Title
An Argument for Knowledge Variety in Evidence-Based Management

Permalink
https://escholarship.org/uc/item/0514w6wx

Author
Martelli, Peter Fabrizio

Publication Date
2012

Peer reviewed|Thesis/dissertation
An Argument for Knowledge Variety in Evidence-Based Management

By

Peter Fabrizio Martelli

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

in

Health Services and Policy Analysis

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Stephen M. Shortell, Chair
Professor Thomas G. Rundall
Professor Karlene H. Roberts

Fall 2012
Dedication and Acknowledgments

This essay marks the fulfillment of a life of formal education at several schools and in several disciplines. It might not be the most cogent or noteworthy essay I have written, but I am content that I arrived to the contents through an authentic process of research and reflection, and I hope that the insights will help others see the world in a new and interesting way. For that education and desire for knowledge, I sincerely thank my Mother, my Father, my family, and my teachers and professors at PEARLS, the Hackley School, the University of Pennsylvania, Thomas Jefferson University, and the University of California – Berkeley. For the endless support and inspiration, I sincerely thank my wife, Sarah, and her family, and my many dear friends, particularly Geoff Mitelman, Alex Glassman, Elias Greendorfer, Nate Bush and Gordon Shen. For the many years of support and intellectual stimulation, I thank all the members of the UC-Berkeley Center for Catastrophic Risk Management, especially Karlene Roberts and Tony Hare.

During my time as a doctoral student, I met several professors, mentors, and colleagues whose wisdom and friendship helped me grow academically and personally: Among them are Todd LaPorte, Chris Ansell, Jim Lincoln, Ray Catalano, Paul Schulman, Emery Roe, Ian Mitroff, Bob Bea, Tom McKone, Renaud Vidal, Terry Hill, Jody Hoffer Gittell, Sara Singer, Ranga Ramanujam, Tim Vogus, Adam Galinsky, and Brian Uzzi. From my program, I would like to particularly thank Sara “SFT” Taylor Rohe and Karl Rohe, Jenny Liu, Alexis Pozen, and S. Omowale Fowles for their infinite support; Dion Shimatsu-Ong for making the program work; and my committee members, Steve Shortell, Karlene Roberts, and Tom Rundall, whose perspectives shaped me inestimably. To everyone that I have not included here, including my friends and classmates from Berkeley, Oakland, the US Forest Service, and elsewhere, thank you for your support and kindness.

I would also like to thank and note the funders that made it possible to pursue a doctoral degree, especially:
- The Russell M. Grossman Endowment Dissertation Award
- The Gordon and Betty Moore Foundation
- National Science Foundation (NSF) Grant # CMMI 0624296
- Center for Catastrophic Risk Management Research Fellowship
- Support from the France-Berkeley Fund
- Support from the Agency for Healthcare Research and Quality (AHRQ)
- An award from the Center of Evaluation for Global Action (CEGA)
- Block grants from the Health Services and Policy Analysis program
# Table of Contents

§0.0: Introduction .................................................................................. 1  
  §0.1: Motivation .................................................................................. 1  
§1.0: Research-Practice Gap .................................................................. 2  
  §1.1: Translation of Research into Practice .......................................... 3  
  §1.2: Evidence-based Management ...................................................... 5  
  §1.3: Challenges for Evidence-based Management ............................... 6  
§2.0: Rewiring Evidence-based Management .......................................... 8  
  §2.1: Managing Uncertainty ................................................................. 9  
  §2.2: Uncertainty, Complexity and Diversity ........................................ 10  
  §2.3: Diversity, Requisite Variety and Knowledge ............................... 13  
  §2.4: Knowledge Variety .................................................................... 15  
    §2.4.1: Making Inferences from the Environment is a Social Process ... 16  
    §2.4.2: Knowledge, not Information, is the Object of Regulation .......... 24  
    §2.4.3: Knowledge Diversity and Knowledge Variety ......................... 29  
  §2.5: Evidence-based Management in the Context of Knowledge Variety .. 33  
    §2.5.1: What is the Best Available Evidence? ..................................... 34  
    §2.5.2: Organizational Attention and Sensemaking ............................ 39  
    §2.5.3: Best Practices and Overadoption ......................................... 42  
  §2.6: Concluding Thoughts on Theory ................................................ 46  
§3.0: Investigating Knowledge Variety in Practice .................................. 48  
  §3.1: Healthcare as a Setting ............................................................... 49  
  §3.2: Scope of the Project .................................................................. 50  
  §3.3: MATCH Project Description and Findings ................................... 51  
    §3.3.1: MATCH Toolbox Dissemination ........................................... 52  
    §3.3.2: MATCH as Motivation and Comparison ............................... 54  
  §3.4: BEACON-Berkeley Project ......................................................... 55  
§4.0: Methods and Findings .................................................................. 56  
  §4.0.1: BEACON-Berkeley Project Funding ....................................... 57  
  §4.0.2: Institutional Review Board ....................................................... 57  
  §4.1: Empirical Setting: Details, justification and limitations ............... 58  
  §4.2: Operationalization of Variables ................................................ 62  
  §4.3: Survey Overview: Details, justification and limitations ............... 66  
    §4.3.1: Survey Sampling .................................................................. 67  
    §4.3.2: Survey Procedures ................................................................ 69  
    §4.3.3: Survey Responses ............................................................... 71  
    §4.3.4: Survey Results ................................................................... 74  
  §4.4: Interview Site Selection: Details, justification and limitations ....... 93  
    §4.4.1: Interview Sampling ............................................................. 97  
    §4.4.2: Interview Results ............................................................... 99  
  §4.5: Concluding Thoughts on Practice ............................................. 111  
§5.0: Discussion and Conclusion ............................................................ 117  
REFERENCES ....................................................................................... 121  
APPENDIX ............................................................................................ 139
List of Figures:

1) Risk of Death from Shark Attack.................................................................4
2) Ashby’s Model............................................................................................13
3) Theory, Theorizer, Aspect ........................................................................22
4) Continuum of Voice I ..............................................................................36
5) Continuum of Voice II ..............................................................................37
6) Proposition................................................................................................48
7) Informed Decisions Toolbox Screenshot ..................................................52
8) Use of Informed Decisions Toolbox Website ...........................................54
9) Simplified Model of Evidence-based Management Knowledge Transfer ....59
10) BEACON Collaborative Member Hospital Site Locations in 2011 .............61
11) Survey Demographic Statistics.................................................................75
12) Survey Descriptive Statistics.................................................................77
13) Time/Money Currently Spent by Organization to Collect Evidence ...........78
14) Sources Trusted and Frequently Sought ..................................................79
15) Total Mentions of Evidence by Category ...............................................79
16) Knowledge and Rogers’ Diffusion Characteristics......................................81
17) Knowledge Preferred Under Time/Money Constraints .............................82
18) Knowledge Box Plots...............................................................................85
19) Organizations by Respondents...............................................................86
20) Histograms of Normalized HHI ...............................................................87
21) Relationship of Knowledge Variety and Years Since Most Recent Degree 87
22) BEACON Meeting Attendance...............................................................95
23) California Dept. of Public Health Complaints/Incidents/Enforcement Actions 97
24) Comparison of MATCH and BEACON Evidence Sources .....................115

List of Tables

1) Gasson Knowledge Mapping ....................................................................27
2) NORFACE Knowledge Typology..............................................................28
3) Knowledge Typology Illustration..............................................................29
4) Established Processes for Evidence-Based Management ...........................33
5) Common Definitions for Evidence-Based Management ...........................34
6) Sample of MATCH Comments ..................................................................52
7) MATCH Problematic Comments..............................................................55
8) Suggested Instances of Evidence .............................................................63
9) Complements and Substitutes for BEACON ...........................................80
10) Knowledge Types by Source Category..................................................83
11) Definition of Evidence-Based Management ..........................................92
12) Definition of Management Best Practices (Survey Data) .........................93
13) BEACON Meeting Attendance................................................................95
14) Executive Participation in BEACON Site Visits .......................................96
15) Hospital Site Mission and Vision Statements .........................................97
16) Interviewees by Organizational Title ......................................................100
17) Examples of Strategic Priorities...............................................................101
18) Definition of Management Best Practices (Interview Data) .....................109
Abstract

An Argument for Knowledge Variety in Evidence-Based Management

by

Peter Fabrizio Martelli

Doctor of Philosophy in Health Services and Policy Analysis

University of California, Berkeley

Professor Stephen M. Shortell, Chair

Evidence-based management has been proposed as a method to impose rationality on decision-making, and thereby link research more closely with practice. Yet, knowledge, not evidence per se, is the means by which organizations map the uncertainty of their environments and develop strategies accordingly. This essay develops a theoretical argument to support the construct of knowledge variety, and argues that knowledge variety is a preferable operationalization of evidence-based management for organizations facing the epistemic uncertainty of complex environments. The first section sets a foundation drawing upon diversity in socio-behavioral regulation, requisite variety, the social structure of knowledge, epistemic knowledge categories, rhetorical persuasion, sensemaking, and organizational attention. The second section describes a mixed-methods project to examine evidence-based management in terms of knowledge variety across 42 hospitals associated with a common knowledge intermediary. An original proposition intending to establish a relationship between the maintenance of variety and an organization’s ability to customize diffusing best practices was unable to confirm or disconfirm a relationship. Nevertheless, a series of qualitative and quantitative findings on evidence-based management and best practices are presented, including a relationship between knowledge variety and the number of years since last academic degree, suggesting the importance of revisiting the Aristotelian notion of practical wisdom (or, phronesis). Reflections, limitations, and suggestions for future research are described throughout.
§0.0: Introduction

What is *evidence-based management* and what role should it play in strategic management? In posing these questions, this essay intends to provide a general overview of the topic, how it is and is not being used by top management, and some creative perspectives on the topic drawn from adjacent literatures.

The overarching goal of the inquiry is to find a common ground linking the process of evidence-based management to the study of behavioral risk and reliability. It will be argued that the formation of mental models and the sociological process of sensemaking are necessarily driven by the characteristics of the *sense* itself — and that the characteristics of sense are promoted or inhibited by the characteristics of evidence. In order to apply evidence-based management in a strategic context considering safety and reliability, this inquiry suggests that organizations should deliberately structure their decision-making around types of knowledge.

In the following sections, I will describe some of the theoretical, practical and methodological issues posed by evidence-based management in this context. The earlier sections will address the intent of evidence-based management, challenge the current definitions, and suggest new perspectives and propositions on the topic. The later sections will describe an investigation into the perception and use of evidence-based management and best practices by high-level decision-makers, followed by findings, limitations and reflections on the study.

§0.1: Motivation

In a collection of short essays entitled “How to Travel with a Salmon,” Umberto Eco, the Italian author and semiotician, illustrates a theoretical situation encapsulated by the title of its essay: “On the Impossibility of Drawing a Map of the Empire on a Scale of 1 to 1.” In this whimsical piece, Eco highlights the challenges of specifying a link between theory and reality so completely that a human could perfectly determine phenomena (i.e., where he or she was in the Empire) given the relevant theoretical predictions (i.e., the point on the map) without any additional special human inference.

What lessons can be drawn from this essay? I have often thought of Eco’s commentary on the special links between reference and the referrer when considering the difficulties in the translation of research into practice. Generally speaking, the translation of research into practice is a concept related to evidence-
based management focusing on the barriers and facilitators to the use of research findings in practical applications. An often-cited, and never updated, finding is that “it takes 17 years to turn 14 per cent of original research to the benefit of patient care,” with the majority of time absorbed in bibliographic indexing, summarizing in reviews and guidelines, and in attempts at implementation (Lenfant 2003) – in other words, a clinically applicable finding is applied poorly and narrowly.

Among the many interesting issues to consider is the restriction that this impossibility imposes on our use of maps, whether literal or theoretical. If a 1:1 map is impossible, then we must suffice with some lack of fidelity in the mapping – or in terms of translation, some deficiency between what we know and how we begin to apply it. This deficiency can take multiple forms, from a paucity of truly-corresponding information to a shortfall of our concepts and theories of how facts are linked to actions. Lacking information drives decision-makers into tradeoffs between achieving completeness and acting on available information. The question of how decision-makers attempt to achieve completeness in their information despite individual and organizational barriers is one strand of translational research addressed later in this essay. Likewise, lacking concepts and theories about the mapping drives decision-makers to make tradeoffs on whether and how to apply the information assembled by actors within the system. Even given modest attempts to overcome this deficiency, a trap of “increasing relative ignorance” (LaPorte 1994) with respect to the growing complexity of the world being mapped frustrates optimal decisions. This shortfall also raises questions on the extent to which decision-makers are aware of the impossibility of finding a perfect map, and what they believe about how information should be collected and used under this constraint.

If the fidelity of our “scaled” maps is inherently compromised, then what would it mean if decision-makers behaved as if they weren’t? What would result from a prescriptive theory of translation that implicitly implied a perfect, fixed mapping? What would this mean in the context of complex, cooperative endeavors in an uncertain environment? Is there an alternative to translational approaches that imply that we could know the exact thing to do, if not for some personal and social barriers? These, and other concerns, serve as background motivation for addressing this topic.

§1.0: Research-Practice Gap

There is a well-documented gap between research findings and their application in social and clinical settings (Evensen et al 2010; Van de Ven and Johnson 2006; Rynes, Bartunek and Daft 2001; IOM 2001), where the delay in implementation of findings can lag into the decades or even centuries, as in the case
of a strategy to prevent scurvy in the British Navy (Mosteller 1981, quoted in Berwick 2003). Awareness of the research-practice gap is critical because of the opportunity for lost efficiency, harm to workers or clients/patients, and the overall threat to organizational survival over time.

This research-practice gap, and its implications, has been documented and elaborated in a wide range of disciplines. Lomas and Haynes (1988) report that there is a “distressing distance between health care knowledge in general and the practices of individual clinicians for most validated health care procedures” (Dopson et al 2003). The same can be said for education (Robinson 1998), social work (Herie and Martin 2002), library science (Booth 2003), and even physical education (Kirk 1989). Recently, the same argument has been made for the area of management. For example, Rousseau (2006) notes that “less than 1 percent of HR managers read the academic literature regularly (Rynes, Brown and Colbert, 2002), and the consultants who advise them are unlikely to do so either.”

§1.1: Translation of Research into Practice

A general approach to closing the research-practice gap is called the translation of research into practice. Translational research is an overarching area that applies both technological and social processes across the spectrum of research dissemination in order to promote application of research findings to practice in a swift and complete way. In a deep sense, the notion of translation is the manifestation of the scientific endeavor. The very notion of applying practically oriented knowledge toward action is an ancient concern, with roots in the Aristotelian notion of praxis. The underlying implication is that as science generates regularities about the world, individuals should apply this knowledge in practice to generate expected effects.

Yet, translational research is itself still under development, and remains multidisciplinary and pre-paradigmatic. Translation can focus on research-side issues such as linguistic, marketing and information science approaches to make findings more clear and accessible, or on practitioner-side issues such as social psychological and sociological approaches to promote the practical, cost-effective and long-term use of those findings. Much of what we know about translation is grounded in strategies that have reliably failed. The Agency for Healthcare Research and Quality (AHRQ), a leader in supporting and disseminating research on translation in healthcare, notes two “debunked myths” to close the research-practice gap. The first they call the “supply side” myth, in which researchers believe that if they publish more findings, practitioners will be more likely to find and apply those results. The second they call the “better packaging” myth, in which researchers or intermediaries, such as professional societies, condense or bundle
findings into guidelines or other packages of information believing that practitioners will be more likely to access, read and understand those results. Probably the most general conclusion of the translational agenda to date is the realization that praxis cannot be passively attained, not only because of socio-behavioral challenges, but also because of the sheer volume of new research that must be absorbed, to which Mulrow (1994) notes that it has “become clear that reliance on passive diffusion of information to keep health professionals up-to-date is doomed to failure in a global environment in which around two million articles on medical issues are published annually.” And sometimes, of course, evidence is just not as useful as one would hope – see Figure 1 below.

There are probably as many approaches to the translational agenda as there are commentaries on it. Some suggest rejecting the notion of a one-way arrow from research to practice, instead opting for forms of research, like community-based participatory research or practice-based research, that use a knowledge co-production network to blur the lines between the research and practice domains (Nutting et al 1999, Minkler 2005). Others, arguing in a vein similar to this essay, note that the research-practice gap has been an enduring feature of social science because of our mistaken belief that the “relationship between knowledge of theory and knowledge of practice entails a literal transfer or translation of one into the other” (Van de Ven and Johnson 2006). They note that whereas:

“knowledge of management practice is typically customized, connected to experience, and directed to the structure and dynamics of particular situations,... scholarship is committed to building generalizations and
theories that often take the form of formal logical principles or rules involving causal relationships” (ibid.)

Other approaches are more traditional in their orientation. This essay will contend with a particular instance of the translational agenda known as evidence-based management.

§1.2: Evidence-based Management

Evidence-based management is a process that encourages “the systematic application of the best available evidence to the evaluation of managerial strategies for improving [organizational] performance” (Kovner and Rundall 2006). Although its philosophical roots extend back to mid-19th century Paris and earlier, the modern concept developed out of a drive to reduce unexplained variance in clinical practice in healthcare. The push for evidence-based medicine derived from a desire to deemphasize “intuition, unsystematic clinical experience, and pathophysiologic rationale as sufficient grounds for decision-making” (JAMA 1992), having grown out of a clinical learning strategy developed at McMaster University Medical School in the 1970’s and 80’s (Sackett et al 1996, Rosenberg and Donald 1995). The evidence-based process was designed to inject systematically collected, up-to-date science into a practice of medicine that varied widely, and that was based more on the art of case-by-case analysis than on generalizable science. As a testimony to its popularity, since the late 1990’s, this concept has spread to nursing, public health, health policy-making, and other specialty areas in the health sciences (Kovner and Rundall 2006), intended to focus decision-makers on the “strategies best suited to create and sustain high-performing organizations, and [increase] the probability of the successful implementation of evidence-based medicine” (Shortell, Rundall, and Hsu 2007). The concept has also taken root in non-health disciplines such as policing (Sherman 1998) and education (Thomas and Pring 2004). More recently, the applications of these ideas to evidence-based health management (Walshe and Rundall 2001), and later to the general evidence-based management of a firm (Pfeffer and Sutton 2006), have brought the process from health care into the core transactional, operational, and strategic management processes in general business practice. Adherents to evidence-based management argue that managers who use this approach are more likely to develop strategies reflecting current understanding about mechanisms of action and empirical regularities, thereby producing better decision-making outcomes. The apparently straightforward nature of the process,

1 I agree with this characterization, but disagree with the authors on other items. Citing Baum and Rowley, they suggest that “organization theorists have never been positivists,” and argue for a “pluralist” view of the role of science based on a generous characterization of realism in the socio-behavioral sciences. As I suggest later, most of decision theory based on Simon is of a distinctly positivistic character, and a sociology of scientific knowledge perspective derived from a holist framework is probably a better route to addressing the research-practice gap in management.
combined with its reinforcement of managerial agency, seems to appeal to scholars and practitioners alike, providing the potential missing link in the translation research of practice. The refocus on evidence across disciplines has even evoked the utopian vision of an “evidence-based global society” (Donaldson 2009).

§1.3: Challenges for Evidence-based Management

Despite the promise of evidence-based management, it nonetheless faces serious operational and theoretical issues that threaten its usefulness as an organizational decision-making tool. Like translational research in general, evidence-based management is still pre-paradigmatic at its core – a point reflected in continuing challenges to and refinements of the practice in the literature. Issues as basic as the definition, role, uptake, perception and application of evidence-based processes remain open to further development.

The earliest formulations of evidence-based management were based on the design of its forerunner concept, evidence-based medicine. These models favored the increased use of research literature as the main function of the process, arguing not only that the evidence being used is sub-optimal, but also implicitly that much of it is simply not evidence. For example, Isaacs and Fitzgerald (2001) noted that a key function of evidence-based medicine is to curtail the unscientific approaches they argue run rampant, particularly in medical settings. While admittedly in jest, their observation belies the sentiment that eminence-, eloquence-, and confidence- (and other-) based management, marked respectively by radiance of white hair, smoothness of tongue, and bravado, are not borne of scientific and rational ground, and are therefore illogical to admit to a decision process (Isaacs and Fitzgerald 2001). The argument seems so natural that some have quipped: “who would argue that practice should not be based on evidence?” (Hammersley 2001). For those managers who make decisions haphazardly, enforcing a systematic connection to research literature is bound to impose some process, if not outcome benefits. Nonetheless, if we are to consider evidence-based management “a simple idea... It just means finding the best evidence that you can, facing those facts, and acting on those facts” (Pfeffer and Sutton 2006), then it is important to consider, within the management context, what counts as evidence, what the purpose of evidence is, and how evidence fits into the decision-making process.

