IS research and certainly for global IS research as a whole. More specifically, these journals tend to be positivist and are based in North America and may, hence, underrepresent alternative paradigms (Walsham, 2012). Therefore, readers should interpret our results accordingly. Future research

should consider expanding on our efforts to include a broader range of journals that hail from different regions of the world or that promote alternative research philosophies.

6 Concluding Remarks

The scope of inquiry of IS research in general and North American research in particular is dynamic, plastic, and permeable and has, thus, witnessed many changes in its theoretical ecosystem over the past several decades. Consequently, the theoretical core and use of theories in this realm have been elusive and difficult to document or define. In this study, we relied on computational techniques to objectively identify and explore the evolution of the IS theoretical core and to understand the ways in which scholars have used theories the IS community during the past 24 years. The findings shed new light on theory usage in North American IS research and can serve as a springboard for future research and scholarly debate in this domain. We believe that objective computational tools and techniques such as those we used can be powerfully and uniquely illuminating in this regard, and, therefore, we call for further use of such tools and techniques in efforts aimed at examining the past, present, and future of IS research.

References

Agarwal, R., & Lucas, H. C. (2005). The information systems identity crisis: Focusing on high-visibility and high-impact research. *MIS Quarterly, 29*(3), 381-398.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *The Academy of Management Review, 14*(4), 496-515.

Baskerville, R. L., & Myers, M. D. (2009). Fashion waves in information systems research and practice. *MIS Quarterly, 33*(4), 647-662.

Benbasat, I., & Zmud, R. W. (2003). The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. *MIS Quarterly, 27*(2), 183-194.

Bohannon, J. (2011). Google Books, Wikipedia, and the future of culturomics. *Science, 331*(6014), 135.

Borg, I., & Groenen, P. J. (2005). *Modern multidimensional scaling: Theory and applications (2nd ed.)*. New York, NY: Springer-Verlag.

Box, G. (1988). Signal-to-noise ratios, performance criteria, and transformations. *Technometrics, 30*(1), 1-17. doi: 10.2307/1270311

Crystal, D. (1997). *A dictionary of linguistics and phonetics (4th edition)*. Cambridge, MA: Blackwell Publishing.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-339.

Dennis, A. R., Valacich, J. S., Fuller, M. A., & Schneider, C. (2006). Research standards for promotion and tenure in information systems. *MIS Quarterly*, *30*(1), 1-12.

Eldredge, N., & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In T. Schopf (Ed.), *Models in paleobiology*. San Francisco, CA: Freeman, Cooper & Co.

Ferratt, T. W., Gorman, M. F., Kanet, J. J., & Salisbury, W. D. (2007). IS journal quality assessment using the author affiliation index. *Communications of the Association for Information Systems, 19*, 710-724.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Cambridge, MA: Polity Press.

Ginzberg, M. J. (2012). A business dean's perspective on the IS field. *Data Base for Advances in Information Systems, 43*(2), 7-10.

Google. (2015). *Google ngram viewer*. Mountain View, CA.

Grover, V. (2012). The information systems field: Making a case for maturity and contribution. *Journal of the Association for Information Systems, 13*(4), 254-272.

Hassan, N. R. (2011). Is information systems a discipline? Foucauldian and Toulminian insights. *European Journal of Information Systems, 20*(4), 456-476.

Hirschheim, R. (2007). The looming crisis for the IS field: Where have all the students gone? *Wirtschaftsinformatik, 49*(3), 232-234.

Hirschheim, R., & Klein, H. K. (2012). A glorious and not-so-short history of the information systems field. *Journal of the Association for Information Systems, 13*(4), 188-235.

King, J. L., & Lyytinen, K. (2004). Reach and grasp. *Mis Quarterly, 28*(4), 539-551.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Larsen, K. R. (2014). *Theories used in IS research wiki*. Retrieved from http://istheory.byu.edu

Larsen, K. R., Monarchi, D. E., Hovorka, D. S., & Bailey, C. N. (2008). Analyzing unstructured text data: Using latent categorization to identify intellectual communities in information systems. *Decision Support Systems, 45*(4), 884-896.

Lyytinen, K., & King, J. L. (2006). The theoretical core and academic legitimacy: A response to Professor Weber. *Journal of the Association for Information Systems, 7*(10), 714-721.

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Google Books Team, Picket, J. O., Hoiberg, D., Clancy, D., Norvig, P. Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331*(6014), 176-182.

Moody, D., Iacob, M.-E., & Amrit, C. (2010). *In search of paradigms: Identifying the theoretical foundations of the information systems field*. Paper presented at the 18th European Conference on Information Systems, Pretoria, South Africa.

Neufeld, D., Fang, Y., & Huff, S. (2007). The IS identity crisis. *Communications of the Association for Information Systems, 19*, 447-464.

