Communications of the Association for Information Systems

Volume 46 Article 1

1-2020

Improving Usability of Social and Behavioral Sciences' Evidence: A Call to Action for a National Infrastructure Project for Mining Our Knowledge

Kai R. Larsen
University of Colorado, kai.larsen@colorado.edu

Eric B. Hekler
University of California

Michael J. Paul
University of Colorado

Bryan S. Gibson
University of Utah

Recommended Citation

Larsen, K., Hekler, E. B., Paul, M. J., & Gibson, B. S. (2020). Improving Usability of Social and Behavioral Sciences' Evidence: A Call to Action for a National Infrastructure Project for Mining Our Knowledge. Communications of the Association for Information Systems, 46, pp-pp. https://doi.org/10.17705/1CAIS.04601

Communications of the Association for Information Systems

Commentary DOI: 10.17705/1CAIS.04601 ISSN: 1529-3181

Improving Usability of Social and Behavioral Sciences' Evidence: A Call to Action for a National Infrastructure Project for Mining Our Knowledge

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

Eric B. Hekler
University of California

Michael J. Paul
University of Colorado at Boulder

Bryan S. Gibson
University of Utah

Abstract:

Over the last century, the social and behavioral sciences have accumulated a vast storehouse of knowledge with the potential to transform society and all its constituents. Unfortunately, this knowledge has accumulated in a form (e.g., journal papers) and scale that makes it extremely difficult to search, categorize, analyze, and integrate across studies. In this commentary based on a National Science Foundation-funded workshop, we describe the social and behavioral sciences' knowledge-management problem. We discuss the knowledge-scale problem and how we lack a common language, a common format to represent knowledge, a means to analyze and summarize in an automated way, and approaches to visualize knowledge at a large scale. We then describe that we need a collaborative research program between information systems, information science, and computer science (IICS) researchers and social and behavioral science (SBS) researchers to develop information system artifacts to address the problem that many scientific disciplines share but that the social and behavioral sciences have uniquely not addressed.

Keywords: Social and Behavioral Sciences, Ontologies, National Science Foundation, Knowledge Bases, Reproducibility, Organizing Evidence.

This manuscript underwent editorial review. It was received 11/13/2018 and was with the authors for 6 months for 2 revisions. Matt Germonprez served as Associate Editor.

1 Introduction

The social and behavioral sciences (SBS)—disciplines that study human behaviors and social processes—address critical questions such as how to use technology for useful purposes, how to prevent chronic diseases, how to reduce poverty, how to tackle climate change, how to design more livable cities, and how to help people get along better and work more productively. As such, we could safely say that this knowledge plays a foundational role for helping individuals, practitioners, and policymakers make decisions that have a real-world impact on our lives. If we could readily access this knowledge, we could use it to transform society for the better. As such, a strong societal need to advance information systems (IS) tools that support the SBS disciplines exists. However, numerous challenges make doing so a truly daunting task.

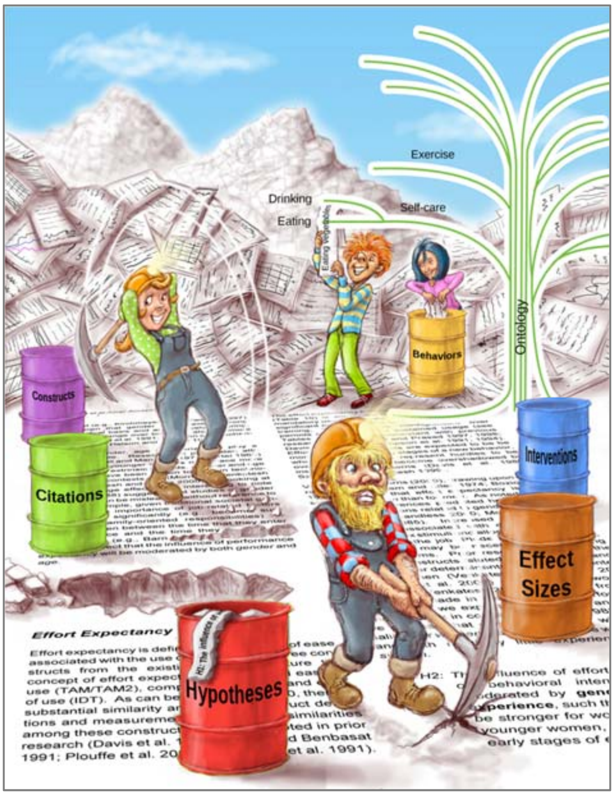

Figure 1. The Knowledge-management Problem

We describe the two primary reasons why the SBS disciplines have a knowledge-management problem. First, given that SBS research now tallies well above a million published papers[1], if we can create the appropriate information infrastructure to unlock it, we may plausibly generalize insights across individuals and contexts and, simultaneously, better understand potential gaps in the literature. We could partly do so by developing methods to extract, organize, and make readily accessible and usable the evidence and information in publications, which could enable more rapid and potentially even automated strategies for systematically reviewing and analyzing the SBS knowledge base. If done well, we could more robustly understand gaps in evidence and open questions and debates grounded in data. While researchers continue to improve methods and reproducibility of findings (Open Science Collaboration, 2015), it would be a waste to not also commit to, integrate, and better understand the knowledge in past research.

[1] We know about no work that has thoroughly evaluated all papers in the SBS disciplines. We state “above a million papers” as a highly conservative estimate given that, from searching Google Scholar for “Likert scale” and “Likert” on 3 April, 2019, we found 500,000 and 1,000,000 papers, respectively. In examining theoretical construct research in the INN database, Larsen and Bong (2019) found that 18 percent of papers contained the term “Likert scale” and 32% contained “Likert”. Given that Likert scales represent a subset of the research that examines theoretical constructs (research that itself represents a subset of the SBS disciplines), we have confidence that a million papers constitutes a highly conservative lower bound and the subset of SBS theoretical construct research alone likely constitutes around three million papers.

