In the INFORMS Workshop on Data Science (WDS), 2017, Houston, TX

UNLOCKING OUR BEHAVIORAL KNOWLEDGE INHERITANCE THROUGH ONTOLOGY LEARNING: A DESIGN FRAMEWORK, AN INSTANTIATION, AND A RANDOMIZED EXPERIMENT

Jingjing Li
jl9rf@comm.virginia.edu
University of Virginia

Kai R. Larsen
Kai.larsen@colorado.edu
University of Colorado at Boulder

Ahmed Abbasi
abbasi@comm.virginia.edu
University of Virginia

ABSTRACT

The accumulated literature base in the behavioral sciences represents the most significant source of knowledge about human behavior, yet the same literature has grown beyond human comprehension, resulting in a knowledge inaccessibility problem. Existing IT artifacts such as search engines have not been able to address this issue and in fact, may have intensified it by both rendering low-precision search results and escalating confirmation biases. Following the design science research paradigm, we propose a novel design framework and an instantiation to unlock behavioral knowledge embedded in large-scale behavioral articles. Based on an *ontology learning layer cake* framework and the state-of-the-art text analytics, we implemented a three-step process of extraction and assembly of behavioral theories through *hypothesis*, *construct,* and *construct-relationship extraction*. Linguistics- vs. statistics-based techniques were evaluated and compared in order to determine the best extraction methods. We also developed an ontology-based search engine—*TheoryOn*—that allows researchers to directly search for *constructs* and *synonymous constructs*, *construct relationships*, *antecedents*, and *consequents*, and to easily integrate related *behavioral theories*. We conducted a randomized experiment comparing four information-retrieval tasks for behavioral literature review between *TheoryOn* and EBSCOhost (a full-text search engine) among 38 IS and Management researchers. We found that *TheoryOn* users are significantly better at retrieving relevant constructs, construct relationships, and theories, suggesting significant benefit of our proposed design artifact for addressing the knowledge accessibility problem.

Keywords: Ontology learning, Behavioral theories, Search Engines, Text Analytics, Randomized Experiment, Design Science Research

INTRODUCTION

Behavioral researchers continually search for and develop theories to improve disciplinary understanding of key phenomena. For example, IS has developed or extended hundreds of theories (Soper and Turel 2015), representing important contributions to real-world IS phenomenon, some receiving tens of thousands of citations (Abbasi et al. 2016).[1] Paradoxically, the rich academic literature on human behavior has become expansive to the point of incognizance over the past decades (Kraemer 1991; Marble 2000; Orlikowski and Baroudi 1991; Weber 2012). Studies have shown that researchers remain largely unaware of the majority of research, especially outside their own disciplines (Larsen and Hovorka 2012), but also within narrow research areas (Colquitt and Zapata-Phelan 2007; Larsen and Bong 2016; Larsen 2002). Larsen and Bong (2016) showed that even for a partial set of full-text articles from two top IS journals, experts, on average, could retrieve fewer than 10% of the articles valuable to a literature review and knowledge acquisition, a retrieval rate likely to change by negative orders of magnitude if all relevant research is searched using industry-standard search engines.

The result is knowledge inaccessibility, an issue that has negatively affected IS research in at least four ways. First, with incomplete access to existing knowledge, researchers are prone to literature fragmentation—reinventing constructs, relationships, or hypotheses already introduced by others, or to contradictory findings across different studies. The result is wasted and redundant research efforts (Spell 2001), as well as "fragmented" (Colquitt and Zapata-Phelan

[1] The Technology Acceptance Model [TAM] (Davis 1989; Davis et al. 1989), the Unified Theory of Acceptance and Use of Technology [UTAUT] (Venkatesh et al. 2003b), and the IS Success Model [ISSM] (DeLone and McLean 1992; DeLone and McLean 2002) have all received more than 10,000 citations; Computer Self- Efficacy (Compeau and Higgins 1995), End-User Computing (Doll and Torkzadeh 1988), Task-Technology Fit (Goodhue 1995), Electronic Data Interchange (Iacovou et al. 1995), and eCommerce Trust (Gefen et al. 2003; McKnight et al. 2002) have each received thousands of additional citations.

2007), “theoretically scattered” (Kraemer and Dutton 1991), “conceptually confused” (Checkland 2011), and “chaotic” (Marble 2000) literature. Second, knowledge inaccessibility prevents the building of cumulative traditions. A cumulative tradition requires that “researchers build on each other’s and their own previous work” and that “definitions, topics and concepts are shared” (Keen 1980). It serves a key role in science (Im and Straub 2012; Tsang and Kwan 1999) and is crucial for the persistent development and progression of a research discipline. Third, knowledge inaccessibility introduces inefficiencies in research processes and knowledge acquisition and construction. These inefficiencies leave the research community vulnerable to rapid change, which is especially common in technological areas (Mumford 2003). Finally, the knowledge inaccessibility issue could accrue tremendous monetary and social costs (Alexander et al. 1991; Bong 1996; Weber 2012). Behavioral research spans multiple disciplines, including behavioral medicine, nursing, psychology, sociology, education, communication, management information systems, marketing, management, and economics. Reducing knowledge inaccessibility and enhancing the quality of behavioral literature will have profound practical implications

Beyond the apparent reasons (i.e., sheer numbers of publications and the lack of available time for researchers to read through them), we argue that the existing IT artifacts, such as full-text search engines, are characteristically limited, and thus, are incapable of solving, and may in in fact, worsen, the knowledge accessibility problem. Full-text search engines like Google Scholar and EBSCOhost have similar characteristics. They manage information at the article-level, provide keyword search of the free text in abstracts or full-texts, and incorporate paper-level citation analysis and usage statistics for the ranking of results (Beel et al. 2010). These characteristics result in severe false positives in returned results (Boeker et al. 2013) and confirmation biases (White 2013), which occur as a result of individual researchers and research fields' proclivity toward "unwitting selectivity in the acquisition and use of evidence" (Nickerson 1998, p. 175).

Following the design science paradigm (Gregor and Hevner 2013; Hevner et al. 2004a; Simon 1996), this study proposes two design artifacts—a behavioral ontology learning design framework and its instantiation, named TheoryOn—to alleviate the knowledge inaccessibility problem in the behavioral sciences and to address the weaknesses of existing IT artifacts. We adopt Weber's (2012) view that a behavioral theory "accounts for some subset of phenomena in the real world" and is a specialized type of Bunge's (1977; 1979) ontology.[2] The constructs of Bunge's ontology share many of the common constructs of a behavioral theory (i.e., constructs, their associations, and the states they cover—meaning, the theory "parts" Weber, 2012, p. 6}. Therefore, we use the ontology learning layer cake (Buitelaar et al. 2005)—a process of extracting relevant parts of ontologies (i.e., concepts, relations, and axioms) from texts by using a collection of techniques and resources—as a kernel theory to guide our design process of extracting behavioral theories from existing, large-scale behavioral publications. We narrow our focus to a manageable initial level by focusing on behavioral positivist research, and specifically those fitting the criteria of Gregor's (2006) theories for explanation and theories for explanation and prediction (natural science types of research). Our design framework and instantiation, however, should be extendable in the future to positivist case studies (e.g., Lee 1989),

[2] There may be alternative notions about the mapping between behavioral theories and ontologies, but such is not the focus of this paper. By adopting Weber's view, many of the ontology-learning tasks and techniques can be nicely adapted to guide extracting behavioral theories from a large-scale behavioral publication.

interpretive studies (e.g., Klein and Myers 1999), and process studies (e.g., Kettinger et al. 1997), as well as other types of theories.

Like Abbasi and Chen (2008), we illustrate the usefulness of the proposed design framework by developing an instance, in this case an ontology-based search engine named *TheoryOn* that extracts *hypotheses, constructs,* and *theoretical relationships* from hundreds of relevant behavioral studies published at *MIS Quarterly*, *Information Systems Research*, and the *Journal of Applied Psychology*—all top journals in their fields (Li et al. 2016a; 2016b; 2017; 2018a; 2018b; 2019). With the extracted theory "parts," *TheoryOn* allows researchers to search directly for *constructs*, *construct relationships*, and *theoretically related constructs* (e.g. antecedents or consequents), as well as to easily integrate *related theories*.