In various ways, this point has been echoed over and over in the research literature. Some have focused on operational difficulties with the evidence-based management process in a management setting, noting that there is large variability across decision-makers in beliefs about the nature and content of the process (Rundall et al 2007) and that evidence-based management can be a disappointingly loose guide to decision-makers, because it “does not prescribe the kind of evidence,
how to obtain it, or what decisions should be made” (Rundall and Kovner 2009). Others have focused on theoretical concerns of using evidence in a management setting, such as the “congruence between properties of knowledge, properties of units, and properties of relationships between units” (Argote, McEvily and Reagans 2003), agreement as to “what counts as evidence in what circumstances” (Nutley, Davies and Walter 2003), and methods to “integrate and synthesize these different types of evidence, including tacit and explicit, quantitative and qualitative” (Glasgow and Emmons 2007). Still others have expressed concern about the endeavor itself, noting that: “First, what counts as ‘evidence’ and, second, what counts as ‘best practice.’ Both are open to interpretation” (Arndt and Bigelow 2009). And some are completely critical of it because of the “context-dependent, constrained by history and non-linear nature” of organizational life, writing:

“A hundred years after [Frederick] Taylor was putting forward similar ideas, [Denise] Rousseau [a prominent theorist on evidence-based management and former President of the Academy of Management] is still writing in the same tradition, believing that it is possible to derive law-like generalisations from the practice of management in organisations, to ‘measure’ what works, to identify independent variables and by doing so to ‘instil best practice’” (Mowles 2011).

In a special forum in Health Care Management Review, Arndt and Bigelow (2009) maintained that research is needed in several foundational areas, including investigating what practitioners make of various forms of evidence, and whether “different definitions of evidence are associated with different outcomes.” They also suggest that research needs to be conducted “that compares organizations that use ‘evidence’ with those that do not,” which reflects the simultaneous call by Rundall and Kovner (2009) to research “what happens when managers use and do not use EB management [under given conceptualizations of the process].”

A full understanding of evidence-based management is not trivial. Among the myriad issues that might play a role in the use of an “evidence-based” process in a management context, a few are listed below to exhibit the complexity:

- Characteristics of the evidence itself, including its ability to represent and control aspects of the world and its stickiness/transferability in an organizational context;
- Characteristics of the evidence source, with special emphasis on legitimacy, status and network position;
- Characteristics of the organizational routines and procedures related to evidence search and incorporation;
- Characteristics of the decision at hand, especially whether focused on discovery (e.g., strategy/innovation/non-routine) or justification (e.g., operational/ routine);
- Characteristics of the decision-makers, including their professional affiliation and dispositional factors (e.g., integrative complexity);
- Variables related to the organization’s capability to translate evidence into action, such as culture, formal structure, and absorptive capacity; and,
- The severity of the outcome errors that might accrue after an evidence-based management process, particularly the immediacy and reversibility of results and the interdependence between target organizational or environmental components.

This essay will not attempt to address all of these issues. Instead, it takes a different approach to resolving the difficulties faced by evidence-based management by reconceptualizing it as a cybernetic attention-directing mechanism. The key focus in this view is the role and use of the knowledge captured by evidence in addressing epistemic uncertainty in a management context. In this way, evidence-based management can more readily reflect the essential contextual, political and relational aspects of knowledge translation. The groundwork for this argument is presented in the following sections.

§2.0: Rewiring Evidence-based Management

Evidence-based management has been promoted in the business literature as a method for managers to make better decisions under the central assumption is that better implementation of better evidence will produce better decisions and improve performance. Underlying this assumption is that “better implementation of better evidence” will control the uncertainty threatening an organization’s attempts to produce reliable outcomes. Yet knowledge, not research evidence per se, is the means by which organizations reduce the epistemic uncertainty in outcomes posed by unpredictable environments. Thus, certain operational challenges faced by organizations adopting the evidence-based management project might not resolve easily by addressing the accidental issues of implementation, but would rather require a realignment of the whole endeavor.

In order to gain traction on the outstanding issues, I propose setting them aside momentarily in order to return to the grounding assumptions behind evidence-based management. As such, this essay will begin with uncertainty and build a new argument intended to “rewire” evidence-based management. My hope is that a new perspective on the topic will, if not useful in its own right, at least help organizations question the assumptions that might lead them into undesirable consequences.
§2.1: Managing Uncertainty

Uncertainty is an abiding feature of organizational life, or as Thompson (1967) puts it: “Uncertainty appears as the fundamental problem for complex organizations, and coping with uncertainty, as the essence of the administrative process.” It is a special concept that derives from the difficulties of making decisions in a future that is unknown (Knight 1921), and has a dual nature, both as something external and independent placing planning and operational constraints on the organization, as well as a feature of the environment perceived in differing levels of severity by members of the organization.

For both performance and survivability, environmental uncertainty is almost always seen as a threat to the organization. A recent historical analysis argues that “the approaches to organization studies, while differing on their view of organizational rationality, share the underlying assumption that organizations exist in a world teeming with uncertainty, an objective fact of life, which poses a real and actual threat to the business enterprise” (Shenhav and Weitz 2000). Because of the relevance of the threat uncertainty poses, organizations attempt to cope by determining the type and severity of the threats and by responding with corresponding strategies, such as buffering and bridging, to control the variability of inputs from the environment (Thompson 1967).

Organizations see uncertainty as driven by a lack of clear information, the long time-spans often necessary for definitive feedback, and the general difficulties in determining causal relationships (Lawrence and Lorsch 1967). In an empirical study, Duncan (1972) describes three features of uncertainty commonly referenced by organizational managers:

“(1) the lack of information regarding the environmental factors associated with a given decision-making situation, (2) not knowing the outcome of a specific decision in terms of how much the organization would lose if the decision were incorrect, and (3) inability to assign probabilities with any degree of confidence with regard to how environmental factors are going to affect the success or failure of the decision unit in performing its function.”

Yet, uncertainty is a special concept which is prone to confusion in common usage, and its character has important consequences for the manner in which an organization registers its potential severity and the strategies to be enacted. In strict logical usage, uncertainty refers to the “absence [or] insufficiency of a certain kind of knowledge” (Mattesich 1978). Wallsten and Budescu (1995) note that uncertainty takes two forms: it may be “due to external, quantifiable, sources of

---

2 Uncertainty should be differentiated from vagueness, “a conceptual problem arising out of the discrepancy between the logical form of a concept and its empirical manifestation,” and from inexactness, “the deficiency of a description or measurement”
random variation (aleatory), or to internal sources such as imperfect, or incomplete, information (epistemic).” The aleatory characterization is akin to the commonly accepted notion of risk, which can be described or measured, and can be predicted if derived from a true underlying distribution. The epistemic characterization, also called Knightian uncertainty, is “radically distinct from the familiar notion of risk,” and entails “action according to opinion, of greater or less foundation and value, neither entire ignorance nor complete and perfect information, but partial knowledge” (Knight 1921).

If the uncertainties affecting organizations are aleatory, then faster, higher quality collection of technical data and more adept statistical analysis are the key features in alleviating the threats posed. When these risks are quantified within confidence intervals, planning can occur as normal. However, if the uncertainty is of an epistemic or Knightian character, then the absence or insufficiency of particular knowledge, and the nature of the knowledge in forming opinion and providing foundation and value are critical features in determining how an organization should represent and respond to environmental threats. As Quiggin (1993) notes, dealing with uncertainty of this sort requires an examination of “the ways in which people collect, process and use information in decision-making,” and by extension, the ways in which organizations do the same. This is a contrast to the aleatory conception of uncertainty, which does not necessarily require special reference to opinion or value. This distinction, and the necessity to consider states of knowledge in epistemic uncertainty, creates a particularly pesky situation for social uncertainties, in which perception and knowledge are wrapped up in the determination of the severity and response to the uncertainties themselves. Lipshitz and Strauss (1997) elaborate 14 overlapping conceptualizations of uncertainty used in the organizational and behavioral decision theory literature; the closest concept for this essay is Weick’s (1979) definition in terms of equivocality as “the multiplicity of meanings which can be imposed on a situation.”

It remains an open question whether management evidence reduces social uncertainty in the same way that scientific and engineering evidence reduces technical uncertainty – and it will be an underlying assumption in the essay that these two forms of uncertainty are in fact different, and that to treat them similarly is to introduce the potential for serious unintended consequences.

§2.2: Uncertainty, Complexity and Diversity

Indeterminism and incomprehensibility are hallmarks of uncertainty, and are caused by the complexity of the organization’s environment. As the level of complexity increases, the ability to measure, comprehend and plan decreases; simple, transparent environments are easily predictable and rarely generate
uncertainty. Complexity is a feature of the world so prevalent and intuitive, yet so perplexing as to have a myriad of conceptions across the physical and social sciences. The Oxford English Dictionary (2nd ed. 1989) defines complexity generally as “composite nature or structure”, or “intricacy”, and specific to mathematics as “a measure of the difficulty of solving a class of problem,… the expected number of computational steps.” In science and engineering, complexity theory refers to the “mathematical study of nonlinear dynamic systems… [which] often follow power law distributions” (Farber 2003). Those in computer science define it in terms of information states; and psychologists as “the amount of information a choice involves” (Dijksterhuis et al 2006). In general, complexity refers to “elaborate temporal and spatial patterns and structures… [that] are hard to describe, explain, or predict” (Page 2011).

In a social context, complexity is “the number and diversity of players who are involved in a project. The more parties involved in a collaboration, the more socially complex” (Conklin 2006). Complex organizational systems are distinguished by (a) increased scale, or number of functional units, (b) increased differentiation, or number of units devoted to distinct tasks or technologies, and (c) increased interdependence, or the number of units coupled in authority-, resource-, or norm-based relationships (LaPorte 1975). Complexity is a structural variable that characterizes both organizations and their environments – with respect to environments, “complexity is equated with the number of different items or elements that must be dealt with simultaneously by the organization” (Anderson 1999). The consequences of complexity include dynamic tasks requiring highly-interdependent coordination, indeterministic processes, and subsystems with decentralized control.

Despite being predicated on a composite structure, complexity is not simply governed by the amount of differentiation, but also by the nature of the relationship between individual parts. Natural and social systems tend to exist in one of three states: simple, complicated, and complex (Rocco and Shen 2012). Dekker, Cilliers, and Hofmeyr (2011) describe the difference between complex and complicated, noting:

“Certain systems may be quite intricate and consist of a huge number of parts... Nevertheless, it can be taken apart and put together again. Even if such a system cannot practically be understood completely by a single person, it is understandable and describable in principle. This makes them complicated. Complex systems, on the other hand, come to be in the interaction of the components... [and are] held together by local relationships only. Each component is ignorant of the behavior of the system as a whole, and cannot know the full influences of its actions.”

---

3 As an example, they note that: “Jet airliners become complex systems when they are deployed in a nominally regulated world with cultural diversity, receiver-oriented versus transmitter-oriented
Some draw a distinction between *component complexity* as the raw number of potential interactions between units dependent on one another, and *combinatorial complexity* as the number of system interactions, including feedback loops and partially or completely delayed effects from decisions (Bush, Martelli and Roberts 2012). Others argue that complicatedness is not a feature of the system *per se*, but relative to the management of a complex system – whereas complexity is “an inherent property of systems..., [complicatedness] is a derived property that characterizes an execution unit’s ability to manage a complex system” (Tang and Saliminen 2001). In the context of uncertainty, all these are insightful; as the number of multiplex relationships between units increases, the more difficult it is to adequately describe the nature of these relationships and in turn, the more difficult it is to describe and manage the system.

Complexity can also be understood in terms of the capacity of a system to assume and retain “a large diversity of states or modes of behavior” (Schwaninger 2009). Diversity is in fact the hallmark of complexity, since a system possessing only one or a few unchanging elements would be relatively simple and describable, and generally predictable. The complexity that results from the increasing connectedness of the human world, “whether it works for us or against us, depends to a large extent on the amount of diversity” (Page 2011).

Diversity can be expressed in three ways: (a) variation within a type, (b) differences between communities or systems, and (c) differences across types (Page 2011). The first conception refers to variation in the statistical sense of a single parameter, such as height, taking measurements across a range. The second conception refers to compositional differences between systems, such that the component elements are the same, but their number and interaction can differ – for instance, members of the hydrocarbon family are all composed of carbon and hydrogen, but differences in number produce the emergent qualitative differences between poison, intoxicant and jet fuel. The final conception, differences across types, refers to a characteristic that can appear in multiple forms despite potential functional differences, such as the common characteristic of circularity shared by a quarter, frisbee and pizza pan, despite their very different uses. This final conception is the colloquial definition, and the one that produces many of the interesting qualities of diversity in complex environments (ibid.).

As Darwin observed in the biological world, environments display fantastic diversity in types. While the sociotechnical environment of organizations is more constrained, it likewise displays numerous actors differentiated not only in communication expectations, different hierarchical gradients in a cockpit and multiple levels of politeness differentiation (Orasanu and Martin, 1998), effects of fatigue, procedural drift (Snook, 2000), varied training and language standards (Hutchins et al., 2002), as well as cross-cultural differences in risk perceptions, attitudes and behavior (Lund and Rundmo, 2009).”
professional specialization, but also in “core processes and technologies, customers and markets, products and product lines, distribution networks, suppliers, or geographical locations” (Dooley 2002). This diversity yields complexity through the multiplex, often unobservable relationships between actors, and holds true in varying degrees for the focal organization as well.

§2.3: Diversity, Requisite Variety and Knowledge

In the midst of World War II, inspired by technical needs such as servomechanisms, and in tandem with scientific progress in the biological sciences, the subjects of homeostasis, systems regulation and the control of behavior in complex environments gained renewed attention. Starting in 1942, a series of interdisciplinary meetings between anatomists, psychologists, philosophers and social scientists sought to reconcile insights on how organizations exist in relation to and under the constraints of complex systems (Dupuy 2000). The field was dubbed cybernetics, deriving from the ancient Greek Кυβερνήτης (helmsman), a term related to steering, ruling and government. The early figures in cybernetics were aware of the importance of their work, believing that cybernetics “offers the hope of providing the essential methods by which to attack the ills – psychological, social, economic – which at present are defeating us by their intrinsic complexity” (Ashby 1956).

In addressing the way in which organisms self-regulate in complex environments, the cyberneticists became fascinated with the way in which organizations sense, measure and respond to the diversity of constraints the environment posed. Drawing on Norbert Weiner’s work on how living systems exhibit control functions and Claude Shannon’s theorem on disturbance in communication channels, W. Ross Ashby proposed the influential law of requisite variety in his 1956 book, “An Introduction to Cybernetics”. His great insight in this model was to focus on the notion of the variety of states and its consequences to a system’s regulation of diverse environmental disturbances.

![Figure 2: Ashby’s Model](image)

- **T** is the system to be regulated
- **R** is the regulator
- **E** is the essential variables of the system
- **D** is the environment that will affect the system **T** and impact its essential variables
In Ashby’s model of requisite variety, the elements D, R and E are defined by a set of possible states described by different values (see Figure 2 above). The *variety* of a system is measured by the number of different values in the set of possible values. For instance, if the system’s essential variable is the outcome “safe or not safe,” then its variety is two – as in the case of a computer bit which takes the two states on and off. The variety of a system can be measured by its *entropy*, which reaches a maximum when all states have equal probability. When the system has only one outcome state in its essential variable, it holds no variety and has zero entropy. Ashby shows that perfect control of a system (that is, no variety in outcomes, or zero entropy in E) is contingent on the system’s capacity to regulate disturbances from the environment. To achieve control and internal order (i.e., *negentropy*), the system regulator must increase the number of possible regulation states, thereby increasing its own variety, and resulting in an equal decrease in the variety of outcomes (Ashby 1956, Ashby 1958). Hence, the law of requisite variety: only variety (in responses) can destroy variety (in disturbances). In other words, the “centuries-old strategy of varying one factor” in simple or complicated systems is supplanted by the special strategy of requisite variety of regulation in complex systems – as Ashby notes, a complex system should not be approached “by analysis, for this process gives us only a vast number of separate parts or items of information, the result of whose interactions no one can predict. If we take such a system to pieces, we find that we cannot reassemble it” (Buckley 1967).

Requisite variety suggests that an organization should similarly meet the variety presented by a complex environment by increasing the diversity of its regulation states. In socio-cultural systems, Buckley (1968/2008) argues that an organization can control external variety by acquiring regulatory features, such as information, that allow it to discriminate, act upon and respond to its environment. The cybernetic view of an organization interacting with an open, complex environment is predicated on the conceptualization of a social system as a “set of elements linked almost entirely by the intercommunication of information” (Zaltman, Duncan and Holbek 1973). A study of general systems by complexity suggests that social systems are distinguished by the fact that “symbol-processing actors who share a common social order organize information from the environment into a knowledge structure” (Boulding 1956, Andersen 1999).

---


5 Entropy $H$ is measured as $H = \sum -p_i \ln(p_i)$, where $p_i$ is the probability of state $i$ occurring. In the two-state example, if the outcome *not safe* has a 1% probability of occurring, then E’s entropy is $H = -1\% \ln(1\%) - 99\% \ln(99\%) = 0.056$; likewise, if *not safe* has a 50% probability of occurring, then E’s entropy is $H = -50\% \ln(50\%) - 50\% \ln(50\%) = 0.693$. When there is only one outcome state, entropy is $H = -100\% \ln(100\%) - 0\% \ln(0\%) = -1 \ln(1) = 0$.

6 Ashby (1958) presents an inequality relating E, D, and R: $H(E) \geq H(D) + H_D(R) - H(R) \geq 0$, where $H_D(R)$ is the entropy of the regulator R, given environmental disturbances D.
Acquiring information and structuring knowledge are the tools by which socio-cultural systems regulate disturbances and respond with structural, cultural and technical strategies to achieve negentropy. In turn, these strategies “match the complexity of an organization’s structure with the complexity of its environment and technology” (Galbraith 1982), allowing the organization to “change its structure to adapt to the environment” (Zaltman, Duncan and Holbek 1973) and to implement core competencies, which if “well suited to the likely disturbances,” diminish environmental disruptions (Page 2011). In this way, socio-cultural systems can be seen as complex adaptive systems, which are open, able to exchange information between components, and depend on feedback of information to achieve increased order in the system (Buckley 1968/2008).

Thus, in order to discriminate disturbances and engage response strategies to manage environmental uncertainty, the organization must have a diversity of information ordered in a knowledge structure governed by actors in the system. This diversity can be determined by variation, type or composition, and its focus can be on reducing aleatory or epistemic uncertainty. When system actors discriminate and act upon disturbances through organizational responses, their capacity to respond is directly proportional to their capacity to perceive those disturbances. Controlling uncertainty by means of knowledge requires at least as much capacity to perceive and “make sense” as there are threats from the complex environment.

Through the lens of requisite variety, the complexity in the environment is “not an unintended side effect of an otherwise orderly world,” carrying with it only threat that needs to be curtailed (Ahlemeyer 2001). Instead, it can be seen as a potential for both threats and opportunities otherwise obscured to an organization unable to perceive and discriminate amongst them. In these terms, the organizational complexity retained by maintaining a diverse set of regulatory knowledge states can be conceived of as a “solution for a problem yet to be described” (ibid.). The challenge is creating and sustaining the right kind of complexity to match the environment.

§2.4: Knowledge Variety

If diversity of knowledge structures and perception are the key to managing uncertainty, then socio-cultural systems should develop capacity in these areas when operating in complex environments. Organizations can approach this task by maintaining technological, structural or procedural structures within the organization to provide the requisite variety of knowledge.

In practice, information on environmental disturbances is collected by organizational actors as evidence for use in strategic, operational and tactical
decisions. As noted above, evidence-based management can be a route to reducing organizational uncertainty and improving the overall quality of decisions, particularly strategic and innovation decisions. I argue that for these types of decisions, a strategy based on knowledge variety is better suited to identify environmental threats and opportunities than evidence-based management as currently operationalized.

Knowledge variety is a special form of the requisite variety argument about information diversity in which each of a set of epistemic knowledge categories is represented in equal balance by each decision-maker in the strategic decision-making process. As a decision-making mechanism, it is in the tradition of the rational decision logic of appropriateness, which is concerned with ambiguity and attention, and not of the rational decision logic of consequences, which privileges intentionality and bounded rationality (Frederickson and Smith 2003). In public administration theory, this distinction runs deep, having caused a long-lasting schism among public policy approaches (Cruise 1997) – and there is no question that the logic of consequences “won” in the academy. As Miller (2008) notes:

“For the past half-century, Simon’s framing of the question of rationality has focused behavioral researchers’ attention on deviations from economic rationality ... his research upheld a perspective on rationality that centered on cognition and decision making, as opposed to action, and treated values as beyond rational consideration.”

To wit, Aaron Wildavsky, founding dean of the UC-Berkeley Graduate School of Public Policy, once “silenced a roomful of candidates for the Nobel Prize for Economics by asking simply, ‘Where do values come from?’” (Chisholm 1995).

Where the logic of consequences is supported by a positivist foundation, the logic of appropriateness supports the inclusion of sociology of scientific knowledge (SSK) perspectives that I believe are important to representing values and enacting successful double-loop (revising processes by learning) and deutero (learning about learning) processes. I take this orientation to knowledge to be critical because of the epistemic drivers behind the concept of knowledge variety. There are two principal issues to consider: first, the philosophy of science providing the foundation for judgments on regulation states; and second, the epistemic typology that provides diversity for regulation states to operate effectively in this context.