Creating robust systems for finding, integrating, and facilitating understanding past work represents no small task. This problem becomes only more complex when these relationships require one to weigh the relative evidentiary value for a given relationship and questions (e.g., weighting insights on a relationship more strongly if it researchers explicitly tested it in a randomized controlled trial compared to one of many correlations that they gleaned in an observational study or, similarly, upweighting research that evaluated results' predictive validity). Beyond these design requirements, one also needs to overcome great hurdles in usability. For example, it is critical that these literature reviews could be conducted in a matter of minutes over the current timescale of months to years and improve their performance metrics (Larsen & Bong, 2016; Larsen, Hovorka, West, & Dennis, 2019).

In this paper, we summarize why we need to organize the SBS literature, key complexities unique to the SBS literature, and how information systems, information science, and computer science (IICS) researchers can help create a literature review tool that helps one rapidly and iteratively interrogate the SBS literature.

1.1 The Problem of Knowledge Scope and Formatting

Given that SBS research now tallies well above a million papers, we need to acknowledge that most authors who have published an empirical paper during the last decade likely did so with knowledge about less than one percent of the past findings that pertained to and likely overlapped with their ideas[2].

Unfortunately, papers' structure, publishing's rapidly increase pace, and individuals' time and memory limitations have long ago overwhelmed researchers' capacity to absorb relevant knowledge. Academic papers follow a structure that one could argue optimally allows individuals to better read individual papers rather than integrate insights across multiple papers. Optimizing papers for human readability makes sense when a discipline contains relatively few papers and when academics largely monitor only key journals, theories, or professional societies that pertain to their work. However, optimizing for human readability has unintentionally made it harder to integrate insights across studies, which one can see in the incredible effort it takes to do robust systematic reviews that, even at the time journals publish them, have already become out of date. To collect, organize, and use knowledge to improve society, we need to optimize knowledge transfer to integrate knowledge across papers.

1.2 Lack of Meta Theory

Despite the obvious need to optimize knowledge transfer and to integrate knowledge from different papers, unique issues with the social and behavioral sciences obstruct efforts to do so. In particular, SBS researchers must start to collect their "things" and organize them in a computable format. The need for these tasks became clear from the workshop presentation by William Riley, the Director of the National Institute of Health's Office of Behavioral and Social Sciences Research (Riley, 2017). However, to complete these tasks, we need to first ask: what things do the SBS disciplines study? While many plausible answers to this question exist, SBS scholars arguably create or discover constructs that represent how and why people behave and interact with one another in contexts (Deci & Ryan, 2010). To better organize such research, we need a taxonomy of behaviors (how people behave), of proposed mediators of behaviors (why people behave in a certain way), and of how these behaviors interact with each other. The myriad theories, constructs, and hypotheses in papers encapsulate these things, and scholars study them using a wide range of research methods, such as randomized experiments, observation, and surveys. When conducting research, scholars often focus on estimating the probability that a concept meaningfully describes some aspect of how people behave and interact in a given context. These concepts, the methods scholars use to study them, and the context in which they study them arguably represent the key things that one must extract from SBS papers.

[2] A typical theory-contributing papers may contain 50 references and examine more than a dozen hypotheses that involve as many as 20-40 distinct relationships when one includes mediating and moderating relationships, each of which tens of thousands of past papers may have examined. For example, any work on the technology acceptance model (TAM) (Davis, 1989; Davis, Bagozzi, & Warshaw, 1989) has over 40,000 papers to draw on, 27 percent of which are empirical papers that contain at least one core TAM relationship (Larsen et al., 2019). As such, to draw on only one percent of the existing knowledge about even one core TAM relationship would require 100 references at minimum. Further, that figure assumes that only papers about TAM contained relevant past knowledge, which could never be true given the many highly similar and overlapping theories related to it (Venkatesh, Morris, Davis, & Davis, 2003). Similar theoretical and paper overlaps exist in all major SBS disciplines (Watts, 2017).

Theories encapsulate and represent current knowledge in terms of key constructs and the relationships between them. Scholars often test and later extend theories in thousands of follow-up studies based on the belief that they can incrementally build findings towards better understanding behavioral phenomena. However, recent findings cast shadows on this belief as even individual theories have expanded beyond comprehension (Larsen et al, 2019). Theory review studies cannot find and integrate more than a small percent of current findings. Further, dozens of theories exist that address the same phenomenon in almost identical ways. Davis, Campbell, Hildon, Hobbs, and Michie (2014) found 82 theories of behavior and behavior change that had substantial overlap and only three that were integrative. Yet, as scholars propose new and possibly redundant theories and such theories gain adherents, old theories continue to thrive. Without a behavioral knowledge-embeddedness and knowledge-integration ontology, researchers remain largely unaware of related findings not only outside their own disciplines but also in them and even in narrow research areas. This lack of awareness prevents the behavioral disciplines from becoming a front-player in new sciences, such as behavioral big data. While studies that build on big data have become more frequent (e.g., studies that examine millions of Facebook users), no single existing behavioral theory can serve an integrative role, which leads to a risk that big data research cannot build on our enormous theoretical base and will be forced to ignore it, a risk that increases by the day. Further, the SBS disciplines study inherently complex phenomena in that they often involve highly dynamic and multi-causal concepts that likely manifest idiosyncratically (Hekler et al., 2019). As such, the SBS disciplines deal with fundamental phenomena that one cannot easily model over time, study causally, and measure with standardized measures. Thus, while the SBS disciplines study complex "things", they can produce real-world insights that drive decision making. For example, Trauer, Qian, Doyle, Rajaratnam, and Cunnington (2015) show how healthcare practitioners have selected evidence-based behavioral interventions such as cognitive behavioral therapy over pharmaceuticals due to findings that showed the former more effectively treat insomnia. We will only be able to extract the signal from the noise from such a complex phenomenon if we advance the appropriate IS infrastructure.