Following the evaluation guidelines by Hevner et al. (2004a) and Gill and Hevner (2013), we evaluate the performance of *TheoryOn* by comparing several existing ontology-learning approaches and conducting a randomized experiment that illustrates the practical importance of *TheoryOn* by comparing it with a full-text academic search engine (expressly, EBSCOhost). Specifically, four typical tasks of a behavioral research literature review (Webster and Watson 2002)—construct search, construct relationship search, antecedent and consequent search, and theory integration—were assigned to both *TheoryOn* and EBSCOhost users. Their behavioral information retrieval performances were compared using precision and recall, which illustrate the tendency to reduce false positives and false negatives in an information retrieval task, respectively. On average, *TheoryOn* users were 16.84% to 72.66% better. Additionally, we found that *TheoryOn* was perceived to be more useful and easier to use than EBSCOhost, as evaluated by Usefulness and Ease of Use scales adapted from the *Technology Acceptance Model* (Davis 1989; Davis et al. 1989).

AN ONTOLOGY LEARNING-BASED DESIGN FRAMEWORK FOR DISEMBEDDING BEHAVIORAL KNOWLEDGE

Design Framework for Disembedding Behavioral Knowledge

This study represents the first effort of extracting behavioral ontologies from texts. Building on work by Wong (2012), Figure 1 represents a first design framework to map behavioral theory extraction into an ontology-learning framework. Accordingly, the ontology learning for behavioral theory could be broken into five tasks: *hypothesis extraction, construct extraction and grouping, theoretical relationship extraction, construct hierarchy building,* and *theoretical relationship discovery*. Each task generates an output corresponding to ontology learning's five outputs: *terms, concepts, non-taxonomic relations, taxonomic relations,* and *axioms*, respectively (Fu et al. 2008; 2010; 2012).

--- Insert Figure 1 about here ---

Hypothesis Extraction

Hypotheses containing all the lexical components delineating constructs and construct relationships represent the basic building blocks for behavioral ontology learning, which are corresponding to the terms in the ontology-learning layer cake. While hypothesis vernacular changes by discipline and some disciplines lack accorded hypothesis formats, in general hypothesis formats for several behavioral disciplines (e.g., IS, management, marketing) are prescribed. However, we admit that hypotheses in many other behavioral research areas may not follow the traditional format (e.g., hypotheses in many nursing studies are nested into the sentence or paragraph supporting the argument and do not have unique hypothesis labels). The differentiating factors may be that certain words or phrases are used more in hypothesis sentences than regular sentences. Therefore, we can use statistic-based techniques to discover this unique pattern. Specifically, a supervised machine learning approach that relies on a training dataset to select linguistic features correlate to hypothesis sentences and build a text classifier to automatically label hypothesis sentences can be used. In summary, we posit that rule-based and supervised machine learning approaches can be used to extract hypotheses, and the choice of which should be determined by the specific disciplines or journals.

Construct Extraction and Grouping

For the purposes of the initial ontology learning, the distinction between constructs and objectively measurable constructs (e.g. demographic constructs)[3] is not substantive. Constructs in behavioral theories correspond to concepts in ontologies (Weber 2012). Therefore, the steps to extract construct instances and group them based on the same latent constructs exactly matches the concepts extraction steps in ontology learning. We assume that constructs in different articles, even when they have identical names, represent different ontological concepts until empirically analysis indicates otherwise. Therefore, the process of grouping several construct instances is initially scoped with a behavioral article where two mentions of a construct almost certainly refer to the same concept.

Due to the diversity of variable embedding forms, we suggest the use of supervised machine learning methods, such as Hidden Markov Models (Rabiner 1989) and Conditional Random Fields (Lafferty et al. 2001), to extract construct instances embedded in the hypotheses.

[3] We use the term "construct" to refer to both behavioral construct and non-construct variables.

After variable instances embedded in the hypotheses from an article are successfully extracted, we need to group construct instances referring to the same underlying concept together. Since researchers tend to use similar phrases to represent variables within the same article, this represents a relatively straight-forward problem.

Theoretical Relationship Extraction

After constructs embedded in hypotheses are extracted, we need to assess the relationships between them. In this step, we focus on extracting non-taxonomic relationships from constructs. We define the correlational or causal relationships among constructs in a hypothesis as theoretical relationships, which combine into the proposed behavioral theories in an article. Specifically, a theoretical relationship could be categorized as *main effect*, *moderation*, and *mediation* (Baron and Kenny 1986). Main effect pertains to a causal relationship between two constructs, whereas moderation and mediation involve relationships among more than two constructs (Figure 2a). Specifically, in a moderation relationship, a moderator is a third construct that affects the strength or direction of the relationship from an independent construct to a dependent construct (Figure 2b); in a mediation relationship, a mediator is a third construct serving as an intermediate construct between an independent construct to a dependent construct (Figure 2c).

--- Insert Figure 2 about here ---

Following shows a hypothesis (annotated with extracted constructs) represent a moderation relationship from Venkatesh et al. (2003a)

H1: [Perceived usefulness] will influence [behavioral intention to use a system] more strongly for [men]$_{gender}$ than it will influence [women]$_{gender}$.

The construct extraction and grouping steps identified two constructs, *perceived usefulness* and *behavioral intention to use a system*, and two construct instances, *women* and *men*, that refer to a demographic construct, *gender*. The theoretical relationship extraction step will identify how these three constructs are connected to form a moderation effect. In this case, *gender* (grouped from the construct instances *women* and *men* in the *construct grouping* step) serves as a moderator for the effect from *perceived usefulness* to *behavioral intention to use a system* (Figure 3).

--- Insert Figure 3 about here ---

To identify such a relationship, one should leverage the syntactic features of the hypothesis, e.g. *perceived usefulness* is the subject and *behavioral intention to use a system* is the object of a verb phrase containing “influence”, as well as the behavioral knowledge about what forms a moderation effect, e.g. “more strongly for men than it will influence women” indicate a comparison of the magnitude of influence between different types of gender— gender is a moderator. Based on a number of studies (e.g., Maynard et al. 2009; Tan et al. 2016), we posit that extracting a domain-specific complex relationship usually requires a hybrid approach that combining statistics- and linguistics-based methods. Specifically, *part-of-speech tagging/sentence parsing* and *syntactic structure analysis/dependency analysis*, usually requiring a combination of statistics- and linguistics-based methods to provide linguistic component for *lexico-syntactic patterns* that describe the types (e.g., main, moderation or mediation) and directionality of the theoretical relationships (e.g., A affects B through C).

Once theoretical relationships are extracted and shared constructs among different hypotheses are identified through construct grouping method introduced in the previous step, a behavioral theory ontology could be automatically extracted from each behavioral articles that contain hypotheses. Figure 4 shows an example of behavioral ontology extracted from nine hypotheses in Venkatesh et al. (2003a).

--- Insert Figure 4 about here ---

Construct Hierarchy Building

Another type of construct relationship in the ontology learning layer cake is the taxonomic relationship, pertaining to “is-a” (hypernym/hyponym) or synonymous/polysemous relationships. This type of relationship is also prevalent among behavioral constructs. For example, the construct *performance expectancy* in Venkatesh (2003a) is synonymous with the *perceived usefulness* proposed by Davis (1989). However, according to Larsen and Bong (2016), taxonomic relationships between constructs in different articles are not widely recognized by behavior researchers. Without knowing the pre-existing construct due to insufficient literature review or constant construct renaming, researchers are unconsciously “re-inventing the wheels”, resulting in the “construct identity fallacy” (Larsen and Bong 2016). Therefore, it is imperative to identify construct taxonomic relationships so as to alleviate the construct identify fallacy and enable researchers to better understand the existing literature and improve the research efficiency. On the other hand, once the synonymous constructs from different behavioral ontologies are identified, we can establish a “nomological network” (Cronbach and Meehl 1955) to identify relatively under studied areas, paving the way for axiom detection.

Theoretical Relationship Discovery

After construct instances embedded in the hypotheses from an article are successfully extracted, we need to use group construct instances referring to the same underlying construct together. Since researchers tend to use similar phrases to represent constructs within the same article, we can use lexical similarity analyses, such as the minimum edit distance (Levenshtein 1966) to identify construct instances with subtle name variations.

THEORYON – AN INSTANTIATION OF THE PROPOSED DESIGN FRAMEWORK

According to the identified behavioral knowledge search needs, we developed four corresponding functionalities. The functionality interface and descriptions are as follows:

a) Construct Search. TheoryOn allows users to specify a construct in a search query and only return articles that containing this construct or its synonymous constructs. The construct information is directly presented in the returned results. Users can also save the related constructs and articles in a sorting hierarchy. Figure 5 shows a search for *perceived usefulness* using a combination of keyword and Latent Semantic Analysis search.[4] Retrieved constructs are shown with citation information and the ability to examine definitions, items, and operationalization origin (citations) as well as to start a new semantic or taxonomic search with the current construct as the starting point. When a nomological network has been extracted from the paper, it is visualized along with the construct information and the target construct marked in yellow. For more details, watch the video "TheoryOn: Synonymous Construct Search".[5]

--- Insert Figure 5 about here ---

[4] When Latent Semantic Analysis is used for information retrieval it is more accurately referred to as Latent Semantic Indexing (Deerwester et al. 1990), but our implementation of the algorithm is in line with how LSA works as construct texts are projected into an existing semantic space.