§2.4.1: Making Inferences from the Environment is a Social Process

The knowledge structures that organizations develop by drawing on information from the environment support theories by which they can test hypotheses and act upon the world. This is the essence of epistemology, an area of

7 In fact, Herbert Simon was a student of Rudolph Carnap, a key figure in logical positivism
philosophy concerned specifically with the study of knowledge and justified belief, and more broadly, about issues “having to do with the creation and dissemination of knowledge in particular areas of inquiry” (Steup 2011). It is also an important part of the philosophy of science, a composite area of philosophy that focuses on the qualities of scientific theories, the function of discovery, hypothesis testing and confirmation, and the processes of establishing truth claims by deduction, induction or other means. It is not necessary to review the long history of philosophical stances in this essay; however, because the debate between Herbert Simon and Dwight Waldo about public administration is so critical to decision science and stems so clearly from differing philosophical foundations (see e.g., Harmon 1989), it is worthwhile to briefly note the major themes inasmuch as they affect the approach to collecting and structuring information from the environment. Thus, an abridged history of relevant argumentation leading from positivism to SSK perspectives is presented.⁸

As above, the motivating issues are uncertainty and how an individual, or a community of individuals, comes to know the unknown given only evidence from the environment. In an empirical context,

“Inductive logic is essentially the logic of uncertainty – that same uncertainty that lately has indelibly impressed itself upon the social and management sciences, leading to the center of decision theory. This created a growing realization that only by probing at the logical and epistemological roots of their disciplines, can administrative scientists bring order, stability, and coherence to their work. Yet inductive logic is still in turmoil and beset by conflict” (Mattesich 1978).

Induction is a “non-demonstrative inference leading to a conclusion referring to some or all members of a class, based on some evidence of these members” (ibid.), which has been “in turmoil” since the work of David Hume, an eighteenth century British empiricist whose work on inductive skepticism was the foundation on which logical positivism emerged. Hume believed that, since knowledge is a sequence, causation is the only relation to think beyond the present moment. His work shifted the understanding of scientific principles to one derived from a world-view based solely upon physical objects and their interactions. Hume held that there are no observable qualities of causation and that, since reason is only the organization of sense data, there is a problem with inductive reasoning. This problem arises because reason only serves to take the empirical evidence and sort it in the mind in a deductive fashion – it does not actually shed light on the world. The implication is

---

⁸The following discussion on philosophical topics is excerpted and revised from Martelli PF. (2001). “A Prolegomena to Justifiable Parsimony: Bridging the Gap between Science and Philosophy of Science.” University of Pennsylvania Philosophy Honors Thesis. That text casts a far wider net, describing the approaches of Lakatos, Hacking, Toulmin, Worrall, Laudan, Boyd and others in greater detail.
that, as experience alone can provide the key to the interpretation of reason, only basic analytic tautologies and synthetic empirical data should reign over philosophy, and thus no amount of analysis of an idea can show the necessary effect of a cause. Perceiving the regularity of an event (experience) is the only way to causally relate issues and predict their outcome. Since regularity is only a tenuous causal relation, Hume believed that probability does no more than to assume that the future will resemble the past through appeals to past success of inductive inferences. Justification is necessary for knowledge, but an inductive generalization cannot be justified (or even accomplished) by reason at all, and thus, the concept of probability itself loses potency. A necessary knowledge of nature, he argues, would “presuppose knowledge of the necessary connectedness of events” (Losee 1993).

Building on Hume, the logical positivist movement undertook a bold project to redefine the values of science in an extreme empiricist light. Positivism argues that the human senses provide the sole form of observational knowledge, and that only synthetic (empirical) statements verifiable through experimentation and observation, such as “on average, men are taller than women”, are admissible to scientific discourse and build scientific knowledge. Moreover, positivism is committed to the unity of science in a deductive hierarchy, or epistemic reductionism, whereby theories about entities at one level of theorizing, such as the socio-behavioral, can be fully explained by theories about entities at a more fundamental level, such as the biological and chemical. Rudolf Carnap, one of the major figures in positivism, proposed a verificationist conception of the growth of knowledge, analogous to trees with growing branches, such that as one proceeds up the tree, each following piece of data subsumes the last in a certain and cumulative manner. Carnap and others asserted that the content of a theory is exhausted by its observational consequences – or in other words, that if two theories entail the same evidence, they have the same meaning. Verification suffers from several notable problems, two of which are important here. The first is that it fails to account for empirical errors accumulated during experimentation, leaving a gap between discrete observational data containing errors and a theoretical language that is continuous and free of error. The second is that there is no distinction between meaningless (non-verifiable) statements and “what we can never know.” In this case, only propositions that can be known and characterized are scientifically meaningful; all theoretical statements not currently scientifically testable (e.g., Hilary Putnam’s famous retort about the possibility of being no more than a brain in a vat with a simulated reality”) are equivocal with verifiable statements (such as, “I am a person”).

Various alternate responses to Hume’s problem of induction and critiques of the positivist agenda appear in the philosophy literature. However, in order to introduce the SSK perspective as a response, it is worthwhile to review three

Motivated by the same concerns for empiricism, but not a positivist, Karl Popper postulated an opposing model. Popper believed that the growth of knowledge can be represented by inverted trees – that is, as the data gains coherence and cogency, the possibilities for their structure are cut off, until one available hypothesis remains. This refutation by experiment, or falsification, trims the list of alternatives along a deductive tree until a plausible hypothesis remains. While it tackles inductive skepticism beyond Carnap’s verification, the falsification process is still not able to confirm or justify a given hypothesis – though it does allow hypotheses about the world to be rigorously tested via crucial tests. Deductive testing of this sort (via modus tollens) uses the lack of supporting evidence to reject and eliminate a hypothesis from a group of competing hypotheses, leaving only the explanatory premise consistent with post-experimental phenomena. Thus, in an improvement over Carnap, bold and speculative theories are subjected to a Darwinian approach of conjecture and refutation, with remaining theories considered as the best alternative. In this way, the progress of science occurs by convergence-to-truth through stages of critical testing and refutation. While empiricist, Popper is often considered a precursor of social epistemological perspectives because, as Longino (2011) notes, the “practical sense of falsification refers to the efforts of scientists to demonstrate the inadequacies of one another’s theories by demonstrating observational shortcomings or conceptual inconsistencies ... [which] is a social activity.”

Both verificationism and falsificationism enforce a strict empiricism on the scientific method, retaining problems that can only be resolved by adopting a philosophy of science above the level of the hypothesis. Promoted by a reissuing of Pierre Duhem’s classic “The Aim and Structure of Physical Theory” around the time of Popper’s writings, two newly controversial ideas took root: that scientific theories meet experience as a unit, and that definitive falsification was not possible (Duhem 1954). In the text, Duhem described complicated cases where several hypotheses are involved in the prediction of a phenomenon. He then noted that if the predicted phenomenon failed to be observed, the conjunction of the hypotheses must be falsified. At that point, a scientist may go back and review the hypotheses to retain, replace, or modify them as necessary. When he applied this analysis of the logic of disconfirmation to crucial tests, Duhem concluded that an experiment is only crucial if every explanatory hypothesis except one was conclusively eliminated. In this way, he argued that no hypothesis is ultimately falsifiable, and that there can be no truly crucial tests. In his account, Duhem rejected the formulations of disconfirmation, and introduced the concept of local holism, where holism is the idea that a principal hypothesis is irrevocably conjoined to a more inclusive (and potentially infinite) disjunctive list of auxiliary hypotheses and background/initial
This argument was later expanded by W.V.O. Quine to include a more general, global holism, which stated more definitively that no evidence can conclusively refute any empirical hypothesis. This more radical position, known as the Duhem-Quine postulate, brings with it many other unresolved questions about the underdetermination of theories by data (i.e., identical sets of data can be confirming instances for disjoint theories).

In response to both Carnap and Popper and building on global holism, Thomas Kuhn (1962/1996) redefined theories of the growth of knowledge by proposing the paradigm as the basic unit of scientific progress. Kuhn defines a paradigm as a theory of science showing a certain reproducible pattern, deriving from a shared understanding of what the basic problems, methods and practices of a discipline should be. This pattern, however, is not simply a means for replication alone, but rather serves in the manner of a court precedent to a legal question – to offer a standard to be heeded. Paradigms are composed of four basic features: 1) general laws (symbolic generalizations) such as PV=nRT and F=ma; 2) the “common values” of empirical accuracy, consistency, broad scope, simplicity, fruitfulness; 3) metaphysical assumptions; and 4) exemplars.

From time to time in the course of discovery under a normal science, novel, unexpected facts are generated, including certain serious anomalies that cause trouble to the theory. When a paradigm is unable to explain or accommodate observational anomalies, even when it is maximally expanded, a crisis occurs that challenges the normal science practiced at the time, and vies for acceptance. In a revolution, the rival theories are logically incompatible, or “incommensurable,” since the different paradigms under which they operate represent different ways of seeing the world. When one theory wins out over another, a paradigm shift occurs, perception itself is influenced, and the scientific worldview changes. A paradigm shift is a theory change that occurs as a matter of persuasion invoking a psychological gestalt shift; this “religious conversion” is not the result of logic or evidence, but solely of rhetoric. According to Kuhn’s original premise, logic is not persuasive because scientists under different paradigms inhabit different worlds and speak different languages, so there is no neutral language of observation. Therefore, different paradigms have theory-laden data, which allows for no objective method of analysis in choosing a particular theory. Only “common values” provide scientists with a means to communicate via translation through shared beliefs.

---

9 General laws encapsulate phenomena that are included in the lexicon of a particular science. Values are the agendas of a theory, and place emphasis on a particular methodological inquiry within the paradigm. Metaphysical assumptions, once redacted in the logical positivist models, allow a paradigm to hold a commitment to a sort of model, such as mechanism, teleology, or atomism. Exemplars are examples of solved problems at the heart of a paradigm. A science is a paradigm with a project for the understanding and predicting of nature.
Kuhn’s holist views on the theory-ladenness of science constituted a serious anti-realist counter to the empiricist claims of his predecessors. Anti-realist philosophies of science, including instrumentalism and constructive empiricism, reject realist claims that the aim of science is to discover an objective reality with real relations ordered by independent scientific laws. At the extreme of these positions is social constructivism, which espouses that observation and other aspects of method are totally theory-dependent. Under this view, there is no knowledge of a theory-independent reality, rather only of a socially constructed one. Thus, because our understanding of scientific referents can change through a period of scientific revolution, reality is itself theory-laden. For a strong social constructivist, scientific belief in both the epistemological (i.e., explanatory) and metaphysical (i.e., actual physical presence) of entities is determined by convention among the scientific community. Strong realists, on the other hand, do not accept the Kuhnian notion of revolution through incommensurability, but rather want to show that the regulation of theory acceptance by features of the existing theoretical tradition is a “reliable guide to theory-independent matters of fact” (Boyd 1984).

The propositions of meaning dependence and theory-laden observation spawned an anti-realist zeitgeist about the nature and goals of science. One of Kuhn’s contemporaries (and a likely inspiration) was NR Hanson, whose concern with scientific discovery in a theory-laden context led him to propose a doctrine of meaning dependence, in which a “process of seeing that [i.e., perceiving] involves a broader system than merely experiencing a visual image; it requires and includes a conceptual organization that is bound to be subjectively tainted by previous experience” and through which we inescapably give “preference to one interpretation over others” (Mattesich 1978; emphasis in original). In Hanson’s view, theories provide “patterns of conceptual organization that render data intelligible... [and are] discovered not by inductive but by retroductive inferences (pattern recognition)” (ibid.). Retroduction, also known as abduction, is a form of inference originating with Aristotle’s third inference and interpreted by CS Pierce, which takes the general form: If E provides the best explanation for P, and P is the case,

10 Scientific realism, generally speaking, asserts that: 1) Our best current theories are approximately true; 2) The central terms of our best theories are referential; 3) The history of “mature” sciences shows increasing approximation to the truth; 4) The theoretical claims of science must be read literally, and are true or false; and 5) Science aims at a true account of the world.

11 It is amusing to note that, despite taking opposing views on empiricism and the power of observation, the positivist and social constructivist positions nonetheless have an anti-realist overlap. Although a positivist believes the realist claim that theories progress towards greater explanatory power, he or she would assert that the aim of science is not to represent an actual reality, so much as to sketch out relations of observable entities and regularities of occurrence of data, since the existence of unobservable theoretical entities is denied. The social constructivist takes a harder line, arguing that theories do not refer at all, but are rather constructed and used ad hoc given available phenomena.
then E is probably the case. Hansonian abduction\textsuperscript{12} is a mode of inference that is “intended to be ‘creative’ ... [and] has greater potential in accounting for the early phases of scientific investigation” (Paavola 2006). Hansonian abduction, contrasted with the falsificationist hypothetico-deductive model, provides a pragmatic way for an anti-realists to sidestep the problem of induction, and is a suitable foundation for pattern recognition and feedback (Bayesian-type) approaches to inference. This form of inference is nicely aligned with a notion of scientific discovery based on the perception of facts through values and the ordering concepts within a broader social context.\textsuperscript{13}

It would be interesting to review the wide variety of realist and anti-realist responses on this topic (see Figure 3 above for the various areas of application\textsuperscript{14}), but the highlighted points are sufficient to lay out the argument for knowledge variety. Where the social perspective cannot solve the problem of induction, it instead manages to redefine the issue into how a community comes to perceive and agree upon on a theoretical framework given the evidence available. In contrast to a positivist, or even strong empiricist, orientation to the scientific endeavor, holism and theory-laden observation established “that realist talk of ‘truth’ and ‘objectivity’

\begin{itemize}
\item \textsuperscript{12} Following Paavola (2006), I distinguish Hansonian Abduction from Inference to the Best Explanation.
\item \textsuperscript{13} There are, of course, many realist and middle-ground responses to the issues Kuhn and Hanson suggested. Of these, I particularly note the forms of entity realism/ theory anti-realism) defended by Ian Hacking and Nancy Cartwright. Hacking’s (1983) view particularly promotes intervention, rather than representation, saying, “So far as I’m concerned, if you can spray them [e.g., electrons] then they are real.” Even these close positions differ on their orientation toward Inference to the Best Explanation; Cartwright, for example, suggests Inference to the Best Cause as an alternative.
\item \textsuperscript{14} Credit for this excellent figure belongs to Zoltan Domotor, Professor of Philosophy at the University of Pennsylvania, with whom I took eight classes focused on philosophy of science topics
\end{itemize}

22
no longer has a place in our scientific world-picture” (Norris 2004), or at least, that “every individual choice between competing theories depend[s] on a mixture of objective and subjective factors, or of shared and individual criteria” (Kuhn 1977). Building on this platform, a continuum of perspectives has arisen to reorient the scientific agenda away from empiricism and towards the social dimensions of science. This wide area of study is often referred to as the SSK or as science and technology studies (STS). These social models of scientific inquiry build upon the holist, theory-laden foundation of observation, focusing on the critical role of negotiating shared values in a social context in managing problems of incommensurability and underdetermination. As Jasanoff (2004) notes, the subject’s most fundamental concern is “the emergence and stabilization of new objects or phenomena; how people recognize them, name them, investigate them, and assign meaning to them; and how they mark them off from other existing entities, creating new languages in which to speak of them and new ways of visually representing them.” Other topics include the shared perspectives and standards of a community of researchers, the nature and role of public reasoning, the interplay of science in law and politics, and trust in the authority of the scientific community.

To summarize this section of the argument:

The motivating issues are uncertainty and how an individual, or a community of individuals, comes to know the unknown given only evidence from the environment. In this section, the positions of positivism and the sociology of scientific knowledge are introduced and contrasted along an abbreviated historical continuum. Among the responses to Hume’s inductive skepticism, positivism is a perspective uniquely unable to address serious questions of epistemic uncertainty, especially in a social context. Moreover, because it lacks a framework to incorporate discrete errors accumulated during experimentation into continuous theoretical language, it is ill-suited to the organizational learning through feedback built into cybernetic systems models. As such, a positivist orientation towards empiricism, in which the content of theory is exhausted by observational consequences, is not appropriate for bold conjectures about uncertainty and does not take into account the holist social dynamics that make sociobehavioral regularities distinct from corresponding laws in the physical sciences. The purpose here is not to suggest that organizational managers and researchers are all positivists, but rather to reveal the background assumptions driving empirical research in the traditions of Simon’s logic of consequences and Waldo’s logic of appropriateness based on their distinct responses to the problem of induction, and to assert that approaches under the SSK umbrella are superior in addressing epistemic uncertainty.

Similarly, I want to build an argument that regulation of epistemic uncertainty in an organizational management context depends on social perception. As Mattesich (1978) notes:
“the Weltanschauung-philosophers [e.g., Kuhn, Hanson and Feyerabend] have adopted a holistic point of view (which is also the hallmark of the system experts), and thus are looking at science from the vantage point of a structural perspective. Many systems theorists, in turn, have adopted the subjectivistic attitude which characterizes the Weltanschauung-philosophers.”

Subjective knowledge structures affect how individuals perceive the aleatory or epistemic character of uncertainty of the environment. As Starbuck and Milliken (1988) note, “perceptual frameworks categorize data, assign likelihoods to data, hide data, and fill in missing data.” Moreover, complex environments alter the structure of decision-making, since “the environment in which decisions are made is key, not simply... as a setting but as an embedded entity which forms both ‘substance’ and ‘arena’ for the strategic actors” (Connolly and Koput 1997, quoted in Gore et al 2006). In socio-cultural systems, subjective knowledge structures are formed interpersonally, and these form the regulatory mechanisms by which organizations discriminate, act upon and respond to uncertainty in the environment.

I propose that knowledge variety is a regulatory mechanism that allows organizations to select between alternate models of the world and navigate epistemic uncertainty, despite the problem of underdetermination. Diversity of the type employed in knowledge variety should also promote organizational double-loop and deutero learning by providing organizations with the robustness of regulation states to capture environmental feedback apprehended in knowledge types outside of the expected category.

§2.4.2: Knowledge, not Information, is the Object of Regulation

Knowledge cannot be information alone. That observation is theory-laden is sufficient to show that individual knowledge is distinct from objectively true facts or information about the entities in the world. As noted in the SSK approach, this distinction magnifies in a social context, where the shared perspectives, standards and goals of a community influence the status of knowledge claims.

Instead, knowledge is a quality of the knower, and functions as a “justified [or, justifiable] true belief” about how the world works (Takeuchi and Nonaka 2002). Knowledge is context-specific and relational; tied to a particular stance, perspective, or intention; and is always compiled in support of a particular end. Discrete observers can “know” characteristics of a singular facet of the world differently, but can reconcile those differences through interpersonal communication to build a socially-negotiated model of the environment.
Our current understanding is that knowledge takes various forms ranging from knowledge on the nature of relationships between variables to pragmatic understanding on implementation of action (Ekblom 2002), and can be categorized along several useful dimensions. Whether knowledge is public or private (Uzzi and Lancaster 2003), and the extent to which it varies along tacitness, observability, independence and complexity dimensions (Birkinshaw, Nobel and Ridderstrale 2002), affect the ease with which an organization is able to codify and transfer it (Winter 1987). The distinction between forms of knowledge is long-standing – Aristotle wrote of five “states of virtue [in] which the soul possesses truth by way of affirmation or denial,” namely craft (techne), scientific knowledge (episteme), practical wisdom (phronesis), philosophical wisdom (sophia), and comprehension or intuitive reason (nous) (1139b lines 14-18, cited in Ackrill 1989 translation).

This Aristotelian distinction remains in debates of knowledge translation by means of know-about, corresponding to episteme, and know-how, corresponding to techne. In modern terms, know-about is theoretical scientific knowledge conveying representations of the “nature, formulation, natural history and interrelations of social problems” (Morton and Nutley 2008), whereas know-how is tacit knowledge “relating to the accumulated internalization and recursive interpretation of the skills required for professional competence/expertise through learning-by-doing” (Gasson 2005). In other words, where know-about is like a blueprint, know-how captures a procedural dimension that is important to implementation in organizations (Kogut and Zander 1992). In current usage, the concept of know-how is built on Polanyi’s explication of tacit knowledge and a general background of gestalt psychology. Polanyi argued against the epistemic reductionism and empiricism of the positivists, claiming that personal experience contains more knowledge than can be reduced into sensory data, and that personal judgment and commitments make impersonal objectivity impossible. Instead, tacit knowing is the foundation of personal experience and knowledge of the world. The tacit, or background, knowledge inhering in individuals is prior to focal, or operational, knowledge about a situation and governs the blending of reality with experience (Polanyi 1962), providing a thought anchor to categorize sensory input and helping to explain the “taken-for-grantedness” of entities in the world (Sveiby 1997). In turn, gestalt psychology, a holist approach to perception, was “the outcome of an active shaping of experience performed in the pursuit of knowledge... [and] the great and indispensable tacit power by which all knowledge is discovered and, once discovered, is held to be true” (Polanyi 1966). In this way, he reasserts the importance of different ways of knowing, one of a more intellectual kind and the other of a more practical kind, that is:

“both the ’wissen’ and ‘können’ of the Germans, or the ‘knowing what’ [in this essay, know-about] and the ‘knowing how’ of Gilbert Ryle.15 These two aspects of knowing have a similar structure and neither is ever present

---

without the other. This is particularly clear in the art of diagnosing, which intimately combines skillful testing with expert observation” (ibid.)