1.3 Lack of a Common Language and Open Access

Even when focusing only on English-language research, scholars have a severely limited ability to communicate. First, human behavior is inherently messy and, therefore, difficult to categorize and quantify, which has resulted in a lack of standardized concepts across the literature. This lack of standardization creates challenges for synthesizing new findings. Second, structural issues impede progress on the first point. We believe that the SBS disciplines would benefit from adopting openness practices that have become more widespread in the computational sciences (Foster & Deardorff, 2017; Nosek et al., 2015), though they must do so carefully and in a way that considers the complexity of the phenomena that the SBS disciplines study. As of March, 2017, the Open Science Framework had accumulated 103,000 active users and 142,000 projects (COS.io, 2017), which suggests a strong movement towards open science. However, while open science makes detailed information available on research projects and often enables better replication, it does little to advance the SBS as a whole.

Without both cultural changes towards openness and standardization when possible for a given phenomenon and corresponding IS tools, the knowledge-organization process across the SBS disciplines will remain fragmented and likely driven by currently available search algorithms, particularly Google Scholar. Google Scholar has combined what seems like a vast majority of existing research, which includes SBS research, in an unprecedented way. This ability to see "all" in one search has caused scholars to trust in Google Scholar at proportions that do not match its precision and recall performance (Boeker, Vach, & Motschall, 2013; Bramer, Giustini, & Kramer, 2016; Gehanno, Rollin, & Darmoni, 2013). We know of few if any equally opaque literature systems. It has no application programming interface, and, as scholars who conduct systematic reviews have experienced, employs anti-robot algorithms so aggressive that they often mistake researchers for robots. Further, like the general Google search algorithm, Google Scholar employs network centrality measures that steer results towards older and primarily positive results at the expense of new and negative findings to improve the extent to which users believe it produces precise results. For progress, the research community must take back ownership of its research content and create transparent search engines. As we argue in this paper, search engines may well solve our problems, but, until such search engines better delineate and tag papers' content, such as specifying when the term trust refers to a construct rather than an organization or displaying an ontology of different names for the same construct to address the lack of shared language across the SBS disciplines, progress will remain elusive.

1.4 The Potential for Knowledge Discovery in Existing Reports

The information systems, information science, and computer science (IICS) disciplines could play a significant role in helping to create the knowledge infrastructure needed to address both the common and unique issues that the SBS experiences using the prior literature as a foundation. In particular, we believe the IS discipline will have a critical role in creating such an infrastructure.

While may see the existence of vast sets of overlapping and undiscovered research as a weakness, it actually constitutes a source of future strength for the SBS disciplines. Current SBS methods often rely on self-reports and human observation and interpretation, which leads to insufficiently accurate findings and high variability if looked at across the literature. This inaccuracy and high variability could prove valuable for unpacking the complex, dynamic, multi-casual and idiosyncratically manifesting phenomena that the SBS disciplines study. A central pathway forward involves building tools to help scholars examine hundreds of results on any phenomenon or relationship, which would enable them to model and visualize this wide variance when conducting iterative literature reviews. Collecting and analyzing statistics on an unbiased set of hundreds of past relationships between even two variables constitutes a major research project in itself, and our evaluations suggest that an average SBS paper tests 148 relationships and specifies more than 12 such relationships in hypotheses. Without support from the IICS disciplines, understanding what prior work has learned represents an insurmountable problem. With sufficient support from the IICS disciplines, we could feasibly translate this high degree of variance into tools that enable scholars to rapidly understand and explore how variations in constructs, construct definitions, construct operationalizations, study participants' attributes, use contexts, and study methods all co-influence to define potentially trustworthy relationships and, simultaneously, that enable scholars to identify systematic gaps in the literature from these myriad relationships.

In Table A1 in the Appendix, we share a reasonable starting point for the kinds of knowledge that exist in especially empirical SBS papers. A particularly difficult challenge for the SBS disciplines involves determining what "things" one should collect. In Table A1, we describe the types of knowledge that one typically finds in empirical SBS papers, define them, provide examples, and review state-of-the-art knowledge in ontology learning.

While Table A1 serves as a good starting point for defining the things that one would need to collect, an initial challenge involves critically examining the assumptions we make here and ensuring that one collects and organizes the right things so one can readily use them for more rapid, iterative, literature reviews—our key goal and aspiration that motivates these other efforts.

For tools to succeed in enabling scholars to properly review literature, such tools would need to tighten the iterative feedback loop in that process. For example, robust literature-review tools should likely enable scholars to rapidly explore and calibrate how they question the literature in minutes rather than the months and years the process currently takes. For example, one could imagine a person exploring issues such as "What is the strength of the relationship between A and B when tested via randomized experiments compared to observational work?" followed by a different, but complementary, question such as "How does the relationship seem to differ with one population compared to another?". If one could ask these questions in quick succession, one could start to understand prior knowledge and evidence in a more consilient way.

In Section 2, we further explore and unpack the challenges for the IICS disciplines to support the SBS disciplines and provide some initial directions on available resources.

2 The Challenge and Currently Available Resources

Other sciences have recognized that, among other tasks, they focus on collecting "things" (objects or processes of interest) and that, with a large enough collection of things, they can create organizational structures about those things, such as taxonomies or ontologies (Ashburner et al., 2000). For example, starting with Aristotle, naturalists (biologists) famously collected organisms and specimens from their work. With a large enough collection of things, Linnaeus, who reportedly personally collected 40,000 specimens, built on Aristotle's structure to create a taxonomy of living organisms that had enough flexibility to allow others to add undiscovered species to it. Over time, scholars have improved it to cover domains, kingdoms, phyla, class, order, family, genus, and, ultimately, species as a way to organize life. This collection of things and its subsequent organization has vastly expanded researchers' absorptive capacity in the otherwise overwhelmingly complex biological sciences by providing order and structure to life. From this point of view, the challenge for the SBS disciplines may seem quite a bit easier given they focus on only one species. Some existing resources that would prove helpful in moving an effort like this forward also exist: 1) controlled vocabulary and developed ontologies, 2) ontology portals, 3) data portals, and 4) translational and practice-focused work. We discuss each resource type next.