Peffers, K., & Ya, T. (2003). Identifying and evaluating the universe of outlets for information systems research: Ranking the journals. *Journal of Information Technology Theory and Application, 5*(1).

Rainer, K., & Miller, M. (2005). Examining differences across journal rankings. *Communications of the ACM, 48*(2), 91-94.

Robey, D. (2003). Identity, legitimacy and the dominant research paradigm: An alternative prescription for the IS discipline. *Journal of the Association for Information Systems, 4*(7), 352-359.

Rushing, J. A., Ranganath, H. S., Hinke, T. H., & Graves, S. J. (2001). Using association rules as texture features. *Ieee Transactions on Pattern Analysis and Machine Intelligence, 23*(8), 845-858.

Somers, M. J. (2010). Using the theory of the professions to understand the IS identity crisis. *European Journal of Information Systems, 19*(4), 382-388.

Soper, D. S., & Turel, O. (2012). An n-gram analysis of communications: 2000-2010. *Communications of the ACM, 55*(5), 81-87.

Soper, D. S., & Turel, O. (2015). *Identifying theories used in North American IS research: A bottom-up, computational approach.* Paper presented at the 48th Hawaii International Conference on System Sciences, Koloa, Hawaii.

Suen, C. Y. (1979). N-gram statistics for natural language understanding and text processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 1*(2), 164-172.

Todd, P. (2012). What are they thinking... A view of the IS field “from the dean's office”. *Data Base for Advances in Information Systems, 43*(2), 20-25.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Walsham, G. (2012). Are we making a better world with ICTs? Reflections on a future agenda for the IS field. *Journal of Information Technology, 27*(2), 87-93.

Wastell, D. (2010). Managing as designing: “Ppportunity knocks” for the IS field? *European Journal of Information Systems, 19*(4), 422-431.

Appendix A

The table below provides the field of origin and a brief summary of each core IS theory.

Table A1. Summary of Each Core IS Theory

Theory	Field of origin	Brief summary
Action theory	Sociology	Interconnects the study of social order with micro- and macro-level structural factors and voluntary actions
Agency theory	Economics	Provides insights into situations in which an agent is tasked with acting on behalf of or making decisions on behalf of another person or entity
Capability maturity model	Software engineering	Defines a level of behaviors, practices, and processes which predict varying degrees of success in organizational software processes
Contingency theory	Organization sociology	Describes the best ways in which an organization can be organized in light of its attributes and environment
Control theory	Engineering and mathematics	Addresses the behavior and control of systems with inputs, including modifying system behavior by means of feedback
Deterrence theory	Criminology	Examines the use of punishment and threats of punishment as mechanisms for deterring undesirable behavior
Exchange theory	Sociology	Addresses the roles of subjective cost-benefit analyses and the comparison of alternatives in the formation and maintenance of relationships
Game theory	Economics	Uses mathematical models of conflict and cooperation to study and predict the behavior of rational decision-makers
Institutional theory	Sociology	Describes the processes through which structures are established that become norms or authoritative guidelines for social behavior
Network theory	Computer science	Involves the formal use of graphs to represent and study relationships between and among discrete objects
Prospect theory	Psychology	Describes how people decide among probabilistic alternatives involving risk when the probabilities of the possible outcomes are known
Search theory	Economics	Describes optimal strategies when one must choose from a set of opportunities of random quality, under the assumption that delaying one's choice is costly
Structuration theory	Sociology	Describes how macro-level social systems are created and reproduced through the aggregated micro-level actions of individuals
Technology acceptance model	Information systems	Describes the mechanisms through which individuals decide whether or not to adopt and use a particular technology
Theory of planned behavior	Psychology	Uses intentions to link attitudes, subjective norms, and perceived behavioral control to an individual's behavior
Theory of reasoned action	Psychology	Provides insights into behavior by considering motivations, intentions, and beliefs
Theory of the firm	Economics	Addresses the existence, boundaries, organizational structure, and actions of firms

About the Authors

Daniel Soper is an Associate Professor of Information Systems and Decision Sciences in the Mihaylo College of Business and Economics at California State University, Fullerton. His current research interests lie in the realms of human cognition, interface design, and innovate applications of computational linguistics. His research has appeared in a wide variety of leading IS journals and conference proceedings. In his free time, he enjoys travelling, painting, building websites and software, and attempting to teach himself to play the piano.

Ofir Turel is a Professor of Information Systems and Decision Sciences in the Mihaylo College of Business and Economics at California State University, Fullerton, and a Scholar in Residence at the Brain and Creativity Institute, Department of Psychology at the University of Southern California. His research interests include a broad range of behavioral and managerial issues in various information systems contexts. He has published his work in a wide variety of leading journals, including *MIS Quarterly*, the *Journal of Management Information Systems*, the *MIT Sloan Management Review*, *Communications of the ACM*, the *European Journal of Information Systems*, and *Information & Management*, among many others. Before joining academia, he held senior positions in the information technology and telecom sectors.