Building on know-about and know-how, there is a growing convergence among translation researchers on a larger typology of knowledge to represent the spectrum of tacit and explicit forms of knowledge important to the translation of research into practice. Given that these categories derive from the study of organizational translation, they address the recognition, structuration and application of theoretical knowledge. As such, I argue that this typology is not limited to organizational translation, but can be extrapolated to the wider socio-behavioral context because of the concurrent personal and interpersonal (social) nature of all knowledge, as noted above.

Enhancing the “epistemology of possession” described by Cook and Brown (1999), Gasson (2005) elaborates four categories of knowledge and maps them onto four different “ways of knowing” corresponding to individual/group and tacit/explicit combinations. Her typology is presented in Table 1 below.

---

16 I follow Miller's (2008) lead on categorizing know-how in a typology: “To argue that tacit knowledge is never fully convertible to explicit knowledge is consistent with Polanyi’s thinking. However, the presence of a tacit dimension to knowledge does not negate the possibility of making some tacit knowledge explicit; it merely makes us aware that such conversion always remains incomplete.”
Table 1: Gasson Knowledge Mapping

<table>
<thead>
<tr>
<th>Category</th>
<th>Knowledge included in category</th>
<th>Ways of Knowing (Cook and Brown 1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Know-what</td>
<td>Explicit knowledge relating to organizational facts and conventions.</td>
<td>Concepts represent individual, explicit forms of knowledge – or generically subjective forms (Weick, 1995).</td>
</tr>
<tr>
<td>Know-why</td>
<td>Explicit knowledge relating to global rules and models of behaviour [sic]; Tacit knowledge, relating to local and socially-situated, normative practice.</td>
<td>Stories relate the application of specific knowledge to specific situations, providing a rationale for its application and a cultural set of meanings that legitimize that rationale within the local organizational culture.</td>
</tr>
<tr>
<td>Know-how</td>
<td>Tacit knowledge relating to the accumulated internalization and recursive interpretation of the skills required for professional competence/expertise through learning-by-doing.</td>
<td>Skills embody tacit, individual knowledge, and permit that knowledge to be applied unreflectively, so that it becomes internalized and communicated through shared practice. Genres are shared conventions and practices that make group knowledge identifiable and accessible for a specific community of practice.</td>
</tr>
<tr>
<td>Who-knows-what</td>
<td>Explicit knowledge relating to individuals’ wider social networks of knowledge sources, that may be local or global.</td>
<td>Across community boundaries, process exemplars, claims of expertise, and the use of specific genres are mobilized to access and to influence perceptions of who-knows-what. As groups become more stable other mechanisms may come into play.</td>
</tr>
</tbody>
</table>

Excerpted from Gasson (2005)

A slightly expanded typology is presented in Table 2 below and will be used in this essay. Not only is this typology developed from similar reasoning and corroborated with numerous observational accounts across varied socio-behavioral contexts, but it is also the product of a consensus of translation researchers after the first of a series of in-person, day-long seminars on the use of evidence in practice and policy. In this typology, know-about and know-how are again drawn as distinct
categories, to which three additional categories – *know-why*, *know-what*,\(^\text{17}\) and *know-who* – are added. The translation researchers acknowledged that values, exemplars, and the individuals involved are as important to the process of organizational decision-making as mechanistic or pragmatic knowledge, and as such, have similarly special status among knowledge types. Thus, this typology captures those refined elements of knowledge that have otherwise been subsumed within *know-about* and *know-how* – for instance, *know-what* could be viewed as a special case of *know-about*, but is distinguished because of its particular capacity in directing socio-behavioral systems towards action.

<table>
<thead>
<tr>
<th>Table 2: NORFACE Knowledge Typology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category of Knowledge</td>
</tr>
<tr>
<td><em>Know about problems</em></td>
</tr>
<tr>
<td><em>Know-why (you might implement a change)</em></td>
</tr>
<tr>
<td><em>Know-what (has worked)</em></td>
</tr>
<tr>
<td><em>Know-how (to put a change into practice)</em></td>
</tr>
<tr>
<td><em>Know-who (to involve)</em></td>
</tr>
</tbody>
</table>

This typology was abstracted and augmented from Ekblom (2002) in Nutley, Walter and Davies (2007), and subsequently reproduced in Morton and Nutley (2008) in the Proceedings of NORFACE Seminar #1: Types of Knowledge for Evidence-Based Policy, held 26 November 2007 in Edinburgh, Scotland. I make a few minor additions and one notable change as described.

To illustrate this typology, consider the concept of “safety culture” drawn from the literature on organizational reliability, referring to Table 3 below. In this typology, knowing that safety culture aligns with the values of an organization is distinct from knowing the mechanisms by which safety culture affects communication or examples of how safety culture has worked in other contexts.

\(^{17}\) It is important to note that I have slightly revised the definition of *know-what* from the original, which describes this knowledge type in terms of interventions that “will bring about desired outcomes [my emphasis].” To include the future tense implies a certain fundamentally reproducible causation that is unwarranted and unnecessary here; it is possible to retain the sense of looking to the successes of others without treading on the inductive fallacy by simply changing “know what will work” into “know what has worked.”
Table 3: Knowledge Typology Illustration

<table>
<thead>
<tr>
<th>Category of Knowledge</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Know about problems</td>
<td>Definition of safety culture and the mechanisms by which it affects communication in groups</td>
</tr>
<tr>
<td>Know-why (you might implement a change)</td>
<td>Symbolic, emotional, ethical and cultural meaning of enacting a safety culture intervention</td>
</tr>
<tr>
<td>Know-what (has worked)</td>
<td>Existing safety culture interventions, such as training sessions, that have produced desired outcomes</td>
</tr>
<tr>
<td>Know-how (to put a change into practice)</td>
<td>How to practically implement and evaluate an effective safety culture-focused intervention</td>
</tr>
<tr>
<td>Know-who (to involve)</td>
<td>Internal and external collaborators to advise and support a given safety culture intervention</td>
</tr>
</tbody>
</table>

If knowledge (or knowing) takes multiple forms, then a socio-behavioral regulatory approach to epistemic uncertainty must account for the different types that can present themselves in the environment. Following the arguments above, these types become the specific regulation states maintained by an organization in order to regulate the epistemic uncertainty of the environment.

§2.4.3: Knowledge Diversity and Knowledge Variety

In the above sections, I argued that making inferences from the environment is a social process, and that knowledge, not information, is the object of regulation. In order to regulate the environmental uncertainty, the organization must build knowledge structures using diverse states within a typology of knowledge. As long as the interactions between diverse types “don’t produce exceptionally large negative interactive effects [which in the case of knowledge types will likely be the opposite], then the average performance of diverse collections should perform better than that of more moderate or homogenous collections” (Page 2011). Through a requisite variety lens, Buckley (1968) notes:

“The concept of requisite deviation needs to be proffered as a high-level principle that can lead us to theorize: a requisite of socio-cultural systems is the development and maintenance of a significant level of non-pathological deviance manifest as a pool of alternate ideas and behaviors with respect to the traditional institutionalized ideologies and role behaviors.”

Still, the diversity needed for organizational regulation can be maintained through three routes: through variation, composition and type. Increasing variation within a given knowledge type, such as adding more know-about, is moot as a driver for the organizational regulation of epistemic uncertainty because additional sensor
capacity within a single regulation state will not account for the diversity of type present in the environment. In fact, there remains “the disturbing possibility that when people experience uncertainty and gather information to reduce it, this often backfires, and uncertainty increases” (Dorner 1996, quoted in Weick 2001). In other words, more information is not always better – a knowledge regulation structure is necessary to control epistemic uncertainty.

In socio-behavioral systems, diversity by composition refers to using group structure to bring together specialized individuals into a composite. In this case, composition refers to specialization of knowledge across individuals within a group; in other words, building decision-making units where each individual deeply focuses on one of the five knowledge categories above, later bringing their specializations together. This is a plausible route to developing requisite variety. In the context of diversity and complexity, Page (2011) notes that: “as the task becomes further and further decomposed, the relative advantage of specialization increases to infinity ... specialization creates an advantage within a complex system.”

Several theorists have approached the issue with a compositional focus. In a general sense, organizations can maintain conceptual slack among organizational members, where conceptual slack is the “divergence in analytical perspectives among members of an organization over theories, models, or causal assumptions pertaining to its technology or production processes” (Schulman 1993). In itself, conceptual slack is not inconsistent with other formulations of diversity, as will be seen below. However, some executions of the idea presume between-individual diversity in regulation states. One immediate issue that a compositional solution faces is how to coordinate the knowledge regulated across individuals into a cohesive organizational knowledge structure. All socio-behavioral endeavors face some form of this problem when pursuing interpersonally-negotiated goals; there is a healthy literature describing organizational integration through various coordination methods. However, there a few other interesting caveats that make compositional diversity subordinate to type diversity in this case. Knowledge is both

---

18 In a team context, divergence in perspectives can be measured by means of interpretive ambiguity. Teams with high interpretive ambiguity tend to lack clarity on the level of similarity among team members' perceptions, but achieve high performing teams when exhibiting “both early interpretative ambiguity and late heedful interrelating” (Kilduff et al 2000). In this way, interpretive ambiguity “preserves the requisite variety needed to sense and regulate the variety facing the organization [through compositional diversity]” (ibid.). Conceptual slack can also be represented in terms of the types of knowledge accessed through the networks of organizational decision-makers. Knowledge heterogeneity refers to the “variety of knowledge, know-how and expertise to which a manager has access through her [social] network” (Rodan and Galunic 2004). Knowledge heterogeneity among one’s contacts is important for innovation, and may help performance “even when there are no [social network] holes to be bridged” (ibid.). However, with respect to innovation performance, “the fact that structure only matters in the presence of knowledge heterogeneity would seem to [argue for the mechanism of] idea generation (purely knowledge based) and implementation of those ideas” (ibid.).
the regulator and the object of regulation. As a regulator, it exists both in explicit forms and personal tacit forms, and changes not only in relation to the environment, but also by interpersonal influence within the organization. As such, individuals can maintain specializations, but an organization can never fully decompose the regulation function through individuals. Moreover, the regulation provided by individuals is not definitively specialized because it depends on individual theoretical representations of the environment, which themselves grow and change. As a feature of uncertainty, knowledge is socially understood and socially represented both by the organization and by third parties. Per the philosophical argument above, theories of the world are socially influenced such that elements of the epistemic uncertainty change as individuals react and change. Because it can present in forms that change not only quickly and complexly, but also as a function of the regulation itself, a compositional strategy using specialized regulators will constantly need to adjust the time and location of individual sensors – a strategy requiring deliberation that is bound to fall behind threats and opportunities in the environment. Furthermore, individuals specialized to focus on one knowledge type devote their attention to perceiving one element of the gestalt of uncertainty that they apprehend, which, under the logic of appropriateness, creates an organizational attention issue.

In an excellent piece on sensemaking and boundary-spanning design, Gasson (2005) examines how an organization develops knowledge structures across a knowledge typology. Her project is similar to the one in this essay – but because her argument excludes the cybernetic and inductive skepticism foundations above, she does not explicitly describe the process in terms of knowledge regulation. As such, it is difficult to infer the various forms of diversity assumed by the focal boundary-spanning collaborative group. Although there is some discussion along the lines of a compositional mechanism (e.g., “specific domain expertise” of individuals) and there is a rich discussion of how knowledge structures are negotiated, a discussion of the diversity by which her focal group drew knowledge from beyond the system boundary is unfortunately opaque. Rather, she provides a description of how this organization moves from early to late group representations in a sensemaking process under politics and power constraints. Her conclusion that “progress may be made in distributed design, by manipulating the task and knowledge-sharing focus of the group, or by the use of specific knowledge sharing and elicitation mechanisms” (Gasson 2005) overlaps with the gist of the diversity argument presented here.

19 Although, I suppose, this task could be arranged through some random, non-deliberative rotation of specialized sensors (individuals) across environmental boundary-spanning points, which would sometimes capture and sometimes overlook knowledge to be regulated. To be useful, the rotation would need to take place relatively often and smoothly, which would still require active management of individuals and intensive coordination to create organizational knowledge structures.
Type diversity maintained by individuals across the organization is another approach to regulation, and *knowledge variety* is a special case within this class. Type diversity obtains by having the range of types represented within the regulators, and in this case, by ensuring that individuals within the organization (and especially those critical to the decision-making process) have each regulation state available. Because capacity to capture the full range of knowledge is present in each individual, even if they might not prefer or excel within a given type, the argument turns to ensuring that these individuals are attuned to capturing these elements from the environment.\(^{20}\)

Yet, the individuals could still maintain a large enough specialization toward one type (e.g., know-about) as to reduce this diversity at the group level back to a poorly-implemented diversity of composition. For example, an organizational decision-maker might undertake a decision in which 60 percent of the rhetorical justification was derived from explicitly codified, mechanistic knowledge from journal articles in the decision-making process, 20 percent from explicitly codified “best practice” reports from colleagues, 20 percent from tacit, front-line knowledge from workers, and no other externally-derived knowledge. If this were the case, he or she would be an organizational knowledge sensor that was regulating by knowledge type, and heavily privileging know-about over know-what and know-how. Instead, under the constraint of knowledge variety, each decision-maker maintains an equal (or near-equal) proportion of each of the five knowledge categories. In other words, that organizational decision-maker would constrain his or her decision-making process to draw on (approximately) 20 percent of each of the five knowledge types for justification of a given decision. In turn, the interaction of individuals using this approach should direct organizational decision-attention and weight persuasive evidence more robustly than a scenario in which each individual decision-maker enters with a specified argument, agenda and politics to support his or her conclusions. Some benefits to this approach will be presented below.

The context for *knowledge variety* is that of an organization maintaining a requisite variety of knowledge to address environmental uncertainty, where the regulators of knowledge must be individuals. Therefore, while the construct is measured by robust knowledge diversity within individual, the character of uncertainty control and concomitant outcomes are emergent at the organizational level. This means that there are two levels important to knowledge variety: the first is the individual level of knowledge diversity expressed, and the second is the balance of knowledge diversity across individuals in the decision-making group.

---

\(^{20}\) Perhaps it is important to note that machines today cannot perform these knowledge regulation functions because they are not able to capture the tacit and value-laden quality which inheres in human sentience. For many reasons, I personally do not believe that machines should replace human observation and representation of the world, particularly with respect to values, but I suppose I am agnostic to the possibility that they could.
§2.5: Evidence-based Management in the Context of Knowledge Variety

In the last section, it was argued that knowledge variety is a specific cybernetic operationalization for the apprehension of knowledge and organizational decision-making, designed to take advantage of regulatory knowledge structures for organizational discovery and innovation. This section will examine the concept vis-à-vis models of evidence-based management.

In its naïve form, evidence-based management focuses on the incorporation of codified, mechanistic knowledge above other types. In terms of requisite variety, simply using more information, no matter how refined or compiled, accounts only for the boundedness of rational decision-making, and does not adequately account for complexity. In its least restrictive, but still structured form, evidence-based management is essentially a directive to think critically and systematically about management options. Kovner and Rundall (2006) describe the process in eight steps, and Rundall and colleagues (2007) in six broader categories based on research into evidence-based management implementation. These steps are outlined in Table 4. Ultimately, however, evidence-based management “does not prescribe the kind of evidence, how to obtain it, or what decisions should be made” (Rundall and Kovner 2009).

<table>
<thead>
<tr>
<th>Table 4: Established processes for evidence-based management</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Identification of a problem</td>
</tr>
<tr>
<td>2: Identification of a decision</td>
</tr>
<tr>
<td>3: Allocation of weights to criteria</td>
</tr>
<tr>
<td>4: Development of alternatives</td>
</tr>
<tr>
<td>5: Analysis of alternatives</td>
</tr>
<tr>
<td>6: Selection of an alternative</td>
</tr>
<tr>
<td>7: Implementation of the alternative</td>
</tr>
<tr>
<td>8: Evaluation of decision effectiveness</td>
</tr>
</tbody>
</table>

Evidence-based management has typically been defined with a focus on using the best available evidence to evaluate managerial strategies. Table 5 presents four of well-cited definitions of the process. Because this is an insight that has remained central, it is useful to briefly overview what is meant by the terms evidence, best, and available, and the consequences resultant from differing definitions.
Table 5: Common definitions of evidence-based management

<table>
<thead>
<tr>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The systematic application of the best available evidence to the evaluation of managerial strategies for improving the performance of organizations”</td>
<td>Kovner and Rundall 2006</td>
</tr>
<tr>
<td>“Evidence-based management is a commitment to finding and using the best theory and data available at the time to make decisions.”</td>
<td>Pfeffer and Sutton 2006</td>
</tr>
<tr>
<td>“[T]he conscientious, explicit, and judicious use of current best reasoning and experience in making decisions about strategic interventions”</td>
<td>Kovner, Elton and Billings 2000</td>
</tr>
<tr>
<td>“A paradigm for making decisions that integrates the best available research evidence with decision-maker expertise and client/customer preferences to guide practice toward more desirable results”</td>
<td>Rousseau 2006</td>
</tr>
</tbody>
</table>

§2.5.1: What is the Best Available Evidence?

*Evidence* is “ground for belief; testimony or facts tending to prove or disprove any conclusion” (Oxford English Dictionary 2nd ed. 1989). It is unnecessary to reiterate the justification that information must be socially abstracted to produce knowledge, making knowledge different from data. A similar argument holds for knowledge and evidence – just as knowledge is information tied to a particular stance, perspective, or intention, so is evidence knowledge tied to supporting or undermining a conclusion. In a socio-behavioral context, evidence is, by definition, concerned with perception, belief and persuasion. Whereas knowledge can exist free-form, evidence can only exist as a package of knowledge directed towards a goal. For organizations, this means that evidence is always wrapped up in context, shared meaning, and interpersonal goal reconciliation. Confusing the difference between evidence and knowledge is almost as problematic as confounding knowledge and information. Schwandt (2009), citing Upshur (2002), notes problems with the way evidence is conceptualized, emphasizing the abstraction of evidence from knowledge; he writes:

“At present, we face some difficulty and confusion with understanding the term evidence in evaluation because it is often taken to be synonymous with the term evidence-based. The latter notion, however, has two shortcomings. First, it is narrowly interpreted to mean that only a specific kind of scientific finding – that is, evidence of causal efficacy – counts as evidence. Second, the idea of an evidence base (at least in its original formulation in evidence-based medicine) suggests that evidence is the literal foundation for action because it provides secure knowledge.”

The latter point highlights the perception, belief and persuasion intrinsic to evidence in a social context. Secure knowledge for action does not necessarily obtain because the persuasive value of evidence exists in a range across multiple dimensions. Kuhn first underscored this general point by proposing the common values (i.e., empirical accuracy, consistency, broad scope, simplicity, and fruitfulness) by which individuals are able to discuss and reconcile different
scientific paradigms. Referring especially to evidence-based practice, Donaldson (2009) proposes relevance, coherence, verisimilitude, justifiability, and contextuality as the common values which govern the use of evidence in organizations. Like most social propositions, the dimensions of value in evidence are often in tension – for example, Keller (2009), citing and supporting prior research, suggests that features of salience, credibility, and legitimacy are interconnected such that procedures developing one tend to undermine another. Thus, rhetoric using these values, as with Kuhn’s common values,21 plays a large role in persuading individuals to switch gestalts between positions using an evidence-based process. Drawing on Aristotle's convention of rhetorical persuasion, Van de Ven and Schomaker (2002) suggest that evidence-based medicine has only concentrated on logos, an appeal to the strength and consistency in logical argument, to the exclusion of pathos (an appeal to values, beliefs and experience) and ethos (an appeal to credibility). This insight is equally applicable to evidence-based management, and its focus on privileging particular forms of knowledge, rhetoric, and systems of inquiry.

In this light, it is less secure to determine what constitutes the best evidence. Best has traditionally been established with an underlying assumption of logos, with the “best” evidence meeting the epidemiological standard of the randomized controlled trial (RCT). However, where evidence is better, it is also worse. In evidence-based medicine, virtually all institutional reviewers of evidence grade expert assessment as the lowest strength of evidence,22 The problem with this characterization in socio-behavioral settings is twofold. First, it leads to what some have called the “parachute approach to evidence-based medicine” (Potts et al 2006), referring to an earlier tongue-in-cheek article calling for an RCT to establish definitively whether parachute use prevents trauma due to “gravitational challenge.” This view advocates making policy decisions on “good science” even when RCTs are unavailable; in other words, strength of evidence is one of many considerations including the fiscal and sociopolitical climate within which governments, institutions, and communities operate (Tang et al 2003). This is a progressive position that appreciates the importance of action when “best” evidence of evidence. Using only codified sources of evidence, such as journal-based RCTs, assumes they can act as substitutes for non-codifiable types of knowledge.

---

21 Kuhn softened his hard-line incommensurability argument in his later writings, adding the list of common values by which others are persuaded. We might assume that the act of persuasion through rhetoric is supported by evidence on the part of the persuaders.