2.1 Controlled Vocabulary and Developed Ontologies

A controlled vocabulary constitutes an agreed-on set of words and phrases that scholars can use to tag information units or to develop a shared understanding about entities of interest. Ontologies include controlled vocabularies but also specify their interrelations. For example, in controlled vocabulary, an entity such as exercise may have the definition: "An activity that requires physical or mental exertion, especially when performed to develop or maintain fitness". An ontology could use that definition along with additional information. For example, the entity swimming is a type of exercise; exercise is a type of mobility; and mobility is part of activities and participation. Relevant controlled vocabularies include the psychology ontology available on the NIH Bioportal and the National Cancer Institute's Thesaurus (NCIt). Parts of the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED-CT), such as the social context section, also qualify, as do the Medical Subject Headings (MeSH) thesaurus's psychiatry and psychology sections.

2.2 Ontology Portals

These sites provide key content from theoretical papers, such as constructs, definitions, relationships, and related measures. The NCI Grid-Enabled Measures (GEM) (see https://www.gem-beta.org) portal likely represents the best-known theory portal in behavioral medicine. In line with ontological thinking, this site focuses on creating harmonized measures and constructs through community editing and voting. The site also provides links to datasets and a team collaboration area. The TheoryMaps (see https://www.theorymaps.org/) site enables one to manually draw a theory's constructs and their hypothesized relationships with visual results when searching for theories containing a given construct. The final resource, the InterNomological Network (INN) portal (see https://inn.theorizeit.org/), comes from the University of Colorado's Human Behavior Project, and provides access to constructs and measures from 10 different disciplines and fields including behavioral medicine. It provides synonymy and citation search and contains taxonomic structure for some disciplines. An evolved version of INN, TheoryOn (see http://theoryon.org), shows how one may improve and enrich search results with automatically extracted theory networks. It does so via automatically detecting hypotheses, extracting constructs and relationships, and visualizing papers' construct network.

2.3 Data Portals

The Inter-university Consortium for Political and Social Research (ICPSR) likely holds the status as the largest and most relevant data portal. ICPSR maintains over 500,000 research-relevant files such as surveys and their results. It maintains over four million variables and enables one to search for studies that contain a pair of variables. Another data portal, the MetaBUS project (http://metabus.org/), contains over one million effect sizes from studies in the management discipline. The project uses a taxonomy to tie together the variables for which scholars have collected data. The project focuses on enabling scholars to automatically conduct meta-analyses on relationships between variables.

2.4 Translational and Practice-focused Work

Translational and practice-focused work includes the psychometric instruments carefully developed for the NIH PROMIS project to cover patient-reported outcomes in the physical, mental, and social health domains. Researchers created each instrument bank based on carefully reviewing thousands of questionnaire items. The International Classification of Functioning, Disability, and Health (ICF)—another valuable resource with behavioral science implications—classifies health and health-related domains with a focus on body, individual, and societal perspectives. Finally, the Nursing Care Coordination Ontology at the NIH Bioportal addresses the coordination, people, places, and problems involved in delivering care.

3 A Framework of an Initiative to Advance Social and Behavioral Science Ontology Learning

With these resources, we propose that we should now advance SBS ontology learning and offer an initial framework on how this work could advance. We argue that future progress in the SBS disciplines depends partially on two interrelated tasks: 1) moving the SBS disciplines toward computational, integrative approaches that can capture knowledge embeddedness and illuminate existing knowledge and 2) automatically integrating new knowledge into knowledge bases. These new approaches will enable knowledge integration across disciplines and identify scientifically grounded feedback on issues such as appropriate behavior-change techniques for given individuals to produce a desired health goal. To accomplish these objectives, experts must bring together the various ways to integrate behavioral theory, develop ontology, and process natural language processing to coordinate endeavors to move behavior sciences efforts to the next level. Today, big data impacts practically every area of behavior change, and, unless we integrate findings from big data with existing behavior-change theories, both the theories and the findings will continue to overwhelm researchers and lead to fractured science.

A project that targets those two tasks should cover all the SBS disciplines and could comprise 1) careful sample and collection of a large set of manuscripts (100,000+) to represent all research in the social and behavioral sciences and 2) a virtual lab where ontology learning researchers can access the manuscripts and develop open source software to extract and place the individual knowledge types and instances in a manuscript into knowledge bases. The project should be able to execute software in a virtual lab, should store results in NoSQL databases, should accurately report accuracy measures, and should rank software on leaderboards to help scholars compare different approaches to solving the same problems. Any time new software outperforms past software on the evaluative sets, the code would be executed on the entire sample or on subsets of the sample for which the new code outperforms past code (e.g., for the psychology literature or even as focused as a specific journal in the psychology discipline).

Scholars could advance such in various ways. Indeed, in the UK, scholars have begun complementary efforts to do so. For example, scholars built the Human Behavior Change Project (Michie et al., 2017) on ideas from the IS-based Human Behavior Project (Larsen, 2010). Figure 2 outlines one possible formulation across three stages. In the first stage, the project establishes the infrastructure necessary for such projects to succeed in the future and for fostering collaboration and disseminating findings. Once the project has completed this infrastructure, the second stage begins. In this stage, the project would invite information and computer scientists work in the infrastructure to address ontology learning (e.g., most effectively extracting constructs from the sample and the targeted sources that contain such constructs). The project would also bring private and public funding agencies on board as well to support such efforts. Different teams may compete to provide the best software for a given problem. The leaderboard evaluations would automatically evaluate the software against gold standards and rank results on leaderboards. Once the second stage has proven the code can extract high-quality knowledge types, the third stage begins. In the third stage, the project develops software that would enable manuscript owners and publishers to extract knowledge locally and develop their own products. Alternatively, it would enable them to collaborate with a virtual lab team to create software that would allow them to pool their knowledge types into large knowledge bases. These large knowledge bases would then support existing business models and enable new ones while also making the knowledge bases available to researchers through the kinds of contractual arrangements already in place between content owners and academic libraries.