[5] The four videos in this section have been blinded for peer review.

b) Construct-Pair Search. TheoryOn allows users to specify a construct pair in a search query and only returns articles that containing these two constructs. The constructs (marked in yellow) and their relationships are shown in the extracted theoretical models in the left part of the search results. For more details, watch the video "TheoryOn: Construct-Pair Search".

--- Insert Figure 6 about here ---

c) Theoretically Related Construct Search. This functionality allows inspection of the theoretical models containing a construct of interest (highlighted in yellow) as well as examination of its antecedents and consequents in a list or plot view (Figure 7). This functionality takes the first *n* papers returned by the construct search and displays the antecedents to the searched-for construct. It then does the same for the antecedents. Incremental use of the tool should allow future high-quality categorization of constructs so that the selection of an *n* would not be necessary. This would also likely simplify the display as the three *Ease of Use* constructs could be combined. The same would be true about the consequents as four of the constructs could likely be integrated into a share *Behavioral Intention* group. For more details, watch the video "TheoryOn: Theoretically Related Construct Search".

--- Insert Figure 7 about here ---

d) Theory Integration. All the related theories can be saved in the sorting hierarchy and visualized on the canvas. A user can then integrate theories by clustering synonymous constructs or customize the theoretical networks by editing, deleting or adding any nodes and links. This tool represents an important future goal. As users categorize the variables they need for the future, data about their decisions allows for user-driven ontology development that will benefit the whole discipline after use. For more details, watch the video "TheoryOn: Theory Integration".

--- Insert Figure 8 about here ---

EVALUATION

Experiments to examine text analytics components' performance

According to our proposed design framework, we applied various natural language processing and text mining models to address each step in the behavioral ontology learning framework. We followed the best practice in NLP to construct the most relevant features and select the best model parameters for the chosen models. Due to the page limits, we do not put the details here. We use traditional information extraction metrics, precision, recall and F-measures to evaluate the system performance. In general, the F-measure for hypothesis extraction is around 91%, variable extraction is 72%. Variable relationship extraction is around 84%.

User experiment to examine system effectiveness in supporting tasks

To evaluate the usability and usefulness of *TheoryOn*, we conducted a randomized experiment with 38 Systems and Organizational Behavior Ph.D. students from a variety of programs in the U.S. and around the world. We designed four tasks commonly carried out in a literature review process and evaluated the performance of *TheoryOn* against the control group that used a common full-text search engine, EBSCO-host. We selected the full-text search engine, Business Source Complete database (BSC) powered by EBSCOhost, because it has one of the largest scholarly full-text business databases. BSC also represented, at the time of experimental process, the longest uninterrupted period of full-text coverage for MISQ, ISR, and JAP (1990-2009).

Tasks

To test the *TheoryOn* system's utility, we designed four tasks for each participant to complete: *synonymous construct search*, *construct pair search*, *antecedents and consequents search* and *theory integration*, each of which is a common literature review task for behavioral research and is related to one of the proposed system functionalities of knowledge retrieval. All four tasks are related to one theory, in this case the Technology Acceptance Model (TAM), which demonstrate a natural progression of knowledge acquisition, curation, and integration in a literature review process (see Table 1).

For each task, the participants were given as a starting point, an example of a construct, a construct pair, or a theory, along with necessary details such as, construct definition and items (see Table 1). In order to familiarize the participants with the functionalities of *TheoryOn* and EBSCOhost, a short video tutorial (3–5 minutes) was given for each task. The participants were required to complete each tasks in less than an hour.

--- Insert Table 1 about here ---

Evaluation Methods

According to Hevner et al. (2004b), we evaluated *TheoryOn*'s performance against that of the experimental groups using two types of metrics: objective and perceptual. Objective evaluation compares the construct, article, and theory retrieval performance including precision and recall (Salton 1989), whereas subjective evaluation takes place the realm of human perception and taste and tries to point to the perceived utility of the artifact.

Objective Metrics

We adopted precision, recall, and the F_1-measure commonly used in information retrieval evaluation to assess the task performance. Each participant's submission was compared against a carefully constructed gold standard set. The gold standard for each task was rigorously constructed by a team of two experienced faculty researchers, three doctoral students, and four senior research assistants (RAs with at least 500 hours of experience in construct extraction from behavioral papers). For all gold standard evaluations, the TheoryOn participants outperformed EBSCO participants (see Table 2)

--- Insert Table 2 about here ---

Perceptual Metrics

Following the evaluation guidelines by Hevner et al. (2004a) and Gill and Hevner (2004a), we adapted multiple scales to evaluate the perceptual utility of *TheoryOn*. Specifically, immediately after completing each task, the participants were asked to report on how the system helped with each specific task and whether or not it increased perceived confidence regarding their submissions by a 4-item *Usefulness* scale adapted from Venkatesh et al. (2003b). In addition, for each task, we asked three questions related to *Task Experience* to make sure there were no significant differences in task familiarity between the two experimental groups. After the participants completed all tasks, they were asked to report on their perception of three TAM constructs adapted from Davis (1989) and Venkatesh et al. (2003b): a 4-item *Perceived Usefulness*, a 4-item *Perceived Ease of Use*, and a 3-item *Behavioral Intention to Use* scale.

Across four tasks related to our proposed system functionalities, we found no significant difference in task experience ($p>0.05$), but the Perceived Usefulness of *TheoryOn* was significantly higher than EBSCOhost ($M_t = 6.09$, $M_e = 4.82$, $p<0.001$; $M_t = 5.96$, $M_e = 5.24$, $p<0.05$; $M_t = 6.41$, $M_e = 4.72$, $p<0.001$; $M_t = 5.67$, $M_e = 4.43$, $p<0.01$). As specified in Table 3, *TheoryOn* tremendously helps users to find synonymous constructs, antecedents and consequents, and extension theories with a whopping difference greater than 1.3 points on a 7-point Likert scale. Regarding overall utility perception at the system level, *TheoryOn* was considered to be easier to use and useful ($M_t = 5.88$, $M_e = 4.71$, $p<0.01$; $M_t = 6.20$, $M_e = 5.17$, $p<0.01$), whereas the behavioral intention to use the system was marginally significant ($M_t = 5.53$, $M_e = 4.70$, $p = 0.07$).

--- Insert Table 3 about here ---

CONCLUSION

This study proposes two design artifacts—a behavioral ontology learning design framework and its instantiation, named TheoryOn—to alleviate the knowledge inaccessibility problem in the behavioral sciences. Our contributions are manifold. First, we propose an ontology learning design framework specific for behavioral research to guide incremental development of behavioral theory knowledge-management systems. Second, we outline a research agenda, or map, for the behavioral ontology learning research area. Third, we instantiate the framework into an ontology-based search-engine artifact, namely *TheoryOn*, to show the applicability of the framework. Finally, we demonstrate the value of a knowledge base for behavioral research findings through a randomized experiment. Overall, the knowledge contribution of this research represents an instance of *exaptation* in which we adapted solutions from the ontology-learning field to a new problem of extracting behavioral theories from large- scale behavioral publications (Gregor and Hevner 2013). We believe the work has important implications for disembedding behavioral knowledge in various social science domains including IS and health (Netemeyer et al. 2019; Zahedi et al. 2015; Zimbra et al. 2010), including potential for predicting behavioral relationships (Brown et al. 2015a; 2015b). Our work also advances the state-of-the-art for natural language processing (Kitchens et al. 2018; Deng et al. 2018; Zimbra et al. 2018; Adjeroh et al. 2014; Benjamin et al. 2014) and text analytics (Abbasi et al. 2018a; 2018b; 2019; Ahmad et al. 2019; Khaja et al. 2018).