22 A sample of grading schemes – USPSTF: Level III = Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees; ICSI: R = consensus statement, consensus report, narrative review, and X = medical opinion; SORT: C = Recommendations based on opinion, consensus guideline, usual practice, clinical experience; GRADE: Very low = Any other evidence (except RCT or observational study); Oxford Center: 5 = Expert opinion without explicit critical appraisal, etc.
In fact, expertise itself can be seen as part of a continuum of evidence that plays a special role in organizational decision-making. The social side of evidence exists on a continuum of voice (see Figure 4 below). Consider a thought experiment with a continuum of voice running from one extreme representing pure institutions and the other representing pure democracy. At the low end of the continuum, all rhetorical legitimacy and authority to resolve knowledge claims rests within institutional experts. At the high end of the continuum, local knowledge prevails empowering every individual an equal share on the claims of theoretical reference.

In a socio-behavioral context, neither extreme is tenable – both experts and local knowledge must be incorporated into decision-making. On the one hand, experts play a crucial role in decision-making by providing “valuable and reliable information on the state of the knowledge in their field, how to solve problems, and on the certainty of their answers” (Meyer and Booker 2001). This tacit, background knowledge also “allows individuals to limit the factors which they consider to be important in a decision,” to systematically structure them, and to discriminate among information (Bennett 1998). Experts also use “fast and frugal” heuristics to process information (Gigerenzer and Goldstein 1996), and are able to define a problem space and focus attention to its features (Chisholm 1995). On the other hand, experts are no better than non-experts in prediction and at making judgments outside of their domain, as evidenced by their poor long-term forecasting (Citrin and Tetlock 2008) and “fractionated expertise” (Kahneman and Klein 2009). Moreover, “when predictability is low, experts are often more susceptible to overconfidence” (Griffin and Tversky 1992). Likewise, local knowledge plays an important role in grounding impressions to ideographic conditions, and in allowing easier implementation through co-production, governance and stakeholder buy-in (Ansell 2007). Local knowledge represents the egalitarian view of voice. Individual non-experts possess value-filled, geographically discrete and cognitively idiosyncratic knowledge. This idiosyncratic knowledge provides a local grounding for “naturalistic generalization” from a nomothetic, generalized evidence base.
(Stake 1978). However, compiling and using real local knowledge can be elusive – becoming too discrete in collecting data assumes that each data point is equally accessible, comprehensible, and valid, and furthermore that each individual perspective is equally legitimate to include as evidence. Including either expertise or local knowledge sources separately means including either too many or too few voices and perspectives (see Figure 5).

In addition to ranking, “best” can also be envisioned in terms of a best match, or contingency, between evidence and a situation. One stream of literature promoted by translation researchers associated with the Research Unit on Research Utilisation (RURU) at the University of St. Andrews might be described as a “contingency approach” to evidence-based management (see e.g., Nutley, Walter and Davies 2007). In this framework, studies of organizational implementation successes and failures (i.e., know-what) are aggregated by disciplinary application to suggest combinations of organizational, individual, evidentiary, source and search factors that promote high performance. Although reasonable in stable conditions, this approach is problematic under turbulent environmental conditions. Consider that finding the right evidence to support action given a contingency of multiple social factors depends on knowing what those factors are and knowing when they are permanent features or when (and how often) they change. In other words, it is far easier to put a puzzle together when you know you have all the pieces and you are sure they aren’t changing. The conditions when epistemic uncertainty is the highest are the conditions when an organization is least likely to be able to determine and adequately manage at least some of the necessary factors of contingency. Like the problem of Eco’s map, it becomes almost as difficult to determine where you are in reality as where you are on your map.
Using the best evidence implies that it is available to the decision-maker at the time of the decision. Available can be interpreted in two ways. Evidence is transmitted through sources, yet sometimes these sources are literally unavailable to them in time for a decision. Researchers in translation and information science have documented various common technical barriers and facilitators to compiling evidence, such as the cost of journals and difficult technological interfaces (Rundall et al 2009). These are important, but comparatively simple issues to address. Available can also refer to what can be comprehended by the decision-maker or what can be attended to by the organization. An individual modeling his or her world under certain assumptions may not be able to conceive of competing knowledge claims and may reject evidence as rhetorically unpersuasive. In other language, boundedly-rational individuals will never be able to conceive of all the possible states of nature (Simon 1947/1997), and individual models of decision choice under uncertainty are subject to the incompleteness hypothesis, which asserts that “because [a decision] model fails to capture all relevant aspects of the problem, it will yield inaccurate estimates of the expected benefits of any given course of action” (Quiggin 2009). Likewise, an organization may have influential individuals or sub-systems that attend to certain types of evidence more than others, leading the organization, through socio-behavioral drivers, to privilege that evidence in rhetorical justification to the exclusion of other knowledge claims. In this case, the evidence similarly becomes unavailable to decision-makers.

Cognitive availability is one of the key benefits of knowledge variety compared to other formulations of evidence-based management or even a general form of knowledge diversity. The relatively equal weight of knowledge types is important because an organization has limited attention available to search and process evidence, where attention is defined as the “noticing, encoding, interpreting, and focusing of time and effort by organizational decision-makers on both (a) issues ... and (b) answers” (Ocasio 1997). Firms faced with “too much data and not enough information,” compel organizational decision-makers to “oversimplify to deal with overload” (Matheson and Matheson 1998). The focus of attention is important for discovery, innovation and strategic action. For instance, both the total number of sources, and the number of sources across several knowledge types used exhibit an inverted-U shaped relationship with corporate innovation (Laursen and Salter 2006) – search breadth alone itself doesn’t yield more robust attention.

A number of mechanisms focus and redirect attention for organizations, including the availability and saliency of cues, which lead to anchoring and encourage confirmation bias (Tversky and Kahneman 1974), the consequences of groupthink (Janis 1972) and echo chambers (Uzzi 2008), the role of categories and critics (Zuckerman 1999), and cultural-cognitive drivers of imitation (DiMaggio and Powell 1983). In general, the “valuation and legitimization of issues and answers” is a principal mechanism focusing the attention of organizational decision-makers.
In the context of reducing epistemic uncertainty, a knowledge variety approach assists the organization in balancing the “valuation and legitimization” of mechanistic, implementation, and policy knowledge, with less danger of becoming anchored to a given framing and placing too much attention on a particular perspective. As noted, decision-making that does not have enough variety in its input states suffers from an overly narrow mapping of the complexity in the environment.

On the one hand, a narrow range of evidence tends to privilege better precision between goals and evaluated strategies, but assumes that the problem has been correctly specified given uncertainty. Under complex conditions, a narrow range approach is particularly apt to produce so-called Type III errors, in which the wrong problem is solved precisely without solving the root cause affecting their particular case. The probability of resolving at the expense of solving a problem, or of “solving the ‘wrong’ problem precisely,” is known as an Error of the Third Kind, or a Type III Error. These sorts of errors are linked to epistemic uncertainty, often occurring because key variables are omitted in the specification of a model (Mitroff and Linstone 1993; Mitroff and Featheringham 1976). On the other hand, the benefits of more evidence breadth include a better ability to catch weak signals of failure or success, but at the costs of redundancy or overlapping signals due to parallel systems of mental models about the uncertainty faced. In this context, weak signals are instances of knowledge that comprise the epistemic uncertainty of the environment, but are known to organizational regulators in highly restricted ways, such as evidence that appears for a brief time, to a small number of individuals, or in times or places removed from the point of decision.

Given the argument for attention, it is possible to view maintaining a variety of knowledge functions as a form of epistemic hazard identification, in which individuals perceive threat and opportunity cues in the form of knowledge from the environment. Inasmuch as uncertainty is governed by identifiable cues, increasing the number of salient cues reduces the unpredictability of outcomes. Instances of knowledge are cues, providing insights into the mechanism, implementation, and policy knowledge necessary to make sense of complex, uncertain environments.

§2.5.2: Organizational Attention and Sensemaking

If knowledge variety prevents organizational attention problems in decision-making, then there must be a mechanism to coordinate individual knowledge at the organizational level. As Zaltman, Duncan and Holbek (1973) remind:

“Generally speaking, it is important to note that although the creation of feedback mechanisms [such as knowledge variety] is a prerequisite for adequate handling of emerging implementation problems, it is not in itself a
‘safeguard’ assuring such handling. The information that the feedback mechanisms provide must be interpreted to be useful.”

Because the feedback is multiplex knowledge, the manner by which an individual makes sense of this “information” operates in line with the philosophical arguments above. There are several ways to model these “alternative logics for [individual] decision making in social settings” (Messick 1999); a few are detailed in a note below.23

For organizations to interpret the information from feedback mechanisms – in other words, to build an organizational knowledge structure from the environmental knowledge regulated by individuals through a knowledge variety approach, organizations require interpersonal coordination. Sensemaking is the “mutually-negotiated understanding of how to make sense of the local, organizational ‘world’ of work and interaction” (Weick 1995); it is the process by which individuals in group settings come to a shared mental model of a situation and develop the groundwork for goal reconciliation, frequent and timely communication, and effective joint operation. Mental models are collective cognitive representations that range from a distributed configuration of representations with no overlap between individuals, to overlapping representations, to identical representations among individuals (Klimoski and Mohammed 1994). The transfer of organizational knowledge requires shared mental models, and the “ability to define relevant knowledge-domains is essential for collaborative sensemaking” (Gasson 2005).

At the organizational level, knowledge variety both depends on and complements the sensemaking model in developing shared mental models. To map the environment, the organization acts as a “medium” which interprets the properties of the environment. Having many, independent elements and weak internal constraint relative to external constraint are properties that characterize good mediums (Weick 1976, Weick 1979). Operationalized as knowledge variety, all individuals involved in the decision-making process act as regulation elements, and these individuals regulate environmental uncertainty through balanced knowledge categories, rather than through constrained boundary-spanning structures or processes. Maintaining a balanced set of knowledge categories ensures that they are

23 Messick (1999) nicely summarizes: “Connolly and Koput (1997) called some of these theories, ‘naturalistic decision making’ [NDM] theories, and characterize them as theories that address the inadequacies of expected utility theories and their descendants. This class of theories includes image theory (Beach, 1993), explanation-based theories (Pennington and Hastie, 1988; Lipschitz, 1989), and situation matching theories (Klein, 1989; Noble, 1989). Both image theory and explanation-based theories emphasize narrative coherence and internal consistency as essential components of evidence evaluation and action planning, while situation matching theories resemble pattern recognition programs in which decisions depend on which of a set of pre-stored patterns best match a data profile.” Although I do not intend to argue it here, it is likely that many NDM theories argue inference from a Hansonian abduction reference point.
always available to decision-makers as a “consensually validated grammar for reducing equivocality by means of sensible interlocked behaviors” (Weick 1979).” Moreover, maintaining balanced knowledge allows a “consensually validated grammar” that equally admits knowledge valued in logos, pathos and ethos forms of the rhetoric of persuasion into debates about organizational strategy. In this way, sensemaking communication becomes robust regardless of the “system of inquiry” being engaged. In his extensive study of crisis management, Mitroff describes how different inquiry systems meet different kinds of problems and produce different kinds of new knowledge (Mitroff 2012). While simple problems can be solved through analysis or consensus, the ill-structured, complex and unbounded problems that arise from high epistemic uncertainty carry with them a “mess” of complex inputs, including symbolic, emotional, ethical and cultural concerns, which can only be solved through a pragmatic system, such as appreciative inquiry. While not all strategic or innovation decisions are “messes,” they fall within the realm of systemic epistemic, and not aleatory, uncertainty. Maintaining an equal share of non-mechanistic, tacit know-why knowledge in sensemaking across individuals is the only way to ensure that these concerns are maintained in shared mental models of environmental uncertainty. Given that reliability “is a moving target” (Weick 2010), a regulation approach is particularly important for instances where safety and reliability are concerns.

Of course, maintaining knowledge variety does not mean that individuals think the same things – only that they have a common spectrum of evidence to be used with others during rhetorical persuasion. The sensemaking process is a method by which individuals come to create shared mental models, resulting in some level of convergence or divergence of perspectives. This convergence or divergence can be expressed in terms of “cognitive ties” between individuals.24 Strategies that privilege mindfulness, or a “rich awareness of discriminatory detail and a capacity for action,” use expanding cognitive sets to increase diversity in cognitive perception (Weick 2010). Creating and retaining diversity in knowledge types is a key way of increasing the organizational capacity to recognize patterns of information indicating threats or opportunities inductively from the environment. When evidence-based management is operationalized in a way that emphasizes

---

24 To the extent that two actors see things in the same way, they have a close conceptual connection or strong cognitive tie to each other; likewise, to the extent that they diverge in perspectives, they have a distant conceptual connection or weak cognitive tie to each other. Like most things, strong and weak cognitive ties exist in a tradeoff with each other. On the one hand, strong ties encourage cohesion, trust and private knowledge-sharing, but favor exploitation of existing strategies and create echo chambers among decision-makers (Uzzi 1997, Hansen 1999, Wang, Ashleigh and Meyer 2006). On the other hand, weak ties encourage creativity and are beneficial to organizational search (Burt 2004), but are also resource and cognitively intensive, and may be less acceptable intra-organizationally. Studies within the network, diffusion, and implementation literature parse out a multitude of barriers and facilitators including time of implementation, network position, actor and diffuser characteristics, innovation attributes, and various combinations therein.
narrow knowledge within and between individuals (say, know-about through journal articles or know-what through best practices), it results in restricted interpersonal sensemaking communication and a reduced capacity for the organization to map the environment. In other words, knowledge variety governs the sense made in sensemaking.

§2.5.3: Best Practices and Overadoption

When used heedlessly, evidence-based management can generate undesirable organizational outcomes. It is important to note that evidence-based management has a dual nature – both a suggested method of improving socio-behavioral technologies in the organization, as well as a socio-behavioral technology in itself. This suggests that organizations should be careful to note and distinguish methods of evidence use.

The fact that evidence-based management has a dual nature as a technology used to evaluate other technologies makes it a rich topic for study. One area which allows the exploration of both of these aspects is the concept of guidelines or external “best practices.” Best practices may be considered as developed internally or adopted externally from the firm. An internal best practice is the “replication of an internal practice that is performed in a superior way” (Szulanski 1996); this type of practice implicitly carries intra-firm contextual and relational components of knowledge critical to action, and is more likely to transfer than external knowledge (Darr et al 1995). An external best practice is a set of structural and procedural conditions that are widely-acknowledged to be associated with beneficial outcomes. While they often include information about implementation, they rarely fully capture and transfer tacit, private and contextual knowledge about the site or sites involved. External knowledge is more likely to be preferred by the organization, and valued as unique and special (Menon and Pfeffer 2003). Moreover, for many reasons, managers are predisposed to available answers and are generally more concerned with equifinality than with extensive evaluation, arguing “we don’t want the evidence, we want the answer” (Rundall et al 2009).

There appears to be an attentional issue playing out in the example of how organizations notice and adopt best practices; an issue which could be solved by a knowledge regulation approach. For example, Eddy (2005) suggests that there are two approaches to evidence-based management – that is, through evidence-based practice guidelines (i.e., an externally-developed algorithmic model) and through “evidence-based individual decision-making.” The guidelines approach emphasizes an externally standardized method for practitioners to apply; whereas the decision-making approach advocates a standard process to be applied in each case individually. The former approach is evidence-based management “lite,” in that the
individual decision-maker only selects the best procedure from a supermarket of options. Although Eddy’s argument is applied to the application of evidence-based practices in medicine, the insight provides an analogue to the management of organizations. Guideline models provide general, deductive, and nomothetic knowledge. But, applying externally-developed models like guidelines as observed is problematic because, without a meaningful connection to idiographic conditions of the organization, they encourage the mere copying of standards representing effects averaged across multiple organizations.

Many proponents of best practices, such as the American Productivity and Quality Center (APQC 2008), address and attempt to assuage this concern, noting that:

“There is no single ‘best practice’ because best is not best for everyone. Every organization is different in some way – different missions, cultures, environments, and technologies. What is meant by ‘best’ are those practices that have been shown to produce superior results; selected by a systematic process; and judged as exemplary, good, or successfully demonstrated. Best practices are then adapted to fit a particular organization.”

Yet, even with caveats, there remain problems with the notion of a best practice. First, from a practical point of view, while best practice examples “are very useful in showing that alternatives for malfunctioning practice do exist..., they have a limited value for guiding wide scale implementation” (den Hertog, Groen and Weehuizen 2005). Second, best practices are an example of know-what works knowledge, so they should be part of a larger decision schema, rather than a model to be copied with amendments. Third, there is a reasonable chance that as environmental complexity increases, managers will nevertheless migrate toward the false certainty of best practices, displacing genuine judgment based on a range of knowledge in favor of copying the averaged best practices of others. Finally, along the lines of the critique of best above, even a best practice perfectly matched to an organization would only be a delayed “best” – that is, the findings of that best practice would have been best for a previous environmental situation and given some lag in translation, leaving no assurance that the match for the adopting organization is capturing its current environmental complexity in a similar way.

For a variety of reasons, Greenhalgh, Kyriakidou, and Peacock (2004) prefer a modified definition that relaxes the intensity of the notion of a “best” practice. Citing past research, they comment:

“Two additional concepts should therefore be considered here: ‘best practice’, defined by Zairi and Whymark (2000) as ‘a task, function of behavior which, when carried out, produces above average results’; and ‘potentially better practices’, defined by Horbar et al. (2001) as practices that have been shown (or which are believed) to improve outcomes in one setting, and which can be
selected, modified and applied in unique ways to fit a new situation, which
takes account of the fact that ‘best practice’ in one setting is only potentially
an improvement on existing practice when transferred elsewhere.
Interestingly, in their study of potentially better practices, Horbar et al.
made no attempt to verify whether the practices actually improved outcome –
indeed, they comment that the critical impetus for quality improvement may
be the process of pulling together to implement anything that improves or is
perceived to improve outcome, not the practice itself.”

Setting aside the subtlety between best and better practices, these practices
may cause problems in organizational strategic decisions around innovation. An
innovation is “any idea, practice or material artifact perceived to be new by the
relevant unit of adoption” (Zaltman, Duncan and Holbek 1973). In management
terms, innovation can also operate as an organization-level variable describing the
adoption of “an idea or behavior that is new to the organization,” where newness is
determined by “comparing the organization’s present with its past structures,
processes, and objectives” (Rundall, Starkweather and Norrish 1998). Rogers (1962)
points out that innovation is determined by the “newness of the idea to the
individual [or organization] that determines his reaction to it,” and not “whether or
not an idea is ‘objectively’ new as measured by the amount of time elapsed since its
first use or discovery.” Other definitions exist (e.g., Greenhalgh, Kyriakidou
and Peacock 2004), but all refine the main points that innovation is perceived to be new
and is discontinuous with former practice. According to these definitions, innovation
takes on a different connotation than its colloquially understood one. That is, not all
organizational changes are innovations, not all innovations are completely new, and
innovations can range in scale from major reorganization to incremental
adjustments of existing processes.

In cases of strategic innovation, the organization is seeking to adopt a new
structure or process. In his discussion of diffusing innovations, Rogers (1962)
described not only adoption and rejection of innovation, but also “irrational”
overadoption and underadoption. When an organization selects a strategy based on
best practice, it is in danger of overadoption or “heedless” adoption. Overadoption is
a circumstance in which an organization chooses to adopt an innovation from the
marketplace when it is inappropriate to the conditions at hand. Citing an example
of the overadoption of the 2,4-D chemical weed-killer in 1949-50, Rogers (1962)
noted that:

“Many past researchers have implicitly assumed that to adopt innovations is
desirable behavior and to reject innovations is less desirable. In all cases, this
may not be true. ... Overadoption often results from insufficient knowledge;
overadopters perceive the innovation as a panacea.”

Denis et al (2002) give an example in healthcare management in which private-
practice surgeons in Quebec, pressured by professional and market demands to use
laparoscopic cholecystectomy, rapidly adopted the new technique. In that case, whereas the “benefits of the procedure (shorter hospital stay, smaller scar) were readily observable, the risks (damage to internal organs, need for re-operation) were much less visible” (Greenhalgh, Kyriakidou, and Peacock 2004).

When the organization uses best practice as a model for a new structure or process to be adopted, it must do so heedfully to avoid overadoption. Overadoption is most likely to occur under uncertainty, which acts as a “powerful engine that drives mimetic isomorphism, and is an important and potent force that encourages imitation which accords the organization, and its activities, legitimacy” (Shenhav and Weitz 2000). To maintain legitimacy and gain reputation in uncertain environments, organizations look to and take on characteristics from high-performing or highly-legitimate organizations, and in so doing, they begin to take on common features, look and act alike. This copying is described in the neo-institutional process of mimetic isomorphism, in which the drive towards similarity results not just through the shared norms and values of institutions, but as a result of diffuse, cultural-cognitive “taken-for-granteds” (DiMaggio and Powell 1983). In his discussion of organizational fads and fashions, Abrahamson (1991) describes how these processes drive the overadoption of technically inefficient strategies and technologies by providing symbolic utility to maintain legitimacy and gain reputation. For example, Coleman et al (1966) note that credibility of the diffusing source is positively related to adoption decisions, and Rao, Greve, and Davis (2001), in a study of Wall Street analysts, concluded that “social proof is most influential when decision makers are uncertain about the value of a course of action and when they are able to observe the actions of similar others.” Kaissi and Begun (2008) summarize the literature on fads and fashions in healthcare, citing an example of overadoption suggesting that “hospitals attain legitimacy directly by copying other [exemplar] hospitals that are perceived as arbiters of best practices or indirectly by using consulting firms with specific expertise in this area” (Arndt and Bigelow 1992). They conclude their review with the exhortation that:

“Honest confrontation of bandwagon jumping in health care would alert students and executives to the need to proactively challenge so-called best practices from other organizations (does it really fit here?) and alleged success stories (are the findings really clear and valid and persuasive?)” (Kaissi and Begun 2008).