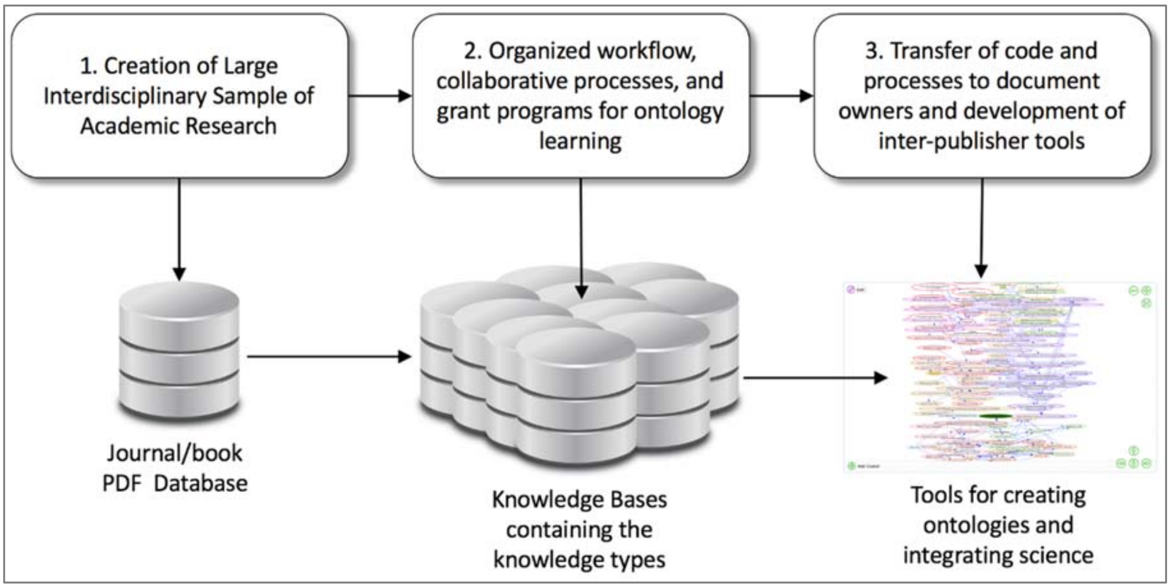

Figure 2. Project Overview

3.1 Tasks in Each Stage

We discuss the tasks in each stage in more detail in Sections 3.1.1 to 3.1.3.

3.1.1 Stage 1: Determine Inclusion Criteria

This stage focuses on collecting machine-readable copies of papers that meet inclusion criteria. As such, the project should set up consistent workflow apps and APIs to collect papers, which will eventually transition into a flow of new papers as they become available in the source databases. To achieve such a workflow, the project first needs to create a complete (or as complete as possible) knowledge graph about SBS sources. Here, the project defines sources to include gray literature, conference publications, books, and journal papers regardless of their journals' citation scores. Alternatively, this work could start with only journal papers to shortcut the process under the assumption that highest-quality knowledge exists in them. To integrate existing ontology learning resources, the project should make sure it includes the ~20,000 papers that the Human Behavior Project, the metaBUS project, and any other large projects cover. As such, the project needs to work with content owners, librarians, and lawyers to develop the legal framework under which this project may exist.

3.1.2 Stage 2: Collect Papers and Set Up Infrastructure

In the second stage, the project should store all papers that it gathers in a database (e.g., NoSQL database). The project should also set up related big data-capable databases for each defined knowledge types in Table A1. The project needs an infrastructure that allows individuals to share code designed to extract knowledge types and the ability to run such code for all or subsets of journals with leaderboards that indicate which code units perform best for specific knowledge types. For example, a team may possibly have created a system that extracts hypotheses from psychology journals better than any other code but that this overall improvement in performance via supervised machine learning decreases the performance on disciplines such as nursing, which adheres to a non-standard hypothesis format. The team may then scale back the new software to focus only on journals from the disciplines it works best for, and the team may adjust the previously best software to address only the non-standard disciplines, such as nursing. Over time, we would expect to see solutions that rely on existing knowledge bases for other types to improve ontology learning. For example, one can find the constructs/variables in a paper in several different ways: 1) one can extract them from hypotheses that other teams extracted, 2) one may extract them from correlation tables, 3) one may extract them from construct definitions sections, and 4) one may infer them based on the citation structure to existing theories. As the accuracy of ontology learning for any one of these approaches increases, others that use the findings from those algorithms will also likely improve through a virtuous cycle.

One promising direction involves further improving the virtuous cycle. In citizen science, an area of work that IS researchers have specified (Levy & Germonprez, 2017; Lukyanenko, Parsons, Wiersma, & Maddah, 2019a), researchers split scientific asks into smaller pieces and parcel it out to citizens (Lukyanenko & Parsons, 2019). For example, one could divide the process to evaluate how well algorithms extract knowledge types into small tasks and shared (via a specially designed online platform) across a network of citizen scientists. While the task involves many tedious and difficult tasks, their nature allows non-experts to evaluate them. Citizen science has begun to accumulate an impressive track record of tackling complex and esoteric problems, such as classifying galaxies, folding proteins (see Clery, 2011; Khatib et al., 2011; Lintott et al., 2008), and designing platforms and tasks in innovative ways. For a review from an IS perspective, see Lukyanenko, Wiggins, and Rosser (2019b).

In our domain, for example, a scholar may display a table to an individual and ask that person to specify whether it contains statistical results or not. If it does, the scholar could give another citizen that table and ask the citizen to identify instances of variable names. The scholar may then give a third person the marked-up table and ask the person to draw a grid around the statistical results, which would provide algorithms better training material for extracting correlation matrices and other types of statistics.

Finally, due to the papers' value and journal owners' requirements, the project would need to set up a process for approving research teams. Such a process may also require a shared virtual private network (VPN) and a security regime to control access to workflow systems and databases.

3.1.3 Stage 3: Tools for Ontologies

In this stage, the project develops processes to help publishers and journal owners use the results of the ontology learning in their databases and products. To enable them to do so, the project needs to develop other processes to help publishers to use the code developed in the project, such as with large publishers with access to teams of programmers. To ensure the program succeeds, research teams may want to work with journal owners on creating cutting-edge tools to search and integrate science. Doing so would generate worthy research projects that would likely make a real difference for the social and behavioral sciences in and of themselves on top of any benefits derived from publishers improving their tools. One potential success would involve small, medium, and even some large publishers deciding to create combined search engines built on top of the ontology knowledge developed for the project. Accordingly, they could provide end users access to products created for micro-information and automatic payment to the relevant papers owners.