FIGURES AND TABLES

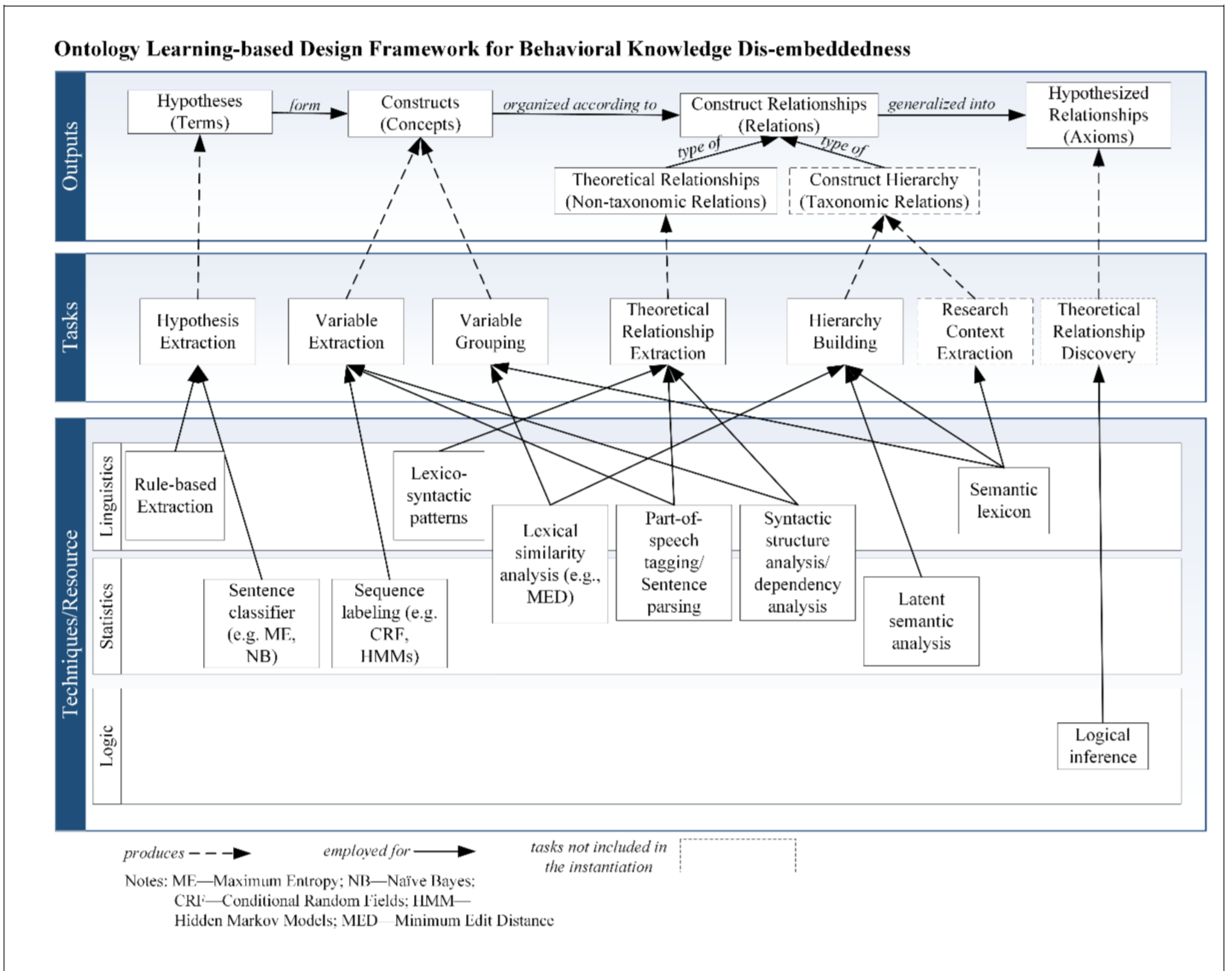

Figure 1. Proposed Design Framework for Disembedding Behavioral Knowledge.

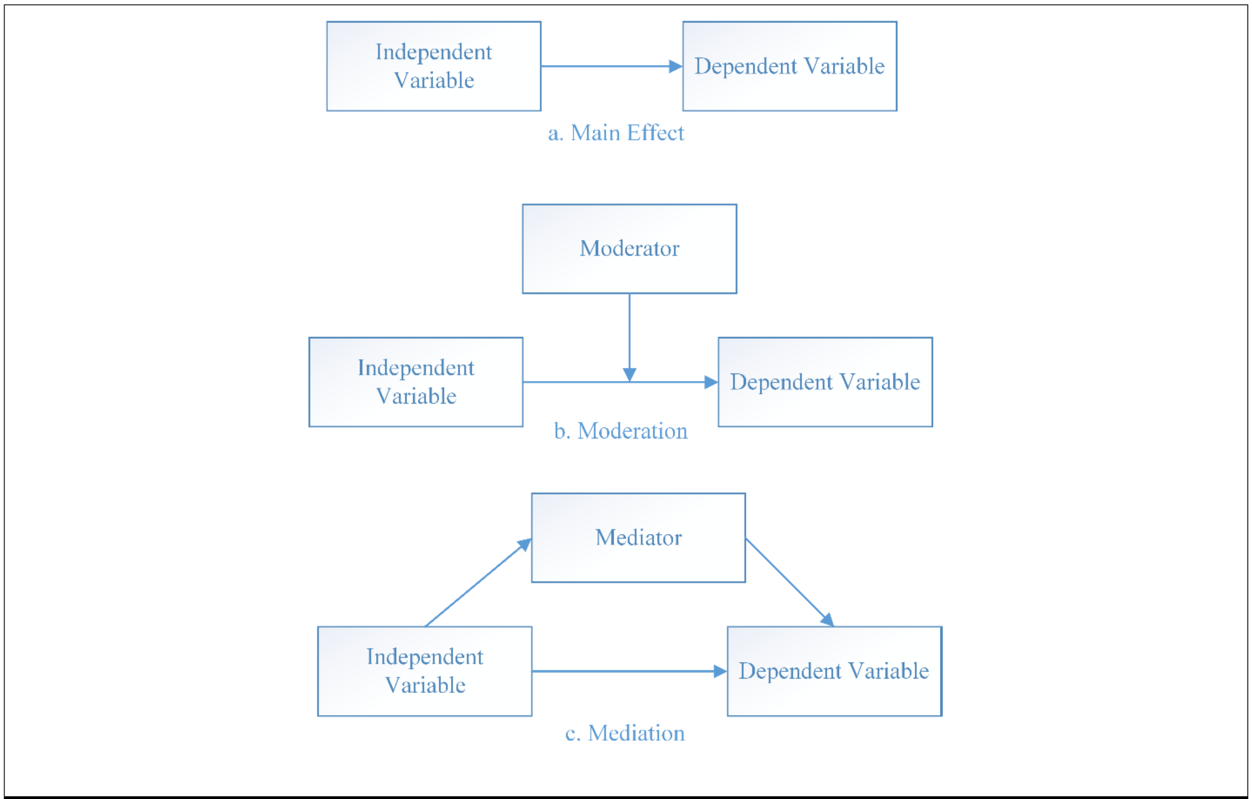

Figure 2. Three types of theoretical relationships.

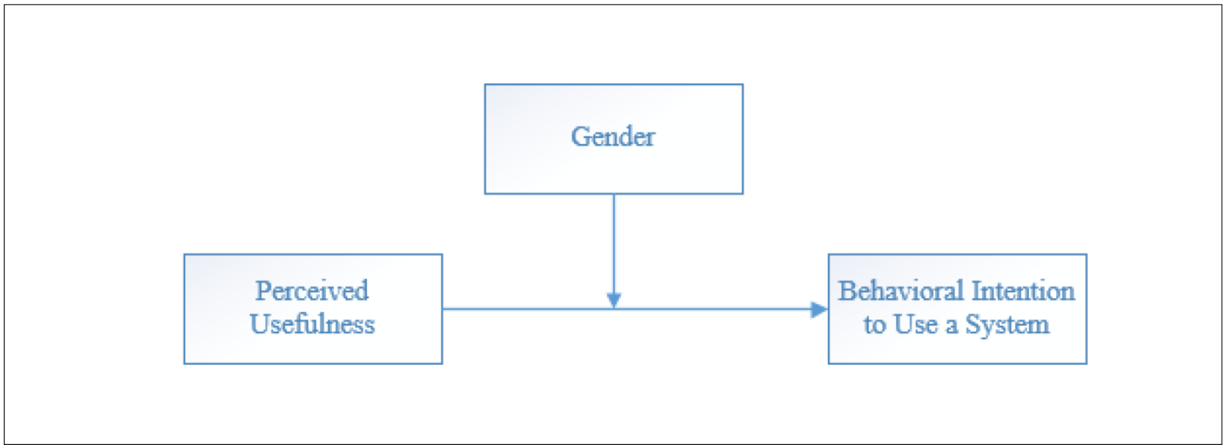

Figure 3. Theoretical relationship embedded in H_1 in Venkatesh et al. (2003).