However, not all organizations engaged in strategic innovation blindly copy the success of others. For example, organizations that adopt innovations early in the process of diffusion tend to adopt for rational justifications and customize those innovations to fit the organizations needs, whereas later adopters tend to adopt for normative justifications, taking on innovations as observed without customization (Westphal, Gulati and Shortell 1997). In fact, this mechanism itself is subject to fine differences: Kennedy and Fiss (2009) argue that “early adoption is associated with
opportunity framing and motivations to achieve gains, both economic and social, while later adoption is associated with threat framing and motivations to avoid losses, again in both economic and social terms.” Thus, opportunity framing is associated with customization, whereas threat framing is associated with the lack of customization. Moreover, the lack of customization is associated with normative justifications, such as increasing institutional legitimacy.

In cases of overadoption in innovation, insufficient knowledge, poorly-observed consequences, and a sense of “panacea” leads to copying the practices of successful or legitimate exemplar others. The drivers of this copying are normative justifications generated through diffuse, cultural-cognitive “taken-for-granteds.” Yet, these taken-for-granted views of the world also structure the information generating models of an uncertain environment. It would be an error to assume that generating rational justifications or becoming an early adopter would solve the problem of overadoption, or lack of customization of best practices. Rather, increasing knowledge, observing consequences, and dispelling the sense of panacea should be the key solution. If increasing knowledge related to the uncertainty being faced is at the heart of the solution, then increasing the diversity of environmental regulators through knowledge variety, and not through variation diversity, is a critical solution.

§2.6: Concluding Thoughts on Theory

In the sections above, I have argued for the construct of knowledge variety by appealing to diversity in socio-behavioral regulation, requisite variety, the social structure of knowledge, rhetorical persuasion, epistemic categories, and several ways in which this strategy is better than non-regulator based strategies.

To that end, it overlaps with, but ultimately differs from two existing theories of note. First, the argument is akin to Social Information Processing (SIP) theory (Salancik and Pfeffer 1978), which argues that individuals process environmental information based on the salience and relevance of cues, and that these cues structure individual attention and interpretation of needs and preferences, particularly with respect to motivation. At the individual level, knowledge variety argues in the same direction, adding a balanced epistemic typology as a suggested mechanism of interpreting the environment. However, at the organizational level, the cybernetic and sensemaking frameworks diverge from an SIP understanding of how an organization interacts with its environment. Second, the argument shares common ground with triangulation, one of the four key concepts of qualitative research (Gittelson 1996). Triangulation generally refers to the incorporation of several different data collection methods (Jick 1979), although it could refer to data,

25 The others being flexibility, iteration, and the creation of context.

46
investigator, theory or methodological triangulation (Denzin 1978). Webb et al (1966) suggest that “once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. The most persuasive evidence comes through a triangulation of measurement processes.” To the extent that maintaining diversity of knowledge types provides multiple perspectives on a proposition, knowledge variety and triangulation overlap. However, knowledge variety expounds a larger argument about the manner in which individuals triangulate triangulated knowledge through sensemaking, and the ensuing consequences for organizational attention.

Moreover, there is an implicit assumption in the argument. Strategies based on diversification contribute to the resilience and robustness of systems, which increases survivability. Yet, while “systems with more types will be more robust because they are less likely to have bad outcomes..., [this] does not mean that diverse approaches will be the top performers. To the contrary, the best portfolios in any given year won’t be diverse” (Page 2011). Therefore, the implicit assumption is that complex organizations in complex environments should be principally concerned with survivability, robustness and reducing “bad outcomes,” rather than taking risky chances for risky payoffs. In management lingo: we can only chance upon the black swan, but we are always running from the red queen.

There is one important addendum. To the extent that we know what drives performance, “then we should select the best collection on the basis of that information... [otherwise,] if we are not sure of what we’re doing, we should err toward greater diversity” (Page 2011). As management goes, I leave this determination up to the reader. The received scientific knowledge of socio-behavioral regularities is still very limited, and as we aggregate to a systems level of analysis, our theories are underdetermined and value-laden. Having studied slow-moving and socially-created catastrophes, as well as patient safety and quality concerns in the healthcare setting, I lean towards robustness in the face of uncertainty.

Moreover, I do not disagree with some readings of the translation of research into practice and evidence-based management projects. For instance, Briner, Denyer, and Rousseau (2009) describe evidence-based management in a measured way, using specific instances of evidence to essentially increase regulation states; they conclude:

“Evidence-based management is about making decisions through the conscientious, explicit, and judicious use of four sources of information: practitioner expertise and judgment, evidence from the local context, a critical evaluation of the best available research evidence, and the perspectives of those people who might be affected by the decision.”
Or, writing about “translation,” Greenhalgh and Wieringa (2011) suggest that it is “time to drop the ‘knowledge translation’ metaphor,” because it will “inadvertently close our minds to alternative framings of this complex field.” Instead, they propose “phronesis” (practical wisdom) [a term that I will also use below], ‘mindlines’, ‘knowledge intermediation’ and even ‘language games’” (ibid.). Or perhaps, as Rousseau (2006) writes: “Evidence-based practice is not one-size-fits-all.”

§3.0: Investigating Knowledge Variety in Practice

The theoretical sections above argue that knowledge variety is a preferable operationalization of evidence-based management for organizations facing the epistemic uncertainty of complex environments. The following sections will lay out a project to examine the features and possibility of knowledge variety in practice. As noted above, knowledge variety plays an important socio-behavioral role in generating interpersonal knowledge structures that are robust to uncertainty, maintain organizational attention, and prevent heedless overadoption. This suggests that a knowledge variety approach produces better organizational outcomes than static versions of evidence-based management for innovation decisions.

One arena in which this might be observed is in the capacity of knowledge variety to prevent the overadoption of best practices. Variety in evidence-based management is responsible for balancing attention among a group of decision-makers, and when employed in a manner avoiding heedlessness, should lead to the customization or rejection of best practice adoptions. Customization becomes an organizational marker of modularizing best practices from legitimate exemplars, such as competitors or knowledge intermediaries, to meet idiographic conditions.

Therefore, the guiding research question is how knowledge variety is expressed in organizations and how this expression affects the adoption of best practices. The level of knowledge variety across decision-makers leads to decisions with more acute organizational attention. In turn, higher organizational attention results in more careful tailoring of adopted innovations to the ideographic conditions. In terms of a proposition:

An increase in the variety of knowledge during decisions to adopt best practices is associated with an increase in the customization of those adopted best practices.

Organizational knowledge variety of decision-makers

Attention effect

Customization of adopted best practice
In order to establish this relationship, two more fundamental research questions will be examined. First, the issue of how individuals view the operation of evidence-based management and best practices, and the extent to which those perceptions have changed over time will be examined. And second, the extent to which decision-makers perceive evidence sources as providing given types of knowledge will be examined.26

§3.1: Healthcare as a Setting

Healthcare provides a perfect setting for the study of these concerns. First, there is “plenty of evidence that a research practice gap also exists in health care policy and management (Lomas 1997), and that the problems of overuse, underuse, and misuse” also apply to health care management (Walshe and Rundall 2001). Second, healthcare represents a form of complex service organization in which uncertainty is present (Plsek and Greenhalgh 2001, Plsek 2003), and failure is never desired, though highly likely (Edmondson 2010). Third, health services and hospitals compose a knowledge-intensive, knowledge-centered industry, in which speed of change and expertise play critical roles (Brint 2001). Fourth, in the delivery of healthcare, “complexity is reflected in the number, variety, and fragmentation of producers involved,” including mutually interactive, dynamic and non-linear relationships between system parts (Begun, Zimmerman and Dooley 2003). Moreover, decision-making in this domain is “quasi-scientific, in a particular sense: competent decision-making requires scientific knowledge, but scientific knowledge is not sufficient to make decisions” (Turner 2004). Finally, while medicine operates in an “environment of fairly high validity,” where validity refers to the stability of relationships between “objectively identifiable cues and subsequent events or between cues and the outcomes of possible actions,” (Kahneman and Klein 2009), the management of health care, like management in general, is more likely operating in a low validity environment.

It is important to note that, just as health care management is not the provision of health care, evidence-based management in healthcare is not evidence-based medicine. Medicine is a highly-professionalized field, replete with strong social norms and governing bodies, whereas management is not (Rousseau 2006); in general, management has no “common body of knowledge, no system of certification, no commitment to use knowledge for the public good, and no code of ethics” (Khurana et al 2005).27 Managers often operate “on the basis of ‘folk

---

26 In my dissertation prospectus, I take the argument in a slightly different direction, building a 2x2 table of decision-making team convergence/divergence of mental models and a narrowly/broadly diverse implementation of the evidence-based management technology. The requisite variety foundation of the argument remains the same.

27 Some fields have specialized professional organizations to provide education and advocacy services to members. For instance, the American College of Health Executives (ACHE) has a certification
theories or conventional wisdom” (Mintzberg 1973, quoted in Aldrich and Pfeffer 1976). Where medicine is a practice based predominantly on the physical sciences, management rests fundamentally on the social sciences, especially psychology and sociology. In fact, it has been argued that management is a black box in which there is “no settled idea of what constitutes management effectiveness” (Learmonth and Harding 2005). As Walshe and Rundall (2001) note:

“Overall, the tightly defined, well-organized, highly quantitative, and relatively generalizable research base for many clinical professions provides a strong and secure foundation for evidence-based practice and lends itself to a systematic process of review and synthesis and to the production of guidelines and protocols. In contrast, the loosely defined, methodologically heterogeneous, widely distributed, and hard-to-generalize research base for health care management is much more difficult to use in the same way.”

On the one hand, Pfeffer and Sutton (2006), argue that they “believe that managers (like doctors) can practice their craft more effectively if they are routinely guided by the best logic and evidence – and if they relentlessly seek new knowledge and insight, from both inside and outside their companies, to keep updating their assumptions, knowledge, and skills.” On the other hand, Learmonth and Harding (2005) argue: nevertheless, “the basic doctrine of evidence-based management remains one appropriated from evidence-based health care: that a consideration of evidence will increase the rationality and thus the effectiveness of managers’ decisions.”

§3.2: Scope of the Project

The application portion of this essay is limited to strategic and high-level operational decisions in healthcare management, and particularly within the hospital setting.

Strategic concerns in healthcare include the organization, structuring and financing of care. These concerns include, but are not limited to positioning the organization to respond to one or more challenges; prioritizing, investing and leading strategies with limited resources; and implementing strategies in time to meet change drivers (Shortell and Oxendine 2008). Examples of strategic and high-level operational questions facing the healthcare setting include:

exam covering ten knowledge areas (e.g., finance, regulatory frameworks, management, ethics, and health information technology), as well as a fellowship category of membership. However, this should not negate the general point about management versus medicine.

They proceed to argue that, in this way, evidence-based management “tacitly reinforce[s] and legitimize[s] managerialism.”
What are the barriers and facilitators to the implementation of an electronic medical record?
How can needlestick injuries in health care setting be prevented?
What managerial interventions can improve physicians’ and nurses’ care of their diabetic patients?
What are the effects of telemedicine on health care costs, outcomes, and patient satisfaction?

Most questions from the survey portion of the research are couched in terms of quality, safety and risk, while findings from the interviews cover both these topics and healthcare management more broadly. Focusing on quality, safety and risk might naturally favor the concept of knowledge variety, which itself is most relevant in conditions requiring organizational resilience. However, this is tempered by the larger discussion of knowledge in a social context, as discussed above.

In general, the findings should be interpreted carefully for use outside of this setting, though findings broadly indicative of empirical regularities can be analogized to the particularities of other settings and sectors.

§3.3: MATCH Project Description and Findings

To promote evidence-based management in health care and bridge the “gap between research [e.g., peer-reviewed, journal-based] evidence and organizational decisionmaking,” the Agency for Healthcare Research and Quality (AHRQ) funded a study called the Methods for Developing Actionable Evidence for Consumers of Health Services Research (MATCH) Study (AHRQ 2008). The results of this study were published and presented in several venues, including Rundall et al (2007). I participated in this study as a Graduate Student Researcher (GSR), collected and analyzed data, reported results, and began to develop a perspective on the role and operation of evidence-based management.

The MATCH team convened four dynamic, semi-structured, full-day discussion groups with a purposive sample of 31 senior healthcare managers, representing a range of organizations distributed nationally. These peer-to-peer discussion groups were asked questions about high-priority strategic decisions, current levels of research evidence use, organizational barriers and facilitators to evidence use, and what would help them make better informed decisions in the future (Rundall et al 2007). Results suggested that decision-makers rarely used research evidence, desired the “best available evidence,” had broad definitions of acceptable evidence, and desired tools to help them gather and use evidence (AHRQ 2008). Their definitions of evidence differed from traditional definitions as understood by health services researchers, and emphasized colloquial, experience-
based, and tacit knowledge. Decision-makers also reported that their trust in any
given piece of evidence was increased to the extent that it was reviewed, critiqued,
and ultimately approved (or “vetted”) by a respected source, and reported concerns
about engaging in collaborations to collect new evidence. Distilling the qualitative
data (see Table 6), the team argued that evidence should follow the Four A’s –
namely, be it should be accessible, accurate, actionable, and applicable. In a
summation of the research, one participant commented that, in his opinion, “what
people really want is not the evidence, they want the answer.”

<table>
<thead>
<tr>
<th>Table 6: Sample of MATCH comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>“For the vast majority of questions I have, I have found the answers are not out there. Research is not being done in a way that answers my questions.”</td>
</tr>
<tr>
<td>“It’s amazing how much information really does exist and it never gets digested and put into the hands of people that can find meaning in it.”</td>
</tr>
<tr>
<td>“There’s an issue of transparency [both literature and knowledge broker recommendations] - both in the evidence and in the way it is gathered. Most people don’t have the resources to do this on their own.”</td>
</tr>
<tr>
<td>“We want a deeper understanding of the sources of evidence and vetting that evidence. Management research is more empirical. Can we get evidence in a timeframe that is actionable?”</td>
</tr>
<tr>
<td>“If you don’t make evidence available, easily available, in the face of the person, at the moment that they need it, they don’t get it. They’re too rushed.”</td>
</tr>
</tbody>
</table>

§3.3.1: MATCH Toolbox Dissemination

Together with journal articles and conference presentations, a series of tools
for individual practitioners seeking support on evidence-based management was
produced to disseminate the research findings. The tools were compiled in the
Informed Decisions Toolbox (IDT) and made partially available in journal and book
publications, and in full on the Internet at http://toolbox.berkeley.edu (see Figure 7
for a screenshot). At the time of writing, the IDT website was still accessible online.
Among other motivations, the toolbox was created to provide a “vetted” and “legitimate” source of support for healthcare decision-makers attempting to engage in evidence-based management processes. It outlines six steps to assist the health care decision-maker to:

1. “Acquire the best available evidence when making management decisions;
2. Assess whether evidence is useful, defined as accurate, applicable, actionable, and accessible; [and]
3. Improve the process by which evidence is used in decision-making.”

The six steps are:
1. Framing the Question Behind the Decision
2. Finding Sources of Information
3. Assessing the Accuracy of Information (Evaluating the Evidence)
4. Assessing the Applicability of Information (Evaluating the Evidence)
5. Assessing the Actionability of Information (Evaluating the Evidence)
6. Determining if the Information is Adequate

A review of site traffic statistics suggests that access to the online version of the IDT has remained constant and very low throughout its history. Because the site is hosted on the same server as another UC-Berkeley site, several IDT-specific statistics, such as hits on the homepage, are not easily retrievable. However, it was possible to compile page hits and referrals by month for certain sub-sites (see Figure Group 8 below). The largest number of referrals to the Toolbox were received from AHRQ sites, including <http://www.ahrq.gov/research/decisiontoolbx.htm> and <http://innovations.ahrq.gov/disclaimer.aspx>; at its maximum, there were 54 referrals from these combined sites. On average, the most viewed IDT sub-site was the “/tools” page which contained the index for the six outlined steps noted above; this site had an average 79 hits per month and a maximum of approximately 160 hits, achieved on two occasions. Over 37 monthly periods, there was an aggregate count of 2,927 non-unique visits to the “/tools” page. This amount is miniscule in Internet traffic terms; by comparison, consider that in December 2011 alone, the AHRQ.gov website received 225,191 unique visitors to its homepage (citation: http://siteanalytics.compete.com/ahrq.gov/).
The lack of traffic on the IDT page does not necessarily suggest that it hasn’t made an impact on the practice of evidence-based management in hospital settings. To make a better determination, it would be useful to know who has accessed the site, from which organizations and how many times. Moreover, it would be useful to ask those organizations the way in which the IDT was marshaled in the organization to effect change. Unfortunately, following this line of inquiry is impossible from the available data. Instead, based on the absolute traffic to the site, it is likely that the online toolbox was ineffectual as a dissemination mechanism to practitioners.

§3.3.2: MATCH as Motivation and Comparison

At the end of data collection, I wrote a letter to the Harvard Business Review reporting these preliminary results of the project:

“Our results show that leaders, particularly in the public sector, seldom use management research evidence when making decisions and that they prefer externally vetted, concise, on-hand evidence rather than evidence that meets classical standards of rigor. Our key finding, therefore, is that useful evidence
must follow the four A’s: It must be accessible, accurate, actionable, and applicable” (Martelli 2006). In response, Pfeffer and Sutton, authors of the recent Harvard Business Review cover article on the evidence-based management, responded, saying that “We agree completely with the four guidelines Peter Martelli proposes for making evidence-based management a way of life...”

However, as time elapsed from the project, several issues remained problematic to me, particularly underscored by three comments captured during the focus groups (see Table 7). These comments suggested that the practitioners’ concept of evidence were inconsistent with our current models of evidence-based management, and that solutions promoting the use of research evidence would be insufficient to meet the challenges they mentioned. Our notion of the definition of evidence, where, when and how it arrives to the decision-maker, and the role it plays in determining organizational decisions seemed to need a deeper exploration. In Rundall et al (2007), we had argued, explicitly citing Simon and bounded rationality, that it was critical for rationality to push back in the evidence-based process. Yet, for the reasons now profiled above, it nonetheless seemed like evidence-based management was missing key aspects of the logic of appropriateness, which would require a reconceptualization of the process.

Table 7: MATCH Problematic Comments (emphasis mine)

| “We use evidence in the deployment of key strategic decisions. We'll use books, vendors, best practices, to the extent that there's literature available. Not sure if it qualifies as evidence, but then, we haven't defined what evidence is.” |
| “A result with 70% confidence now is better than one with 95% in three years.” |
| “What people really want is not the evidence, they want the answer.” |

This became the motivation to continue the study of evidence-based management through an updated project. Because the MATCH study concluded five years before this project, it would be possible to review and compare many of the results, while also addressing the new topics around knowledge variety. By design, the setting, respondents and scope are roughly the same – high-level healthcare decision-makers describing the use of evidence in strategic decisions, especially around innovation.

§3.4: BEACON-Berkeley Project

In order to further understand the role of evidence-based management, a new project was undertaken. The project was broadly designed to capture multiple aspects of the structure and process of evidence by decision-makers in hospitals. The intention was to examine the extent to which decision-makers use a knowledge
variety approach, the factors immediately associated with it, and what it entails in their setting. To that end, the project was not meant to be confirmatory, but exploratory.

As such, a mixed-methods study was designed to further explore the use of evidence-based management and best practices. The first stage of the study consisted of a theory-generated survey designed to capture general relationships, and to indicate individual sites where further qualitative exploration would be interesting. These sites would be chosen based on the site-specific aggregate measure of individual knowledge variety, with high- and low-average sites chosen. Subsequently, a series of interviews with the top management on the use of knowledge in decision-making was conducted at the chosen sites and exploratory reflections captured.

In the late summer of 2009, Bruce Spurlock, MD, director of the BEACON Collaborative, contacted me to discuss the possibility of conducting an evaluation of Collaborative’s value to its member organizations. This Collaborative, which functioned as a knowledge intermediary transmitting best practices to participants, offered a valuable background setting to examine the topic of knowledge variety. Over the subsequent year, a shared vision emerged to examine the perceptions of evidence-based management and best practices, the epistemic qualities of evidence desired for strategic decisions, and the character of BEACON as a valued source of knowledge.

The project design and implementation period ran from August 2010 until August 2011, with analyses and interpretation occurring over the following year.

§4.0: Methods and Findings

This study was designed to build on the findings of the Methods for Developing Actionable Evidence for Consumers of Health Services Research (MATCH) Study, and to expand the frontiers of theorizing on the role and function of evidence-based management in the healthcare context. It employed a mixed-methods model consisting of two integrated portions – a principally quantitative survey portion followed by a qualitative interview component. This section will present the details, justification and limitations of the research design and data analysis for each of these components.