Through the proposal in this paper, we attempt to make a difficult process possible and increasingly understandable. We see IS researchers as the natural leaders for such efforts due to the discipline's focus on the intersection of technology adoption, systems development, design science, and broad social and behavioral research.

4 Conclusion

An effort such as the project we justify in this paper could transform society's ability to build on prior knowledge when making decisions that the SBS disciplines could impact by tackling the absorptive capacity problem—by making it possible for one person to stay up to date on all work relevant to their interests. For example, it would enable scholars to more easily conduct (possibly in a fully automated manner) literature reviews, meta-analyses, and syntheses across studies and scientific domains to advance our understanding about complex systems in the social and behavioral sciences. Experiences from smaller projects in this domain, such as metaBUS (Bosco, Steel, Oswald, Uggerslev, & Field, 2015) and the Human Behavior Project (Larsen, 2010; Larsen & Bong, 2016), have shown that collecting SBS "things" allows new modes of analysis and enables scholars to answer questions we do not currently even know to ask. We believe we should now advance an SBS national infrastructure project for improving the usability of SBS evidence, and we believe the IS discipline must act as an integrator and collaborator in such work.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., & Eppig, J. T. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, 25(1), 25-29.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

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), 1-12.

Bosco, F. A., Steel, P., Oswald, F. L., Uggerslev, K., & Field, J. G. (2015). Cloud-based meta-analysis to bridge science and practice: Welcome to metaBUS. *Personnel Assessment and Decisions*, *1*, 3-17.

Bramer, W. M., Giustini, D., & Kramer, B. M. R. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, Medline, and Google Scholar: A prospective study. *Systematic Reviews*, *5*.

Cane, J., O'Connor, D., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science*, *7*(1), 37-53.

Clery, D. (2011). Galaxy zoo volunteers share pain and glory of research. *Science*, *333*(6039), 173-175.

COS.io. (2017*). Vital stats*. Retrieved from https://cos.io/about/brief-history-cos-2013-2017/

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982-1003.

Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2014). Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review. *Health Psychology Review*, *9*, 323-344.

Deci, E. L., & Ryan, R. M. (2010). *Self-determination.* New York, NY: John Wiley & Sons.

Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association*, *105*(2), 203-206.

Gefen, D., & Larsen, K. R. (2017). Controlling for lexical closeness in survey research: A demonstration on the technology acceptance model. *Journal of the Association for Information Systems*, *18*(10), 727-757.

Gehanno, J.-F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, *13*(7).

Hekler, E. B., Hekler, K., P., Chevance, G., Golaszewski, N. M., Lewis, D., & Sim, I. (2019). Why we need a small data paradigm. BMC Medicine, *17*.

Hovorka, D. S., Larsen, K. R., Birt, J., & Finnie, G. (2013). A meta-theoretic approach to theory integration in information systems. In *Proceedings of the Hawaii International Conference on System Sciences.*

Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., & Popović, Z. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology*, *18*(10), 1175-1177.

Larsen, K. R. T. (2001). Antecedents of implementation success: A comprehensive framework. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*.

Larsen, K. R. (2010). *Human behavior project*. Retrieved from www.theorizeit.org.

Larsen, K. R., & Bong, C. H. (2016). A tool for addressing construct identity in literature reviews and meta-analyses. *MIS Quarterly*, *40*(3), 529-551.

Larsen, K. R., & Bong, C. H. (2019). *Inter-nomological network search engine.* Retrieved from http://inn.theorizeit.org

Larsen, K. R., & Hovorka, D. S. (2012). Developing interfield nomological nets. In *Proceedings of the Hawaii International Conference on System Sciences.*

Larsen, K. R., Hovorka, D. S., West, J. D., & Dennis, A. R. (2019). Understanding the elephant: A discourse approach to corpus identification for theory review articles. *Journal of the Association for Information Systems, 20*(7), *887-927.*

Larsen, K. R., Nevo, D., & Rich, E. (2008). Exploring the semantic validity of questionnaire scales. In *Proceedings of the Hawaii International Conference on System Sciences.*

Larsen, K. R., Voronovich, Z. A., Cook, P. F., & Pedro, L. W. (2013). Addicted to constructs: Science in reverse? *Addiction*, *108*(9), 1532-1533.

Lasser, K., Boyd, J. W., Woolhandler, S., Himmelstein, D. U., McCormick, D., & Bor, D. H. (2000). Smoking and mental illness: A population-based prevalence study. *JAMA*, *284*(20), 2606-2610.

Li, J., & Larsen, K. R. (2011). Establishing nomological networks for behavioral science: A natural language processing based approach. *In Proceedings of the International Conference on Information Systems.*

Li, J., & Larsen, K. R. (2013). Tracking behavioral construct use through citations: A relation extraction approach. In *Proceedings of the International Conference on Information Systems.*

Li, J., Larsen, K. R., & Abbasi, A. (Forthcoming). Theoryon: Designing a construct-based search engine to reduce information overload for behavioral science research. *MIS Quarterly*.

Levy, M., & Germonprez, M. (2017). The potential for citizen science in information systems research. Communications of the Association for Information Systems, *40*, 22-39.

Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., & Andreescu, D. (2008). Galaxy zoo: Morphologies derived from visual inspection of galaxies from the Sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, *389*(3), 1179-1189.

Lukyanenko, R., Larsen, K. R., Parsons, J., & Gefen, D. (2019). Toward creating a general ontology for research validity. In *Proceedings of the 38th International Conference on Conceptual Modeling*.

Lukyanenko, R., & Parsons, J. (2019). Beyond micro-tasks: Research opportunities in observational crowdsourcing. In M. Khosrow-Pour (Ed.), *Crowdsourcing: Concepts, methodologies, tools, and applications* (pp. 1510-1535). Hershey, PA: IGI Global.

Lukyanenko, R., Parsons, J., Wiersma, Y. F., & Maddah, M. (2019). Expecting the unexpected: Effects of data collection design choices on the quality of crowdsourced user-generated content. *MIS Quarterly*, *43*(2), 623-647.