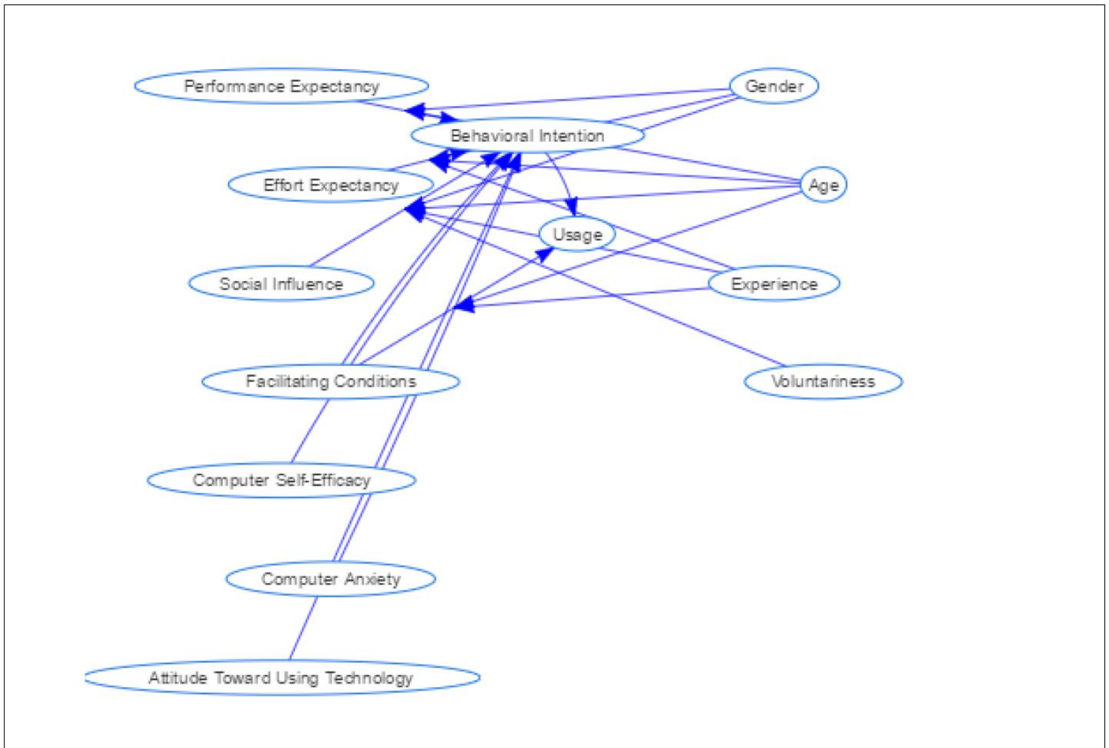

Figure 4. Behavioral ontology automatically extracted from Venkatesh et al. (2003a).