Due to operational constraints on the project, it was not possible to run the project as a longitudinal study; thus, the design of the study is cross-sectional. However, paired with and compared to the MATCH study, it takes on the quality of a repeated cross-section for a subset of questions. If there were a long-enough time
horizon, it would be possible to capture some of the consequences of using different representations of evidence-based management as a management approach over time (for example, in the form of avoiding Type III errors). Because the operational outcomes remain unobserved, a series of “proximal” outcomes of knowledge variety related to decision-making confidence are also explored.

The data presented here are suggestive of general patterns, and serve an exploratory function for future studies.

§4.0.1: BEACON-Berkeley Project Funding

Funding for the project was secured under an independent contractor agreement with the Center for Quality Systems Improvement (CQSI). CQSI is a not-for-profit organization based in Roseville, CA under the executive direction of Bruce Spurlock, MD, and is the administrative home for BEACON Collaborative grant funding provided by the Betty Irene Moore Nursing Initiative of the Gordon and Betty Moore Foundation.

Although members of the BEACON staff provided support in design, recruitment and reporting, the contract expressly stipulated that the study team has ultimate control over design and implementation of data collection and analysis in this project, and that no member of BEACON or of the Center for Quality Systems Improvement will have access to individually-identifiable data collected in the study.

§4.0.2: Institutional Review Board

All study details were submitted to the UC-Berkeley Institutional Review Board (IRB) at the Center for the Protection of Human Subjects (CPHS) in advance of any procedures beginning. The protocol was filed under the title “On Evidence-based Management” under the supervision of Principal Investigator Stephen M. Shortell. The file was approved under the Protocol Number 2010-03-984 on November 9, 2010, and further approved one year later under a continuing review.

---

29 An initial consulting contract was provided by CQSI and was reviewed by a legal adviser at the UC-Berkeley School of Public Health for issues related to project control and intellectual property. Further review clarified terms related to researcher rights and responsibilities, and the operational budget (software, transportation, etc) for the project. A condition of funding was that the investigator remains enrolled as a graduate student at UC-Berkeley throughout the project period, or until graduation. As such, most of the funding provided was earmarked to reimburse enrollment fees.
until the expected formal closure of the project in November 2012. A copy of the file is available in the CPHS offices.³⁰

§4.1: Empirical Setting: Details, justification and limitations

The empirical setting for this study is the BEACON Collaborative and its 42 member hospitals throughout the San Francisco Bay Area.

The BEACON Collaborative was formed in June 2005 as the Bay Area Patient Safety Collaborative of the Hospital Council of Northern and Central California in order to support the local implementation of the Institute for Healthcare Improvement’s (IHI) 100,000 Lives campaign. The Collaborative persisted after the campaign’s formal completion through two rounds of infrastructure and project funding from the Gordon and Betty Moore Foundation totaling nearly eight million dollars. With this additional funding, BEACON quickly developed from a communications and support hub for hospitals into a quality improvement organization focusing on developing, disseminating and supporting the data measurement and collection methodologies of participating hospitals. By 2010, the Collaborative had significant independence from the Hospital Council, pursuing a full range of patient safety and quality improvement activities, including technical workshops, site consultations, and quarterly and annual meetings for its members.³¹ In this period, it described its mission on its website in the following way:

---
³⁰ Two cycles of comment and feedback took place between the CPHS reviewers and the investigators to clarify sampling, data collection and other issues. Because of the nature of the funding source for this project, a potential conflict of interest was reported and examined. At the time of submission, Bruce Spurlock was both the Executive Director of the Center for Quality Systems Improvement and the Executive Director of the BEACON Collaborative, and the BEACON Collaborative was listed as a formal institutional collaborator on the CPHS submission form. Following protocol, a separate submission was made to and reviewed by the Conflict of Interest Committee, and forwarded to CPHS before their final review and approval. The Committee reported that participation “does not represent a significant financial conflict under the current Berkeley policy.”

³¹ Throughout the project period, there appeared to be significant and increasing disagreement between the staff of the BEACON Collaborative and partners at the Hospital Council of Northern and Central California. For example, during the recruitment phase, the Hospital Council was reluctant to share member contact information directly with the BEACON Collaborative, presumably, I speculate, over concern that this list would provide BEACON with a means of reducing dependence upon the Hospital Council as a partner and increasing its role as an independent consultancy. After an extended dialogue across the research collaborators, including a description of the responsibilities of the investigators under the IRB approval, the Hospital Council agreed to prepare the expurgated survey panel to BEACON for use in the study. While interesting in its own right, the complicated relationship between BEACON and the Hospital Council fell outside of the scope of this project and the protocol allowable under IRB approval. Furthermore, inquiring on this issue was unwelcome on both sides, and would have created bad
“To improve the quality of acute healthcare and end inadvertent harm to patients by accelerating the implementation of high impact, evidence-based clinical and operational practices in hospitals utilizing a dedicated peer-to-peer learning community that produces credible clinical outcome data and models for success.”

The Collaborative considers itself to be a “peer-to-peer learning community,” featuring in-person meetings and the sharing of best practices. It might also be called an intermediary “knowledge broker” (Shortell 2006; see Figure 9 below) or a type of “evidence-based management co-operative” (Kovner, Elton and Billings 2000). While it shares some commonalities with a community of practice, it should only be considered one by analogy. Communities of practice are groups of people who “share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis” (Wenger, McDermott and Snyder 2002). In this sense, BEACON is a peer-to-peer community focused on the development of common approaches to management practice. Brown and Duguid (1998) describe a community of practice as a “group across which such know-how and sensemaking are shared – the group which needs to work together for its dispositional know-how to be put into practice,” in which individual knowledge comes together “like the parts of individual performers to a complete musical score.” Generally, BEACON is more structured, intermittent and designed (as opposed to emergent) than true communities of practice, which include closely-shared tasks, mutually defining identities and the “absence of introductory preambles” (Wenger 1999). Furthermore, while the Collaborative has a stable set of organizational members, individual attendees change from meeting to meeting. Thus, the analogy to a practice community should be taken broadly.

------------------------

blood amongst the parties and made engaging in the research difficult. Eventually, the rifts between the BEACON Collaborative and the Hospital Council led to a separation between the organizations – formal affiliation ended in July 2011, corresponding with the end of the final grant period from the Moore Foundation. The BEACON name remains registered under the Hospital Council, and as of writing, the BEACON website has moved under the Hospital Council website, serving as the index page for their Quality Improvement and Patient Safety section. The outgoing BEACON staff applied for, and were awarded with a subsequent grant from the Moore foundation, and started a new organization named Cynosure. The goals, structures and processes of Cynosure appear to be virtually indistinguishable from the previous incarnation of BEACON, with of course, the caveat that Cynosure is not affiliated with the Hospital Council. The staff of the Hospital Council continues to run BEACON activities similar to those previously conducted, including a 2012 annual meeting taking place shortly before the Cynosure annual meeting. A future investigation into the causes of the separation would prove a valuable addition to the literature on implementation and collaboration.
Organizational membership in the Collaborative was free and entitled individuals in these organizations to products and services (both free and fee-based), such as in-person and WebEx conferences and workshops, site consultations and clinical support, which were intended to promote education and the sharing of best practices related to the reduction of hospital patient safety errors. Example workshops included “Excel for Quality Improvement: Basics for Beginners,” which focused on data manipulation in Excel; “Practical Skills for Quality Improvement,” which focused on preparing flow charts and plan-do-study-act (PDSA) cycles; and, “Project Based Team Leader Training,” which addressed team leadership, the science for improvement, measurement for improvement, and communication and sharing results.

Upon joining BEACON, all member hospitals assented to participation in two Collaborative-wide projects and two additional projects of their choosing. During the project period, patient falls and sepsis were the system wide projects, and the additional projects were hospital acquired pressure ulcers (HAPU), stroke mortality, acute myocardial infarction mortality, C. difficile associated disease, catheter associated urinary tract infections, and perinatal improvement by reducing birth trauma. Together with the required projects, a non-binding contract listing the hospital’s desired performance improvement goal (e.g., reduce total falls by 20%) was approved by the institutional signatory official. These contracts and individual hospital performance towards them were maintained confidential and given as feedback in BEACON organization-specific reports.

The BEACON Collaborative was chosen as an empirical site for three key reasons. First, its role as a knowledge intermediary offered an excellent venue to explore the broad character of knowledge, where organizational respondents could cite knowledge from their own resources, their comparison hospitals, and the
Collaborative, in tandem with general literature or best practice sources. To that end, the hospital sites could demonstrate the emergence of knowledge variety at an organizational level. Second, its focus on patient safety and quality improvement was valuable to couch the findings of strategic management and innovation in terms of robustness and resilience. Patient safety and quality improvement are areas that involve technical and sociobehavioral systems across the hospital; and, despite a redoubled focus over the past decade, systems lag behind expectations and serious errors still accrue (Wachter 2010, Clancy 2010, Dickey, Corrigan and Denham 2010). The Collaborative facilitated the dissemination of technological and organizational approaches to improving quality and safety, including models of safety from other domains (e.g., human factors engineering) and longer-term, large-scale or resource-intensive models (e.g., developing a “just culture”). This afforded the opportunity to observe how hospitals draw on knowledge in order to reduce the uncertainty around new courses of action or resource commitments in these areas. Furthermore, because all BEACON members participated in at least two common projects, and the Collaborative brokered best practices to these sites related to these projects, it served as a setting to explore best practices, attention and customization. Finally, as a practical concern, the Collaborative was willing to examine these questions and to broadly support the study.

![BEACON Bay Area Hospitals](source: BEACON website)
General data on hospital type, location, size, certification status and performance with respect to quality and safety were collected from publicly available sources and organized into flat files in Excel 2010, then imported into Stata 11 for additional analysis. Data sources included individual hospital websites, the Hospital Council of Northern and Central California, the Health Facilities Consumer Information System of the California Department of Public Health, the Centers for Medicare and Medicaid Services, the Leapfrog Group Hospital Survey, HealthGrades Hospital Database, and the Joint Commission QualityCheck Database. The BEACON Collaborative also provided data comparing member hospitals to all California hospitals using the California Office of Statewide Health Planning and Development (OSHPD); Collaborative hospitals were larger by bed size, were more likely to be teaching hospitals and to belong to a system of three or more, and were less likely to be Disproportionate Share Hospitals than other hospitals in California. These data were used to provide contextual perspective on the hospitals included in the sample. For instance, interview sites, as noted later, were selected partially based on the feasibility of comparison by bed size and services provided, and on distinctions of each site’s history of quality and safety, as supported by formal complaints, awards or citations over the past three to seven years.

§4.2: Operationalization of Variables

Variety:

Knowledge variety is examined at two levels of analysis. First, an individual-level construct is defined as the maintenance of a roughly equal concentration of each of five epistemic knowledge types. Although the term “roughly” suggests imprecision, it is intended to leave leeway for fluctuations in the manner and timing of new evidence being noticed and encoded by the decision-maker; the main point is not a quantitative exactness of balance, but that a given type of knowledge does not play an exclusive role in rhetorical persuasion. Thus, an equal share of five knowledge types would yield approximately one-fifth per share in decision-making. A concentration measure, such as the Herfindahl-Hirschman Index (HHI), can be used to calculate the level of equality or monopoly across the knowledge shares. A mention by a respondent that an element within a given knowledge type was, or should be, considered as evidence towards an organizational decision counts as an admissible instance within that knowledge type. Evidence, and therefore knowledge for decisions, is compiled through a variety of sources. Table 8 displays a range of the sources that convey knowledge.
Table 8: Suggested Instances of Evidence

<table>
<thead>
<tr>
<th>Theoretical or mechanism data</th>
<th>Marketing/opinion poll data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility/implementation data</td>
<td>Surveillance data on trends over time</td>
</tr>
<tr>
<td>Contextual information</td>
<td>Cost and economic data</td>
</tr>
<tr>
<td>Intended primary outcome evidence</td>
<td>Qualitative data</td>
</tr>
<tr>
<td>Unintended or unanticipated outcome results</td>
<td>Local data</td>
</tr>
<tr>
<td>Process of quality-of-care results</td>
<td>Systematic review data</td>
</tr>
<tr>
<td>“Outcome” or clinical data</td>
<td>Simulation data on project impact</td>
</tr>
<tr>
<td>Quality-of-life data and adverse consequences data</td>
<td>Internal validity evidence</td>
</tr>
<tr>
<td>Quality improvement data</td>
<td>External validity evidence</td>
</tr>
</tbody>
</table>

Excerpted from Glasgow and Emmons (2007)

Second, because the outcomes are evinced through interpersonal knowledge structures, there is also an interpersonal conceptualization for knowledge variety. While individuals engage knowledge variety in the compilation of evidence, the created knowledge structures reside and operate through sensemaking between individual decision-makers. Thus, in order to capture the effects of knowledge variety, both the individual and emergent group/organizational levels require observation. An aggregated measure across decision-makers in an organization is a simple average of the knowledge variety concentration measure for those decision-makers. This is a simplistic measure; it risks conflating concentration driven by type A in one person with concentration driven by type B in a second person. Moreover, in larger samples, the aggregation risks regression to the mean, displaying a false balance of types where none exists. Nonetheless, in smaller samples, it is indicative of the general level of concentration across decision-makers, and is a reasonable approach when paired with individual-level data and qualitative findings of organizational attention. Given the ipsative nature of the nature of the epistemic knowledge item, interpersonal comparison by knowledge type is inappropriate, as will be shown below.

Best Practices and Customization:

This study considers a best practice to be an externally derived set of structural and procedural conditions that are widely-acknowledged to be associated with beneficial outcomes. In the case of this study, patient safety practices, and specifically those related to sepsis, explicitly presented by the Beacon Collaborative are considered best practices.

Customization is defined as deviation from the codified characteristics of a practice adopted from an external source. This can be measured in terms of percent difference; for instance, if a best practice undergoing diffusion is composed N structures or practices, and k are removed or essentially modified by the organization during adoption, then the practice is said to be k/N percent customized.
(or conformed to at $1 - \frac{k}{N}$ percent). Where explicitly codified practices or the discrete levels of conformity are unavailable, patterns of deviation from a suggested practice can also be determined through observation of practices and analyses of organizational artifacts.

Confidence:

Confidence is a useful intermediate outcome because of the role it plays in demonstrating an individual’s belief that a set of evidence supports his or her view of the world, and because of the attendant socio-behavioral factors in interpersonal rhetorical persuasion. Incompleteness is related to decision-maker overconfidence. Mattesich (1978) points this out with respect to formal inference, noting that the “confidence we have in a hypothesis is dependent on the evidence with which it is associated, hence it cannot be asserted independently from the strength of the entire inductive argument (evidence plus hypothesis).” And from the perspective of uncertainty reduction, Quiggin (2009) notes that “[e]stimates of project outcomes derived from formal models of choice under uncertainty are inherently incomplete. Incomplete estimates will generally be over-optimistic. The errors will be greater, the less well understood is the problem in question.” To the extent that variety helps increase perspectives, it should also be reflected in the level of confidence displayed across members of the decision-making group. At the organizational level, Mitroff (2009) reports, in a long-term study of crisis management preparedness, that organizations that perceived themselves to be well-prepared for crisis were generally objectively under-prepared, and vice-versa, given retrospective analysis.

In the survey, confidence is measured through a 100-point scale ranging from zero (not at all confident) to 100 (fully confident), with the halfway-mark labeled as “don’t have strong feelings either way”. The items capture an individual’s confidence in the capacity for an evidence-based management approach to lead to various outcomes including decision efficiency, organizational legitimacy, and potential for avoiding unintended consequences. A rough measure of organization-level confidence can be constructed by creating a composite of respondent values within each organization.

Mental Models:

This study was unable to collect explicit measures of sensemaking or the development of shared mental models. However, partial justification for the level of cognitive diversity in knowledge structures for decision-making was captured through cognitive mapping procedures. Schwenk (1995) notes five approaches to mind mapping proposed in Huff (1990). These are: 1) maps showing attention and association, which count the frequency, clustering and changes in word use in documents; 2) maps showing cognitive taxonomies, which follow the hierarchical
relationships between broad categories and subcategories; 3) maps showing causality and system dynamics, or “causal maps,” which capture individuals’ beliefs about how cognitive elements affect each other; 4) maps showing the structure of arguments, assessed through rhetoric trees; and 5) maps showing schemas, cognitive frames and perceptual codes, which assess expectations by parsing the underlying structure of language. Interview data collected from individuals at exemplar hospitals during the qualitative portion of this study were examined for cognitive taxonomies of knowledge and implicit framing of evidence-based management processes.

Two explicit measures are available to measure shared mental models or cognitive diversity. The interpretive ambiguity construct is defined as “lack of clarity within the team concerning the degree to which team members share common attributions concerning, for example, organizational success and failure” (Kilduff, Angelmar and Mehra 2000). In the Kilduff et al study, it is operationalized by asking participants how much agreement there is within their team, and is anchored at “everybody stuck to a different explanation” and at “all members agreed.” Perfect consensus marked low interpretive ambiguity, and vice versa. The knowledge heterogeneity construct is defined as “the variety of knowledge, know-how and expertise to which a manager has access through her network” (Rodan and Galunic 2004), and is operationalized as a network metric measuring the uniqueness of knowledge between an ego and its contacts. It is constructed from two four-point Likert scale questions asking how well an ego’s alters know one another (“especially close” to “distant”) and how similar knowledge is between alters (“very similar” to “very distant”). Using these data, a knowledge distance matrix, $D$, is generated, in which “$D_{ij}$ is the reported dissimilarity between person $i$’s knowledge and person $j$’s” (Ibid.). The measure $D_{ij}$ increases linearly with the distance between contacts. The knowledge heterogeneity argument focuses on access through social connections to individuals, and not yet through connections to legitimate organizations or other sources. Rodan and Galunic (2004) acknowledge this point, noting: “it would be useful if future studies contrast the influence on managerial innovativeness of social networks versus relatively inanimate sources of diverse knowledge.”

Control Measures:

Individual-level controls include general demographic features. Rodan and Galunic (2004) suggest controlling for gender, tenure, level of education, and science or math education. Cascio (2007) citing a survey by Shapiro, Kirkman and Courtney (2007), suggests controlling for time since last degree, given that managers may summarily dismiss some research as “too academic” for practical application. Group- and organization-level controls include the length of time that the decision-making team has worked together and level of correspondence between individual
and organizational styles of decision-making. Organizational culture plays an important part in knowledge transfer and use; this study captures the perceived fit of an individual with his or her organizational culture, but was not able to obtain an overall measure of organizational culture.\(^{32}\)

Three other control variables were considered for inclusion, but were not collected due to the scope of the project data collection. Absorptive capacity is defined as a “set of organizational routines and processes by which firms acquire, assimilate, transform, and exploit knowledge to produce a dynamic organizational capability (Zahra and George 2002).” Need for cognition is a scale that “refers to dispositional differences in cognitive motivation, as it is thought to reflect stable individual differences in the intrinsic motivation to engage in and enjoy effortful cognitive endeavors (Cacioppo et al 1996).” Tolerance for ambiguity refers to the “way an individual (or group) perceives and processes information about ambiguous situations or stimuli when confronted by an array of unfamiliar, complex, or incongruent clues” (Furnham and Ribchester 1995).

§4.3: Survey Overview: Details, justification and limitations

In order to assess the range of thoughts and opinions likely to influence organizational strategy formation, upper management at each BEACON-participating hospital were sampled. Top level management plays a large role in developing and directing strategic plans. Citing studies by Hambrick and colleagues, Ocasio (1997) writes that the “most critical players in attention regulation are typically the CEO and the top management group.” Moreover, top managers are seen as a “symbolic source of institutional leadership, and their actions and dynamics are watched and modeled carefully by others” (Ancona and Nadler 1989). Although not a perfect proxy for the organization, the influence of these individuals in providing institutional legitimacy and resource support for strategic innovations presented the best opportunity to explore the nature of variety in these settings. The top management team is defined as “all executives above the vice-president level (e.g., senior vice president, vice chairman, chief executive officer),” an operationalization intended to include “all major line and staff executives, encompassing all the key activities of the firm at the senior-most level” (Hambrick, Cho and Chen 1996).

\(^{32}\) For instance, the Competing Values Framework is a survey to classify organizational cultures (Quinn and Kimberly 1984). In this framework, organizations tend to track into one of four cultural types: 1) group – based on affiliation and teamwork; 2) developmental – based on innovation and flexibility; 3) hierarchical – based on rules and regulations; or 4) rational – based on efficiency and achievement. A corresponding 20-item instrument is available to measure this characteristic of organizations (Zammuto and Krakower 1991, Helfrich et al 2007).
Design of the survey took place in the spring and summer of 2010, using theoretical concepts derived from a wide range of fields, including previous studies of the translation of research into practice, organizational behavior, and epistemology. Although an effort was made to include existing, validated subscales, such as the tolerance for ambiguity scale, the strict expectation on survey length led to all of these being excised. Instead, items were largely developed from scratch based on survey methodology standards, with an attempt to design factors based on theoretical parameters and to incorporate multiplex questions which could be joined to one or more questions to support multiple hypotheses.