Lukyanenko, R., Wiggins, A., & Rosser, H. K. (2019). Citizen science: An information quality research frontier. *Information Systems Frontiers*, 1-23.

Michie, S., Thomas, J., Johnston, M., Aonghusa, P. M., Shawe-Taylor, J., Kelly, M. P., Deleris, L. A., Finnerty, A. N., Marques, M. M., & Norris, E., O'Mara-Eves A., & West, R. (2017). The human behaviour-change project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implementation Science*, *12*(1).

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (V1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81-95.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., & Christensen, G. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422-1425.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Riley, W. T. (2017). Keynote on establishing the fundamental scientific challenges that must be resolved to advance behavioral science research. In *Proceedings of the NSF International Workshop on Behavioral Ontology Learning.*

Rhodes, R., & Nigg, C. R. (2011). Advancing physical activity theory: A review and future directions. *Exercise and Sport Sciences Reviews*, *39*(3), 113-119.

Soper, D. S., Turel, O., & Geri, N. (2014). The intellectual core of the IS field: A systematic exploration of theories in our top journals. In *Proceedings of the 47th Hawaii International Conference on System Sciences.*

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

Trauer, J. M., Qian, M. Y., Doyle, J. S., Rajaratnam, S. M., & Cunnington, D. (2015). Cognitive behavioral therapy for chronic insomnia: A systematic review and meta-analysis. *Annals of Internal Medicine*, *163*(3), 191-204.

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.

Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour*, *1*, 1-5.

Appendix

Table A1. Details on Knowledge Types in Behavioral Theories and Past Work

Knowledge type (KT)	Definition	Examples (emphasis added)	Past work
1. Theory	Ontologically, it comprises constructs (knowledge type; KT 4), their associations (KT 6, 10), and the states it covers. A paper or papers generally specify and name a theory. They often represent the theory through diagrams or hypotheses (KT 6) depicting constructs (KT 4), construct relationships (KT 10), and research context.	"Next, a unified model, called the **Unified Theory of Acceptance and Use of Technology** (**UTAUT**), was formulated" (Venkatesh et al., 2003, p. 425).	Soper, Turel, and Geri, (2014) used n-gram (1-5) analysis to detect theory names from an existing list of 87 theories. Later, Soper and Turmel (2015) improved the methodology through detection of instances of "theory" and "model". The authors applied the n-gram analysis to Wikipedia titles under the assumption that "nearly any theory of even moderate renown would be likely to have an associated article in the English language Wikipedia" (p. 4950). Davis et al. (2014) conducted A scoping review study and found 82 theories of behavior and behavior change.
2. Theory instance	Successful theories (knowledge type 1) are extended in follow-up papers that cite (knowledge type 11) the original papers or papers. In a paper, it is represented as name of theory extended in the focal article, diagrams, or hypotheses (KT 6) that depict constructs (KT 4), construct relationships (KT 10), and research context (KT 3).	"While the **technology acceptance model** (**TAM**), introduced in 1986, continues to be the most widely applied theoretical model in the IS field, few previous efforts examined its accomplishments and limitations. **This study traces TAM's history**, investigates its findings, and cautiously predicts its future trajectory.	Larsen et al. (2019) proposed an approach for tracking and detecting articles containing theoretical instances.
3. Research context	Study's boundary: industry, functional area, respondent backgrounds, sample size, survey focus.	Industry: "entertainment". Functional area: "Product development". Respondent backgrounds: not reported. Sample size: "54". Survey focus: "Online meeting manager that could be used to conduct Web-enabled video or audio conferences in lieu of face-to-face or traditional phone conferences" (Venkatesh et al., 2003, p. 438).	

Table A1. Details on Knowledge Types in Behavioral Theories and Past Work

4. Constructs	Construct name, synonyms, and definition. This knowledge type is hard to track and specify (Larsen, Voronovich, Cook, & Pedro, 2013).	"Performance expectancy is defined as **the degree to which an individual believes that using the system will help him or her to attain gains in job performance**" (Venkatesh et al., 2003, p. 447). "**Performance Expectancy** (**PE**)" (Venkatesh et al., 2003, p. 459).	Larsen and Bong (2016) examined construct synonymy by examining item similarity using NLP algorithms. The MetaBUS team manually devised a construct taxonomy for the organizational behavior discipline (Bosco et al., 2015). The Human Behavior Project manually created a construct taxonomy for the MIS discipline (Larsen & Bong, 2016). Cane, O'Connor, and Michie (2012) proposed the theoretical domains framework as a framework for the behavioral medicine discipline.
5. Items	Construct measurement items (items).	"I would find the system useful in my job. Using the system enables me to accomplish tasks more quickly. Using the system increases my productivity. If I use the system, I will increase my chances of getting a raise" (Venkatesh et al., 2003, p. 460).	Larsen and Bong (2016) applied NLP similarity algorithms to detect the similarity among items. They applied latent semantic analysis, latent Dirichlet allocation, wordnet-based approaches and proposed a combination: the construct identity detector (CID_1) algorithm.
6. Hypotheses	The supposition or proposed explanation for the relationship between constructs (KT 10).	"H1: The influence of performance expectancy on behavioral intention will be moderated by gender and age, such that the effect will be stronger for men and particularly for younger men" (Venkatesh et al., 2003, p. 450).	Li and Larsen (2011) and Li, Larsen, and Abbasi (Forthcoming) developed algorithms to automatically extract hypotheses from papers. Early works include Larsen (2001).
7. Construct relationships	The individual construct relationships at a pairwise level. Some relationships are between constructs and existing relationships as is the case for moderating relationships. One should break mediating relationships should up into two pairwise relationships, enabling a recreation of the original model.	Relationship 1: performance expectancy (influence on) behavioral intention. Gender (0 = female / 1 = male) (positive moderation of) relationship 1. Age (negative moderation of) relationship 1.	Li and Larsen (2011) and Li et al. (Forthcoming) developed algorithms to automatically extract relationships between constructs as they appeared inside hypotheses, whereas Gefen and Larsen (2017) automated construct relationship evaluation. Early work in IS includes Larsen and Hovorka (2012) and Hovorka, Larsen, Birt, & Finnie (2013).