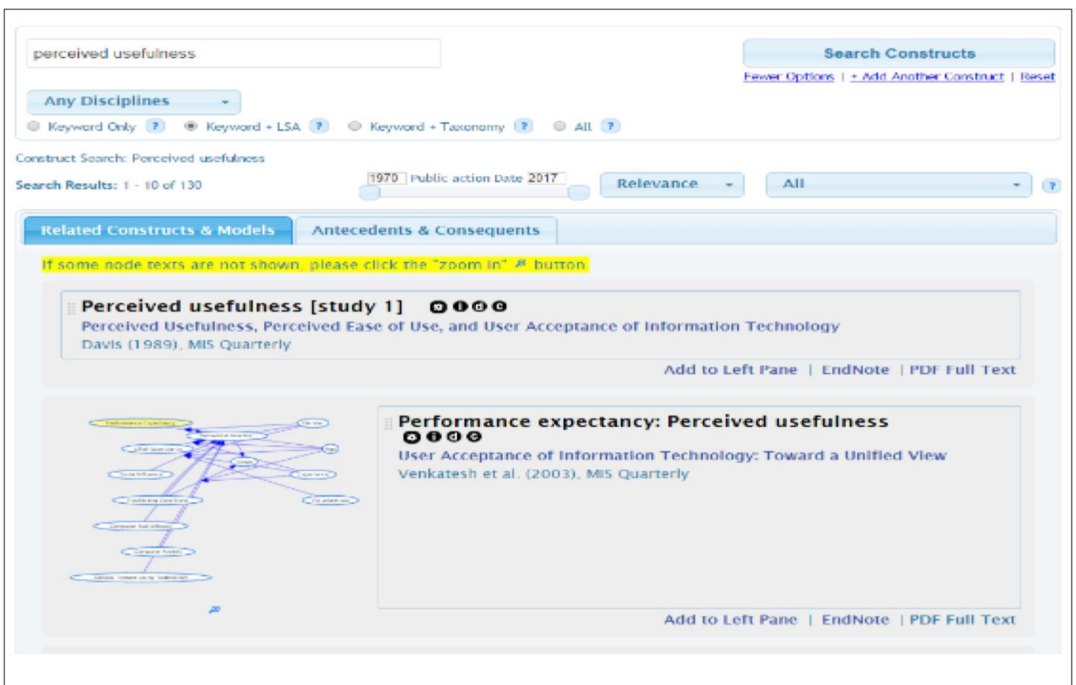

Figure 5. TheoryOn Functionality: Construct Search

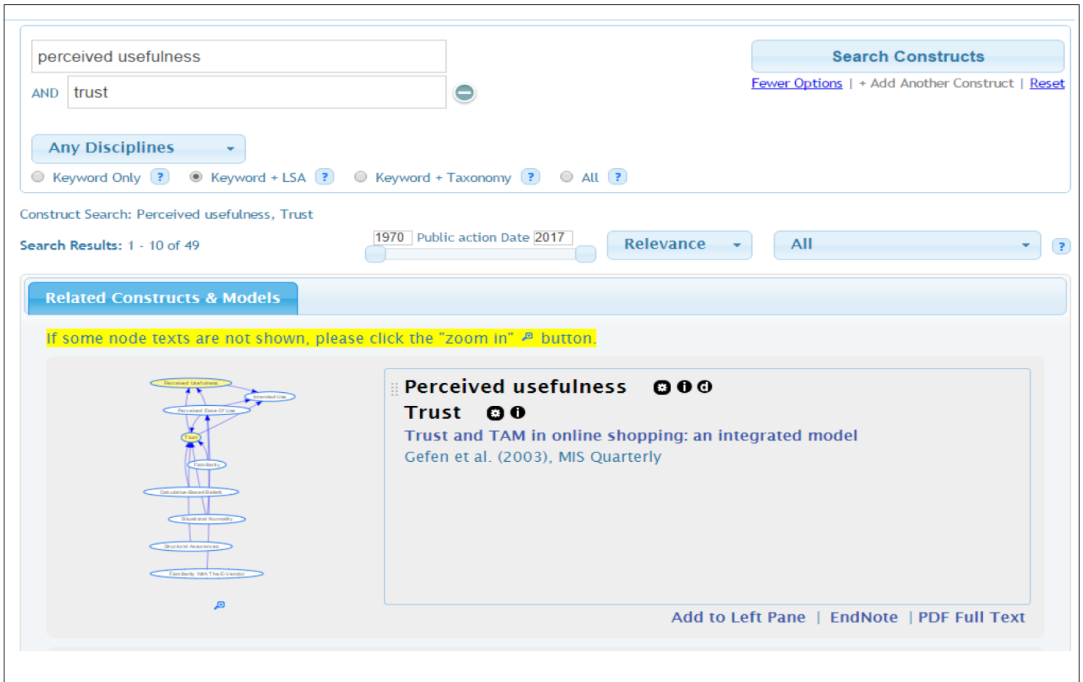

Figure 6. TheoryOn Functionality: Construct Pair Search

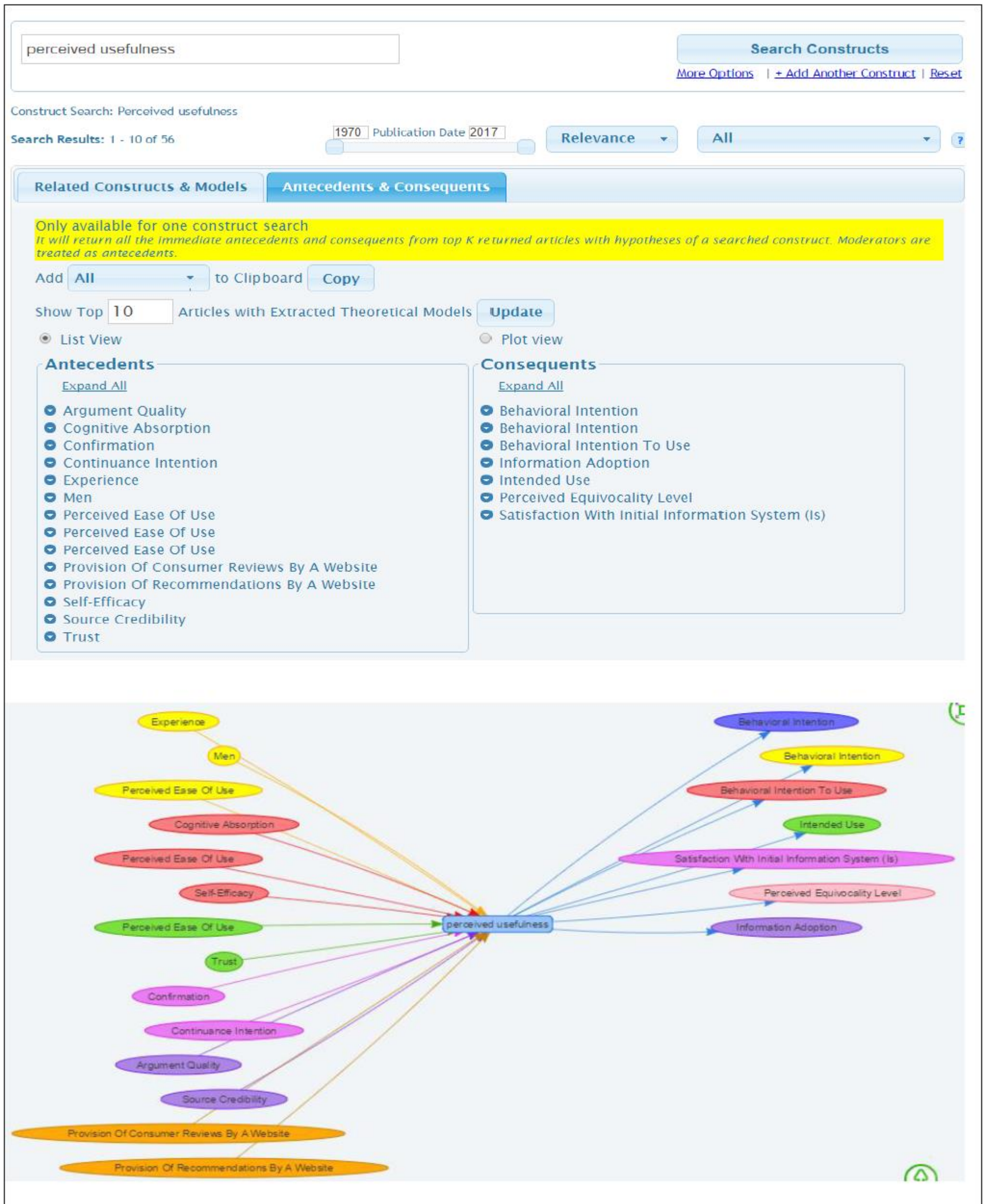

Figure 7. TheoryOn Functionality: Related Construct Search

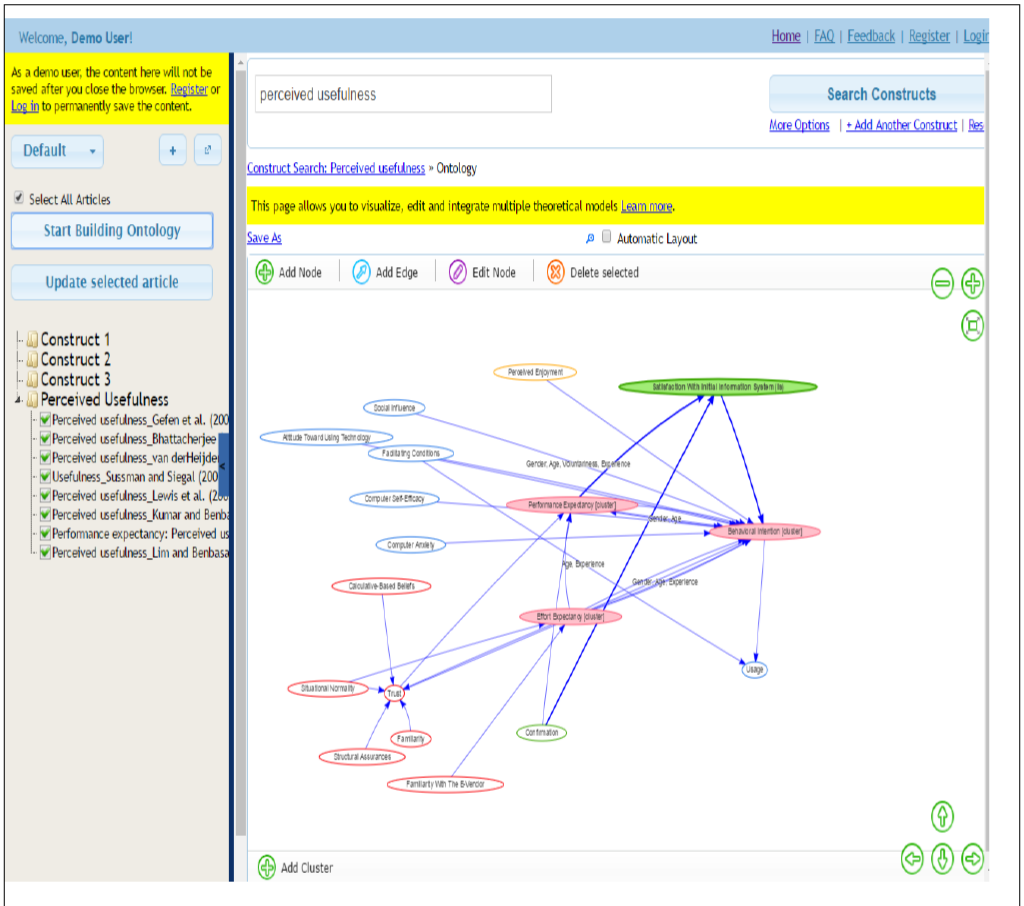

Figure 8. TheoryOn Functionality: Theory Integration

Table 1. Tasks in the Randomized Experiment.

Task Description/Submission	Construct/Definition	Sample of Items
Synonymous Construct Search: Find as many synonymous constructs as possible for **Perceived Usefulness** **Submission**: Synonymous constructs along with their article information	Perceived Usefulness (Davis 1989; Venkatesh et al. 2003): The degree to which a person believes that using a particular system would enhance his or her job performance.	• Using the system in my job would enable me to accomplish tasks more quickly. • Using the system would improve my job performance. • Using the system in my job would increase my productivity. • Using the system would enhance my effectiveness on the job. • I would find the system useful in my job.
Construct Pair Search: Find as many articles as possible that contain both **Perceived Usefulness** (See Task 1 Definition) and **Trust**, including articles that contain both of their synonymous counterparts. **Submission**: Articles containing both constructs (including synonymous constructs)	Trust(Choudhury and Karahanna 2008): A user's beliefs about the reliability, credibility, and accuracy of information gathered through the web.	• I would have greater confidence in the explanations provided by such web sites than in those offered by an agent. • I would trust the validity of quotes provided by this web site more than those provided by an agent. • I believe such a web site would provide more objective recommendations than an agent would provide. • I would feel more confident purchasing the policy through the web than through an agent.
Antecedents and Consequents Search: For the construct **Perceived Usefulness**, find as many immediate antecedents and consequents as possible, i.e., the constructs that are hypothesized to directly influence or be influenced by **Perceived Usefulness**. **Submission**: Immediate antecedents and consequents with their article information	See Task 1	See Task 1
Theory Integration: Extend the original Technology Acceptance Model (TAM) (Davis 1989) by integrating relevant hypothetical relationships through constructs synonymous with **Perceived Usefulness**, **Perceived Ease of Use**, and **Behavioral Intention to Use**. Each article must contain **Behavioral Intention** and at least one construct from **Perceived Usefulness** and **Perceived Ease of Use**. **Submission**: Articles that integrated with TAM and an expanded TAM model diagram	Perceived Ease of Use (Davis 1989; Venkatesh et al. 2003): The degree to which a person believes that using a system would be free of effort.	• Learning to operate the system would be easy for me. • I would find it easy to get the system to do what I want it to do. • My interaction with the system would be clear and understandable. • I would find the system to be flexible to interact with. • I would find the system easy to use.
	Behavioral Intention to Use (Davis 1989; Venkatesh et al. 2003): Participant's intention to use the technology.	• I intend to use the system in the next n months. • I predict I would use the system in the next n months. • I plan to use the system in the next n months.

Table 2. Percentage Retrieval Performance by Task

Task	TheoryOn (n = 18)			EBSCOhost (n = 17)		
	Precision	Recall	F1	Precision	Recall	F1
1. Synonymous Construct Search	95.2	27.3	**40.1**	81.7	16.2	26.4
2. Construct Pair Search	76.7	43.9	**51.6**	72.0+	24.7	34.9
3a. Antecedent Search	86.3	29.3	**41.5**	72.2	13.4	21.8
3b. Consequent Search	80.2	23.8	**34.7**	68.9	16.4	25.3
4. Theory Integration	77.4	25.4	**34.6**	61.9	16.0	23.9

Note: + not significantly different from TheoryOn ($p > 0.05$)

Table 3. Perceived Usefulness Comparison of TheoryOn and EBSCOhost

	TheoryOn (n = 18)		EBSCOhost (n = 17)		
Construct	Mean	SD	Mean	SD	Diff (t-stat)
PU	5.92	0.73	5.01	1.01	3.04**
EU	6.21	0.58	5.47	1.28	2.21*
BI	5.57	1.21	4.84	1.26	1.74
PU1	6.11	0.54	5.21	0.94	3.54**
PU2	5.90	0.75	5.44	1.03	2.14*
PU3	6.44	0.60	4.85	1.44	4.30***
PU4	5.67	0.99	4.72	1.54	2.17**
TE1	5.00	1.19	4.61	1.30	0.93
TE2	5.69	0.92	5.24	1.14	1.29
TE3	5.26	1.35	4.82	1.24	0.99
TE4	4.22	1.46	4.63	1.47	−0.82

Notes: 1. *$p < 0.050$; **$p < 0.010$; ***$p < 0.001$ 2. PU: Perceived usefulness of the system; EU: ease of use of the system; BI: behavioral intention to use the system. PU1–4 are the perceived usefulness for each task. TE1–4 are the prior experience with each of the tasks; diff (t-stat) is the t statistics of EBSCOhost or Google Scholar compared with TheoryOn.

REFERENCES

Abbasi, A., and Chen, H. 2008. "Cybergate: A Design Framework and System for Text Analysis of Computer-Mediated Communication," *MIS Quarterly* (32:4), pp. 811-837.

Abbasi, A., Sarker, S., and Chiang, R. H. L. "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems*, 17(2), no. 3, 2016.

Abbasi, A., Li, J., Clifford, G. D., and Taylor, H. A, "Make 'Fairness By Design' Part of Machine Learning," *Harvard Business Review*, August 5, 2018a, digital article: https://hbr.org/2018/08/make-fairness-by-design-part-of-machine-learning

Abbasi, A. Zhou, Y., Deng, S., and Zhang, P. "Text Analytics to Support Sense-making in Social Media: A Language-Action Perspective," *MIS Quarterly*, 42(2), 2018b, pp. 427-464.

Abbasi, A., Li, J., Adjeroh, D., Abate, M., and Zheng W. "Don't Mention It? Analyzing User-generated Content Signals for Early Adverse Event Warnings," *Information Systems Research*, 30(3), 2019, pp. 1007-1028.

Adjeroh, D., Beal, R., Abbasi, A., Zheng, W., Abate, M., and Ross, A. "Signal Fusion for Social Media Analysis of Adverse Drug Events," *IEEE Intelligent Systems*, 29(2), 2014, pp. 74-80.

Ahmad, F., Abbasi, A., Li, J., Dobolyi, D., Netemeyer, R., Clifford, G., and Chen, H. "A Deep Learning Architecture for Psychometric Natural Language Processing," *ACM Transactions on Information Systems*, forthcoming.

Alexander, P. A., Schallert, D. L., and Hare, V. C. 1991. "Coming to Terms: How Researchers in Learning and Literacy Talk About Knowledge," *Review of educational research* (61:3), pp. 315-343.

Baron, R. M., and Kenny, D. A. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of personality and social psychology* (51:6), p. 1173.

Beel, J., Gipp, B., and Wilde, E. 2010. "Academic Search Engine Optimization (Aseo) Optimizing Scholarly Literature for Google Scholar & Co," *Journal of scholarly publishing* (41:2), pp. 176-190.

Benjamin, V., Chung, W., Abbasi, A., Chuang, J., Larson, C. A., and Chen, H. "Evaluating Text Visualization for Authorship Analysis," *Security Informatics*, 3(10), 2014.

Boeker, M., Vach, W., and Motschall, E. 2013. "Google Scholar as Replacement for Systematic Literature Searches: Good Relative Recall and Precision Are Not Enough," *BMC medical research methodology* (13:1), p. 131.

Bong, M. 1996. "Problems in Academic Motivation Research and Advantages and Disadvantages of Their Solutions," *Contemporary Educational Psychology* (21:2), pp. 149-165.

Brown, D. E., Abbasi, A., and Lau, R. Y. K., "Predictive Analytics: Predictive Modeling at the Micro Level," *IEEE Intelligent Systems*, 30(3), 2015, pp. 6-8.

Brown, D. E., Abbasi, A., and Lau, R. Y. K., "Predictive Analytics," *IEEE Intelligent Systems*, 30(2), 2015, pp. 6-8.

Buitelaar, P., Cimiano, P., and Magnini, B. 2005. *Ontology Learning from Text: Methods, Evaluation and Applications*. IOS press.

Bunge, M. 1977. "Emergence and the Mind," *Neuroscience* (2:4), pp. 501-509.

Bunge, M. A. 1979. "Treatise on Basic Philosophy: Ontology Ii: A World of Systems. Dordrecht, Holland: D." Reidel Publishing Company.

Checkland, P. 2011. "Systems Thinking and Soft Systems Methodology," in *The Oxford Handbook of Management Information Systems: Critical Perspectives and New Directions,* R.D. Galliers and W. Currie (eds.). Oxford Publishing.

Choudhury, V., and Karahanna, E. 2008. "The Relative Advantage of Electronic Channels: A Multidimensional View," *MIS quarterly*), pp. 179-200.

Colquitt, J. A., and Zapata-Phelan, C. P. 2007. "Trends in Theory Building and Theory Testing: A Five-Decade Study of the Academy of Management Journal," *Academy of Management Journal* (50:6), pp. 1281-1303.

Compeau, D. R., and Higgins, C. A. 1995. "Computer Self-Efficacy: Development of a Measure and Initial Test," *MIS Quarterly* (19:2), pp. 189-211.

Cronbach, L. J., and Meehl, P. E. 1955. "Construct Validity in Psychological Tests," *Psychological Bulletin* (52:4), pp. 281-302.

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 319-340.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. 1989. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science* (35:8), pp. 982-1003.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., and Harshman, R. 1990. "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science* (41:6), pp. 391-407.