Table A1. Details on Knowledge Types in Behavioral Theories and Past Work

8. Methods	The type of approach used to analyze the data.	"**Partial least squares** (PLS Graph, Version 2.91.03.04) was used to examine the reliability and validity of the measures" (Venkatesh et al., 2003, p. 439). "**PLS** was used to test all eight models at the three points of measurement in each of the two data sets. In all cases, we employed a **bootstrapping method** (500 times) that used randomly selected subsamples to test the PLS model" (p. 439).	
9. Descriptive statistics	Quantitative descriptions	"The perceptions of **voluntariness** were very high in studies la and 1b (la: **M = 6.50, SD = 0.22**; 1b: **M = 6.51. SD = 0.20**) and very low in studies 2a and 2b (la: **M = 1.50, SD = 0.19**; 1b: **M = 1.49. SD = 0.18**)" (Venkatesh et al., 2003, p. 439).	
10. Relationship statistics	The statistical findings about relationships between the constructs (KT 4).	"Off-diagonal elements are correlations between constructs" PE (performance efficacy) – BI [behavioral intention]: T1 Results" **.38***** T2 Results" **.41***** T3 Results" **.44***** T1-T3 refers to three time periods evaluated" (Venkatesh et al., 2003, p. 458)	The MetaBUS project employs light-weight NLP approaches to extract correlation matrices from academic papers and has developed a large-scale automatic meta-analysis approach on the resulting one million effect sizes (Bosco et al., 2015).
11. Citations	References to past work that informed the focal paper. Citation analysis and access to citations is common but can only connect one paper to another, a level of analysis that lacks granularity for most ontology learning applications. The real benefits from detecting citations and their location in a focal paper concern being able to connect them to key content in the paper. For example, connecting a given citation to a specified theory.	"One of the most powerful theories of human behavior is social cognitive theory (see Bandura, 1986)"	Li and Larsen (2013) proposed an "automatic construct-level citation extraction system to refine citations from the paper level to the construct level". They extracted citation mentions with an F_1 measure of .92[3].

[3] One calculates the F1 measure as the harmonic mean of precision (true positives / true positives + false positives) and recall (true positives / true positives + false negatives).

Table A1. Details on Knowledge Types in Behavioral Theories and Past Work

12. Quality evaluations (validities)	Evidence of the extent to which the authors incorporated various standard quality evaluations.	"Thus, for PA, a recent development has been the suggestion of two reconceptualizations of the process structure specific to *PA*, both of which exhibit **factorial validity** across *gender*, *age*, and *ethnicity* (33)" (Rhodes & Nigg, 2011, p. 115) "The **discriminant validity** of *self-identity*, compared to the other components of the TPB, was also supported by the findings (Table 2)" (Rise et al., 2010, p. 1095) "The present study examined 40 tests of the **predictive validity** of *self-identity* using meta-analytic procedures and, therefore, constitutes the most comprehensive and systematic analysis of the *self-identity/intention* relation to date" (Rise et al., 2010, p. 1097).	Larsen, Nevo, and Rich (2008) and Lukyyanenko, Larsen, Parsons, and Gefen (2019) developed new thinking about what validities are and could become.
13. Behavior Change Techniques (BCTs)	BCTs are theory-based approaches for changing behaviors' determinants.	"The most commonly observed techniques were as follows: providing instruction on how to perform behavior, modeling/demonstrating the behavior, providing feedback on performance, goal-setting for behavior, planning social support/change, information about others' approval, and goal-setting for outcome" (Conroy et al., 2014, p. 650).	BCTT taxonomy (Michie et al., 2013) provides a hierarchy of behavior change techniques. The BCTT interventions database contains 300+ papers coded using BCTT (see http://www.bct-taxonomy.com/interventions).
14. Behaviors	Behaviors are sometimes considered a subclass of constructs. They are actions that human beings can observe and have consequences for that person's life or life quality. These are often the dependent variable in a theory.	Example of two behaviors: "We used population-based data from the National Comorbidity Survey (NCS) to examine the association between type and severity of mental illness and the likelihood of **smoking** and subsequent **cessation**" (Lasser et al., 2000, p. 2606).	

About the Authors

Kai R. Larsen is an Associate Professor of Information Systems in the division of Organizational Leadership and Information Analytics, Leeds School of Business, University of Colorado Boulder. He is a courtesy faculty member in the Department of Information Science of the College of Media, Communication and Information, a Research Advisor to Gallup, and a Fellow of the Institute of Behavioral Science. He is most known for providing a practical solution to Edward Thorndike's (1904) Jingle Fallacy and for his contributions to the semantic theory of survey response (STSR), which holds that results of surveys using attitude scales primarily measure the linguistic relationships between survey questions.

Eric Hekler is the Director of the Center for Wireless & Population Health Systems within the Qualcomm Institute at the University of California, San Diego (UCSD), an Associate Professor in the Department of Family Medicine & Public Health at UCSD, and a faculty member of the Design Lab at UCSD. He is a transdisciplinary researcher, educator, and practitioner as his work is at the intersection of clinical health psychology (original training), human-centered design, public health (my primary affiliation), and control systems engineering. His research is broadly focused on advancing methods in the design, creation, optimization, evaluation, and reuse of digital health technologies. He has over 120 publications that span the many disciplines he contributes and has an active federal and foundation funding. He is recognized internationally as an expert in the area of digital health.

Michael Paul is an Assistant Professor in the Department of Information Science at the University of Colorado Boulder. He received a PhD in Computer Science from Johns Hopkins University in 2015 and a BS in Computer Science from the University of Illinois at Urbana-Champaign in 2009. His research areas include machine learning and natural language processing with applications to health sciences.

Bryan Gibson, DPT, PhD works in the Department of Biomedical Informatics at the University of Utah. His research interests center around the translation of evidence from the psychological sciences into the consumer health Informatics. In particular, he is interested in the information infrastructure needed for the design and evaluation of technology-based behavioral interventions, with a focus on interventions intended to help individuals prevent or manage chronic disease.