DeLone, W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), pp. 60-95.
DeLone, W. H., and McLean, E. R. 2002. "Information Systems Success Revisited," *Hawaii International Conference on System Sciences*: IEEE, pp. 1-11.
Deng, S., Zhang, P., Zhou, Y., and Abbasi, A. "Using Discussion Logic in Analyzing Online Group Discussions: A Text Mining Approach," *Information and Management*, 56(4), 2019, pp. 536-551.
Doll, W. J., and Torkzadeh, G. 1988. "The Measurement of End-User Computing Satisfaction," *MIS Quarterly* (12:2), pp. 259-273.
Fu, T., Abbasi, A., and Chen, H. "A Hybrid Approach to Web Forum Interactional Coherence Analysis," *Journal of the American Society for Information Science and Technology*, 59(8), 2008, pp. 1195-1209.
Fu, T., Abbasi, A., and Chen, H. "A Focused Crawler for Dark Web Forums," *Journal of the American Society for Information Science and Technology*, 61(6), 2010, pp. 1213-1231.
Fu, T., Abbasi, A., Zeng, D., and Chen, H. "Sentimental Spidering: Leveraging Opinion Information in Focused Crawlers," *ACM Transactions on Information Systems*, 30(4), 2012, no. 24.
Gefen, D., Karahanna, E., and Straub, D. W. 2003. "Trust and Tam in Online Shopping: An Integrated Model," *MIS Quarterly* (27:1), pp. 51-90.
Gill, T. G., and Hevner, A. R. 2013. "A Fitness-Utility Model for Design Science Research," *ACM Transactions on Management Information Systems (TMIS)* (4:2), p. 5.
Goodhue, D. L. 1995. "Task-Technology Fit and Individual Performance," *MIS Quarterly* (19:2), pp. 213-236.
Gregor, S. 2006. "The Nature of Theory in Information Systems," *MIS Quarterly* (30:3), pp. 611-642.
Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337-355.
Hevner, A., March, S., Park, J., and Ram, S. 2004a. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75-105.
Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004b. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75-105.
Iacovou, C. L., Benbasat, I., and Dexter, A. S. 1995. "Electronic Data Interchange and Small Organizations: Adoption and Impact of Technology," *MIS Quarterly* (19:4), pp. 465-485.
Im, G., and Straub, D. 2012. "Building Cumulative Tradition in Organization Science: A Methodology for Utilizing External Validity for Theoretical Generalization." Georgia State University, p. 36.
Keen, P. 1980. "Mis Research: Reference Disciplines and a Cumulative Tradition," *Proc. of International Conference on Information Systems (ICIS)*, pp. 9-18.
Kettinger, W. J., Teng, J. T., and Guha, S. 1997. "Business Process Change: A Study of Methodologies, Techniques, and Tools," *MIS quarterly*), pp. 55-80.
Khaja, H. I., Abate, M., Zheng, W., Abbasi, A., and Adjeroh, D. "Evaluating Semantic Similarity for Adverse Drug Event Narratives," *In the 11th International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction (SBP)*, Washington D.C., 2018.
Kitchens, B., Dobolyi, D., Li, J., and Abbasi, A. "Advanced Customer Analytics: Strategic Value through Integration of Relationship-Oriented Big Data," *Journal of Management Information Systems*, 35(2), 2018, pp. 540-574.
Klein, H. K., and Myers, M. D. 1999. "A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems," *MIS Quarterly* (23:1), pp. 67-94
Kraemer, K. L. (ed.) 1991. *The Information Systems Research Challenge: Survey Research Methods*. Boston, MA: Harvard Business School.
Kraemer, K. L., and Dutton, W. H. 1991. "Survey Research in the Study of Management Information Systems," in *The Information Systems Research Challenge: Survey Research Methods*, K.L. Kraemer (ed.). Boston, MA: Harvard Business School, pp. 3-57.
Lafferty, J., McCallum, A., and Pereira, F. 2001. "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," in: *Proceedings of 18th International Conference on Machine Learning*. Williamstown, MA: pp. 282-289.
Larsen, K. R., and Bong, C. H. 2016. "A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses," *MIS Quarterly* (40:3), pp. 529-551; A521-A521.
Larsen, K. R., and Hovorka, D. S. 2012. "Developing Interfield Nomological Nets," in: *Hawaii International Conference on System Sciences*. Maui, Hawaii: IEEE.
Larsen, K. R. T. 2002. "A Taxonomy of Antecedents to Is Success: Variable Analysis Studies," 2002-003, University of Colorado, Information Systems, Boulder, pp. 1-95.

Lee, A. S. 1989. "A Scientific Methodology for Mis Case Studies," *MIS Quarterly* (13:1), pp. 33-50.
Levenshtein, V. I. 1966. "Binary Codes Capable of Correcting Deletions, Insertions, and Reversals," *Soviet Physics Doklady* (10:8), pp. 707-710.
Li, J., Larsen, K., and Abbasi, A. "TheoryOn: Designing a Construct-based Search Engine to Reduce Information Overload for Behavioral Science Research," *In the 11th International Conference on Design Science Research in Information Systems and Technology (DESRIST)*, St. John's, Canada, May 23-24, 2016a.
Li, J., Abbasi, A., Cheema, A., and Abraham, L. "Path to Purpose? Impact of Online Purchases' Hedonic and Utilitarian Characteristics on the Customer Journey," *In the 26th Workshop on Information Technologies and Systems (WITS)*, Dublin, Ireland, December 15-16, 2016b.
Li, J., Larsen, K., and Abbasi, A. "Unlocking Knowledge Inheritance of Behavioral Research through Ontology Learning: An Ontology-Based Search Engine," *In the 26th Workshop on Information Technologies and Systems (WITS)*, Dublin, Ireland, December 15-16, 2016c.
Li, J., Abbasi, A., Ahmad, F., and Chen, H. "PyNDA: Deep Learning for Psychometric Natural Language Processing," *In the INFORMS Workshop on Data Science*, Phoenix, AZ, November 3, 2018a.
Li, J., Abbasi, A., Ahmad, F., and Chen, H. "Deep Learning for Psychometric NLP," *In the 28th Workshop on Information Technologies and Systems (WITS)*, Santa Clara, CA, December 15-16, 2018b.
Li, J., Larsen, K., and Abbasi, A. "TheoryOn: A Design Framework and System for Unlocking Behavioral Knowledge through Ontology Learning," *MIS Quarterly*, conditionally accepted.
Marble, R. P. 2000. "Operationalising the Implementation Puzzle: An Argument for Eclecticism in Research and in Practice," *European Journal of Information Systems* (9:3), pp. 132-147.
Maynard, D., Funk, A., and Peters, W. 2009. "Using Lexico-Syntactic Ontology Design Patterns for Ontology Creation and Population," *Proceedings of the 2009 International Conference on Ontology Patterns-Volume 516*: CEUR-WS. org, pp. 39-52.
McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. "Developing and Validating Trust Measures for E-Commerce: An Integrative Typology," *Information systems research* (13:3), pp. 334-359.
Mumford, E. 2003. "Information Systems Research and the Quest for Certainty," *Journal of the Association for Information Systems* (4:1), p. 7.
Netemeyer, R., Dobolyi, D., Abbasi, A., Clifford, G., and Taylor, H. "Health Literacy, Health Numeracy and Trust in Doctor: Effects on Key Consumer Health Outcomes," *Journal of Consumer Affairs*, forthcoming.
Nickerson, R. S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of general psychology* (2:2), p. 175.
Orlikowski, W. J., and Baroudi, J. J. 1991. "Studying Information Technology in Organizations: Research Approaches and Assumptions," *Information Systems Research* (2:1), pp. 1-28.
Rabiner, L. 1989. "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," in: *Proceedings of the IEEE*. pp. 257-286.
Salton, G. 1989. "Automatic Text Processing: The Transformation, Analysis, and Retrieval Of," *Reading: Addison-Wesley*).
Simon, H. A. 1996. *The Sciences of the Artificial (3rd Ed.)*. Cambridge, MA: MIT Press.
Soper, D. S., and Turel, O. 2015. "Identifying Theories Used in North American Is Research: A Bottom-up Computational Approach," in: *48th Hawaii International Conference on System Sciences*. Kauai, HI: IEEE, pp. 4948-4958.
Spell, C. S. 2001. "Management Fashions - Where Do They Come from, and Are They Old Wine in New Bottles?," *Journal of Management Inquiry* (10:4), pp. 348-373.
Tan, S. S., Lim, T. Y., Soon, L.-K., and Tang, E. K. 2016. "Learning to Extract Domain-Specific Relations from Complex Sentences," *Expert Systems with Applications* (60), pp. 107-117.
Tsang, E. W. K., and Kwan, K.-M. 1999. "Replication and Theory Development in Organizational Science: A Critical Realist Perspective," *Academy of Management Review* (24:4), pp. 759-780.
Taylor, H. A., Henderson, F., Abbasi, A., and Clifford, G. D., "Cardiovascular Disease in African Americans: Innovative Community Engagement for Research Recruitment and Impact," *American Journal of Kidney Diseases*, 72(5)(Supp 1), 2018, S43-S46.
Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003b. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.
Weber, R. 2012. "Evaluating and Developing Theories in the Information Systems Discipline," *Journal of the Association for Information Systems* (13:1), pp. 1-30.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review." JSTOR, pp. xiii-xxiii.

White, R. 2013. "Beliefs and Biases in Web Search," *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*: ACM, pp. 3-12.

Wong, W., Liu, W., and Bennamoun, M. 2012. "Ontology Learning from Text: A Look Back and into the Future," *ACM Computing Surveys (CSUR)* (44:4), p. 20.

Zahedi, F. M., Abbasi, A., and Chen, Y. "Fake-Website Detection Tools: Identifying Design Elements that Promote Individuals' Use and Enhance their Performance," *Journal of the Association for Information Systems*, 16(6), 2015, pp. 448-484.

Zimbra, D., Abbasi, A., and Chen, H. "A Cyber-Archaeology Approach to Social Movement Research: Framework and Case Study," *Journal of Computer-Mediated Communication*, 16, 2010, pp. 48-70.

Zimbra, D., Abbasi, A., Zeng, D., and Chen, H. "The State-of-the-Art in Twitter Sentiment Analysis: A Review and Benchmark Evaluation," *ACM Transactions on Management Information Systems*, 9(2), 2018, no. 5.