The survey was pretested over the summer of 2010 by selected top administrative staff at two hospitals and by members of the BEACON Collaborative itself. One hospital was within the catchment area of BEACON, but was not a member of the Collaborative; the other was informally affiliated with BEACON, but was located in southern California. Given the exploratory nature of the survey, the purpose of the pretest was to solicit thoughtful comments on the design and content, focusing on site access, survey length, item clarity, and ordering, and data output quality, rather than to aggregate data for factor analysis or other statistical testing. Initial piloting suggested that the survey was functional, clearly worded and ordered, and that completion would take between 15-20 minutes. Other comments were incorporated into the survey design before survey roll-out.

§4.3.1: Survey Sampling

A request was made to partners at BEACON for contact information for all upper-level management of the following job titles at each affiliated hospital and regional network:

- Top management in the following titles:
  - Chief Administrative Officer (CAO)
  - Chief Executive Officer (CEO)
  - Chief Financial Officer (CFO)
  - Chief Information/Technology Officer (CIO/CTO)
  - Chief Medical Officer (CMO)
  - Chief Nursing Officer (CNO)
  - Chief Operating Officer (COO)
  - Senior Legal Counsel

- Highest level title available (eg, Chief, Senior Vice President, Vice President or Director) in the following areas:
  - Compliance
  - Diversity/Ethics
  - Human Resources
  - Knowledge/Learning
- Quality
- Patient Safety
- Marketing
- Risk Management
- Strategic Planning

- Other upper management staff positions not already covered that are involved in patient safety or quality decisions, at the discretion of the contact information providers (eg, Chief Process Officer).

The BEACON Collaborative contacted institutional colleagues at The Hospital Council of Northern and Central California to provide the contact information. The Hospital Council has access both to internal membership lists as well as to the California Hospital Association’s comprehensive Membership Directory of over 6,500 “key players”, including all hospital top management in California. Based on the above criteria, a list of 410 individuals was developed and transferred for use in the project.

Due to data sharing limitations between BEACON and the Hospital Council, the contact list provided only the title and email for the sampled individuals. In order to properly describe the sample by hospital, I imputed the additional contact information using a best guess of first and last name from the email handle and organizational affiliation from the email address domain, and cross-referenced the information against available directories and listings, including organizational websites, Hoover’s Online database of company profiles, LinkedIn profiles, professional membership affiliations, and archived copies of conference presentations. Of the drawn sample of 410 email addresses and titles, only 17 individuals were unattributable to a particular brick-and-mortar facility, and all but one of these individuals were affiliated with Kaiser Permanente in various capacities, as evidenced by their properly-formatted address in the domain “@kp.org”. On average, eight people were contacted at each sampled hospital site, with a range from one to 14 individuals within each site.

The sample of 410 individuals represented 42 Bay Area hospitals, and selected regional and corporate offices overseeing these facilities. All hospitals in the sample were general acute care facilities ranging in licensed bed capacity from approximately 50 beds to over 600 beds. A full list of surveyed and responding organizations is provided in the Appendix.

A letter drafted by the investigators and jointly-signed by the BEACON Collaborative and the Hospital Council of Northern and Central California was sent via email in mid-November 2010 to all 410 sampled individuals. The letter described the nature and purpose of the study and introduced the investigators to the recipient.
§4.3.2: Survey Procedures

The survey was an internet-based questionnaire created using the Qualtrics survey tool, a free subscription to which was available through the School of Public Health at UC-Berkeley. A choice was made to use an internet-based questionnaire, rather than the more traditional mailed paper survey not only because of the low cost, but also because of advances in the technology of internet-based surveys and the general familiarity of users with these feedback mechanisms in recent years. An original draft of the survey was coded using a professional subscription to SurveyMonkey, another online questionnaire site – however, Qualtrics offered design, distribution and data management functionality that exceeded the capacity of SurveyMonkey. Among the useful features was the ability to scale items as constant-sum using a sliding-bar interface, and a well-designed survey panel function that allowed easy uploading and manipulation of individual information for the 410 sampled individuals into a recipient list.

One week after the introductory contact by BEACON, an email directly from the investigator was sent to each sampled individual explaining the study and inviting the recipient to participate in the study survey by clicking on a link provided. The email was sent through the Qualtrics survey panel distribution function, which allowed scheduled automated emailing, tracking of responses by individual, and reminder emails based on response status.

Before beginning the survey, respondents were presented with an informed consent form and specific instructions. In the specific instructions, the respondent was told that the survey was intended to be completed by the recipient only, and that it should reflect the respondent’s personal opinions, not an institutional perspective. Furthermore, it clarified and primed the respondent on two key points – first, that the survey focuses on decision-making related to strategic management decisions, where these decisions involve determining long-term goals that produce business value, aligning competencies with environmental opportunities and constraints, allocating resources to meet goals, implementing managerial courses of action to meet goals, and evaluating implemented courses of action; and second, that the survey focuses on management decisions, focused on outcomes like productivity, absenteeism and employee or patient satisfaction, and not clinical decisions, focused on outcomes like reduced nosocomial infections or safer care transitions. Finally, it told the respondent that their answers are important from an organizational point of view, and that their response is important even if they think that they don’t have direct influence over decisions like these.
The questionnaire itself was composed of 30 questions divided into eight sections presented over ten pages (see the Appendix for a full copy of the survey instrument), organized in the following order:

1) Informed Consent
2) Demographic questions, such as gender, age range, education and tenure;
3) Decision-making questions, such as which other BEACON organization the respondent admires for its ability to make strategic management decisions, and how often the respondent contacts co-workers when involved in strategic management decisions;
4) Evidence-definition questions, which ask the respondent for his or her personal definition of “evidence-based management” and of “best practices”;
5) Knowledge type questions, such as which knowledge type from a list of five (e.g., know-how) is most compatible with the respondent’s organization;
6) Knowledge source questions, which ask the respondent to list frequently-used sources of evidence and describe their qualities;
7) BEACON-related questions, such as what sources complement or substitute for use of the BEACON Collaborative when making strategic management decisions related to patient safety; and,
8) Outcomes questions, such as the respondent’s confidence that an evidence-based decision will lead to beneficial clinical and organizational outcomes.

Due to the constraints around securing contact information for the sample, the survey was initiated directly before the Thanksgiving holiday period. As a result, the response to the first request was minimal, and many automated vacation (also known as, “out of the office”) email responses were received in this mailing. Three email reminders were sent to survey non-responders, excluding those who had removed themselves from the study using a link provided in every email correspondence. These reminders were sent on 2 December 2010, two weeks after the initial survey invitation, on 16 December 2010, after an additional two weeks had passed, and finally on 13 January 2011, allowing for individuals to return from the holiday break period.

Subsequent to the final email reminder, an effort was made to reach out to executive assistants to the sampled top management individuals at each sampled hospital using personal contacts and professional contacts from BEACON, the Hospital Council and other executive assistants within the same hospital facility. Each contact was explained the details of the survey, the importance of participation, and that both their assistance and the sampled individual’s participation were completely voluntary. Strong care was taken to reach out to individuals without providing names of other participants who may or may not have responded at the time of inquiry. Of the contacts responding, each one was notably
straightforward and helpful, and on the whole, they were able to secure an additional group of responses from top management without compromising the spirit of voluntary participation.

§4.3.3: Survey Responses

From the sample of 410 individuals contacted, 103 survey responses were received, representing a crude response rate of 25.1 percent. As a criterion for further analysis, survey responses in which the respondent completed only the introductory demographic section and answered no other questions were culled from the sample. The demographic section of the survey accounted for 18.37% of the total survey length and contained no substantive questions related to the study hypotheses. A decision was made to retain one respondent in the sample despite an overall survey completion rate less than 18.37%, based on the quality of this respondent’s answers to the non-demographic questions completed.33 A total of ten responses were removed from the sample, yielding an analytic sample of 93 responses. Therefore, the effective response rate was 22.7 percent.34

33 There were a wide range of survey durations to completion, ranging from one minute to over four hours. The survey was installed with time trackers to log the amount of time spent on each page before advancing to the subsequent page in the survey. The principal purpose of these trackers was to gauge whether respondents had spent enough time on the context, priming and instruction pages to suggest that they had read and internalized the conveyed information. Due to a software error, these counters were reset to zero early in the study and no longer measured time as intended. As a result, no inference can be made along the lines noted above. Instead, it is worthwhile to note that over half of the individuals (52 of 103 responses) spent between 10 and 30 minutes logged-on to the survey, within the timeframe suggested by the survey pilot.

Unfortunately, due to the error, this figure only covers overall time spent with survey pages open using that login, and does not imply that individuals completed all – or any – of the questionnaire items. For surveys taking over 30 minutes, and especially those taking over an hour, we might infer that respondents left the survey window open in their browsers while attending to other duties and returned later to attend to the survey. There were 16 individuals who spent less than 10 minutes on the survey. It is most likely that these respondents completed very few of the survey items, and that their answers may have been excluded from the analytic sample already. If these respondents substantially completed the survey, their answers are suspect given the amount of time spent on each survey question – a survey completed in 10 minutes would allow a maximum of 20 seconds per question, including open-ended questions and allowing no time for reading instructions or other framing information. Given the limitations of the time tracker and the inability to correspond surveys to time, this is a moot question and remains an assumption of the analytic sample and a limitation of the survey.

34 Respondents represented a wide range of job titles and organizational levels among the sampled population. Approximately one-fourth of the top management job titles CEO, CMO, CNO, and COO responded to the survey request. Nonetheless, it proved difficult to differentiate between job titles. Often, slight differences in nomenclature across organizations created immediate misclassification – for instance, Vice President in one hospital might have the same operational level as a Chief Officer in another. In the demographic section, a partially close-ended list of titles (with an “other” fill-in option) was offered to respondents; while this helped structure the data, it also created some inadvertent classification problems. For example, one respondent self-identified
After adjustment, thirty-eight of 49 potential organizational classifications (77.6 percent) were represented in the final analytic dataset.\textsuperscript{35} Despite receiving a suitable between-hospital response, the response rate stratified by hospital ranged from 10 to 60 percent of individuals sampled, and in raw numbers yielded no more than one or two respondents in most hospitals.\textsuperscript{36} As a result, the within-hospital response rate was insufficient to allow rigorous quantitative analysis of variety by organization. Given larger numbers, a preliminary multi-level, or hierarchical, model using the variety metric for top management teams within hospitals would have been used. Instead, the compiled variety metrics by facility were used in an exploratory fashion.

It is important to note that low response rate does not necessarily equate to response bias. Although response rates of 60 or 70 percent are often considered typical for publication purposes, there is increasing statistical evidence that low response rates do not yield response bias. In a Pew analysis comparing a high-yield to a low-yield political opinion survey over two time points, in which the low response rate group was in the same range as the rate for the BEACON survey, the survey results produced no significant differences by response rate across 77 of 84 comparable items and small differences on the others (Keeter et al 2006).

\textsuperscript{35} Individuals represented a diverse range of organizational affiliations, including Kaiser-Permanente and Sutter Health network hospitals, hospitals belonging to other for- and not-for-profit networks, public hospitals, and independent hospitals. Within the original sample transferred, several individuals held regional or system-level positions. In most cases, these individuals also held a brick-and-mortar hospital affiliation in a top management position, and this became their principal affiliation for the purposes of this study. However, due to the state of demographic information in the original sample, five people were attributable to a system-level position only; of these five, only one person responded to the survey portion of the research, which posed little problem for analysis. Also, individuals from two Kaiser Permanente-affiliated hospitals were included in the original sample, but were not member to the BEACON Collaborative; none of those surveyed at these hospitals responded to the survey. Analytically, respondents from St. Luke’s Hospital were indistinguishable from other California Pacific Medical Center (CPMC) facilities, so any individuals representing this BEACON member were folded-in to a CPMC category.

\textsuperscript{36} Given the extensive directory available to the Hospital Council, the representation of the top management of the sampled hospitals was very complete, and it is unlikely that the sample excluded individuals of interest to the study.
years of the Survey of Consumer Attitudes, Groves et al. found no effect of excluding any respondent groups on cross-sectional estimates (Curtin, Presser and Singer 2000). Moreover, a meta-analysis by Groves finds “no consistent relationship between response rate and non-response bias” (Groves 2006, cited in Keeter et al 2006).

However, largely independent of response rate, response bias might obtain based on differential inclusion of respondents in the sample. One way to explore response bias is to examine whether and on what characteristics systematic differences appear between responders and non-responders. Because of the nature of the information included in the original sample, only certain variables could be used to test differences across these groups. Compiling and comparing survey responders to survey non-responders using Chi-squared or Fisher’s Exact tests, there appear to be significant differences among certain categories. Stratifying by job level\textsuperscript{37} displayed significant differences overall, with more Chief and Director level individuals included than expected. While unintended, the oversampling in these categories, as opposed to area/regional level or more operational level staff, suits the intention of the study to explore strategic decisions at the brick-and-mortar level of analysis, and ensures that the survey better reflects the perceptions of upper-level hospital management. Stratifying by job type\textsuperscript{38} displayed significant differences overall, with more individuals in quality roles sampled and far fewer individuals in IT roles sampled than expected. Different coding schemes aggregating roles under administrative and clinical job headings produced similar differences across response groups. Again, there is an “oversampling silver lining” for the study foci as individuals in quality roles were represented at more than double the expected frequency. Stratifying by gender, more females responded than expected by chance overall. This gender discrepancy may cause a degree of overall sample bias, however analysis by job level shows that more Chief level respondents were male and fewer female, and more Director level respondents were female than expected by chance. This finding is not surprising given the preponderance of males in Chief positions and the aforementioned oversampling of quality positions, which accounted for most of the Director level titles included and where individuals were observed as more likely to be female. Without data on the distribution of males and females employed at the BEACON hospital sites, it remains unclear whether this is an important contributor to bias.

\textsuperscript{37} Job Level was hand-coded based on job title. The job level codes were: Area/Region, Chief, Director, Manager, VP and Unknown.

\textsuperscript{38} Job Types were hand-coded based on job title, cross-referencing against publicly available databases as necessary to determine primary function. The job type codes were: Clinical (excluding medicine and nursing)*, Executive*, Finance, HR, IT, Legal, Marketing, Medicine, Nursing, Operations, Quality, Risk/Compliance*, Safety*, Strategy, and Unknown (* denotes primary function)
A final potential problem is missing data across variables within the dataset of responders. In this case, the responses received were very complete – 76.3% of respondents completed the survey and the average rate of completion among responders by facility was 67.7%. Because of the linear web layout of the survey, nearly all the missing item data appeared in the final questions of the survey. Furthermore, the excluded cases were removed because of near-total non-response to the survey questions. The characteristics of the ten excluded cases closely matched a random draw from the distribution of the larger sample: there were seven females and three males, a mix of job titles and levels (although notably, none from the Chief level), and they were employed at nine different facilities, two of which were at the area/region level. In a generic dataset, it would be appropriate to check for missing data created by incomplete cases dropped before statistical analysis by using t-tests comparing values between included and non-included samples. If excluded cases represent a small percent of the total data (often, <5%), then a complete case analysis can be performed using the dataset with exclusions.

The data provided by responders is likely to be representative of individuals who consider evidence-based management a topic of substance, whereas those not responding might either consider the topic unworthy of serious debate or lack the time or motivation to respond to a survey shaping the debate on this topic. Thus, it could be that responders already had a bias towards the importance of investigating evidence-based management, whether for or against, and that this survey may in fact be capturing more extreme cases of individual perceptions for and against the concept than would normally be present in a complete dataset. It could also be that responders over-represent individuals who are more analytically-minded and familiar with the audit-and-feedback culture exemplified by these types of surveys, and that these individuals would, on the whole, be more likely to have strong opinions on the topic of evidence-based management.

In sum, although the overall response rate is not likely to contribute to response bias, the systematic bias caused by the oversampling of certain subpopulations noted above requires care be taken in generalizing analytic inferences widely beyond the populations. With respect to this study of evidence-based management, it appears that whatever systematic bias might exist due to non-response in the survey panel, at least by the available metrics, is unlikely to derail the empirical goals of the study. Given the overall exploratory use of the survey dataset, these limitations seem reasonable to pursue analysis.

§4.3.4: Survey Results

All response data were downloaded from Qualtrics in a comma-separated format (.csv) file and imported into the Excel version 2010 and Stata version 11.
software packages for analysis. Supplementary analyses were also conducted using the R version 2.12.2 software package. The dataset was reviewed for technical errors and cleaned by the investigator, then merged with existing organizational data for the respondents.

Demographic statistics are presented in Figure Group 11 below. The sample was overwhelmingly female, middle-aged, and well-educated, and there were roughly equal numbers of Chief level and non-Chief level individuals represented. Respondent tenure in his or her current position ranged up to 24 years, with a median tenure of 3.25 years.

![Figure Group 11](image)
General descriptive statistics are presented in Figure Group 12 below. Nearly all respondents reported that their personal decision-making style was a good or excellent fit with the culture of the organization, and that they frequently or very frequently contacted co-workers when making strategic decisions. Most respondents reported believing that their organization used BEACON as a source of evidence for innovation decisions about the same as a comparison organization they personally listed. Given an example of strategic management decisions to implement a new program or policy around sepsis, respondents attributed a wide range of importance to the BEACON Collaborative, ranging from critical to peripheral, with the weight of responses on “equal in importance to other sources.”

Respondents were also asked two questions about their orientation to strategic decision-making. The first asked “Which of the approaches indicated on the [seven-point Likert-scaled] continuum below do you generally prefer when making or supporting strategic management decisions?” with the anchors being “strongly prefer to copy ‘tried and true’ models [1]” and “strongly prefer to experiment from scratch [7].” This question was intended both to capture individual risk preference, with seven indicative of seeking and one indicative of aversion, and to yield insight into the tendency to look to and copy external exemplars. The overall distribution of responses was relatively symmetrical, although there were more individuals responding at the “copy” extreme than the “experiment” extreme; while no statistical conclusions can be drawn, it is interesting to note that six of the eight respondents in the “strongly copy” group were chiefs or directors of quality. The second question asked “Which of the following items ['a critical piece of evidence collected' or 'the balance of all evidence collected'] more often leads you to decide in favor of implementing a program or policy new to your organization?” This question was intended to capture individual preference for search stopping rules, with the strong majority (86%) responding in favor of the balance of evidence. Subgroup analysis suggested that education level might be associated with this
preference, but given the small number of responses for a critical piece of evidence, further exploration is best to establish a relationship.
In response to the five-point Likert-scaled question, “How much time and money does your organization currently spend to collect and process the evidence necessary to make informed strategic management decisions?” respondents reported that on average their organizations spent about the right amount of time, but too little money on collecting evidence (Figure 13).

Respondents were asked to name up to three sources of evidence, in order of preference, that they “especially trust and frequently seek” when making or supporting strategic management decisions. A note was made that examples of a source were included “a particular website, journal, person, or professional society.” The results are shown in Figure 14 below and summarized in aggregate categories in Figure 15 below. Individual sources mentioned included: the Institute for Healthcare Improvement (IHI), The Advisory Board, the University Healthcare Consortium, experts within the system, senior leadership, Harvard Business Review, and the National Association of Public Hospitals. It is interesting to note the overall dependence on quality/standards organizations, such as IHI, the increasing importance on literature as a source, and the importance of corporate data and reports as a first-consulted source. While some of the mention of quality organizations might be a response bias to the association of the survey with BEACON and patient safety and quality concerns, the overall impression of the important of quality organizations in strategic management decisions persisted in the qualitative portion of the study. Despite the framing of the survey with BEACON, it is interesting to note that the Collaborative represented only a fraction of overall mentions of frequently used and trusted sources, and that these mentions occur only as the second consulted source.
With respect to BEACON as a knowledge intermediary, respondents were also asked what sources of evidence would complement the knowledge collected from the Collaborative and what sources would substitute for that knowledge, if BEACON were inaccessible (see Table 9). The results suggest that individuals either have a wide range of alternatives for evidence similar to that provided by BEACON or that these individuals are relatively indifferent to the sources of
knowledge. One finding of note is the strong reaction to advocates/advocacy group sources of evidence, suggesting that individuals both do not perceive of BEACON as source of advocate-type evidence and that advocates and advocacy groups are not desired as complements to the knowledge provided by BEACON.

A series of questions were asked to elicit individuals’ knowledge preferences when collecting evidence for decisions. Before any question related to knowledge types was asked, an extensive description of the knowledge types and a compressed
definition and applied example of the knowledge types was presented in tandem with the question was presented to the respondent; and concurrent with each question, an option was given to review a shortened reminder list of knowledge typology definitions. The respondents were re-primed to the project scope with the phrase “Consider a decision to invest resources in a program or policy that is new to your organization.”

A first set of questions asked individuals to select the knowledge type that they thought best suited their organization for four aspects of Rogers’ (1962) diffusion characteristics: namely, compatibility with the decision-making style of your organization; trialability of decisions allowing them to be tested out and modified; observability of decisions, in terms of easily observable and measurable outcomes; and “relative advantage” (see Figure 16). Note that the final category was modified from Rogers’ original description of advantage of a proposed innovation over a previous process into a measure describing an innovation that would provide “advantage” over competitors. This altered definition was implemented here not just to capture the advantage of innovation, but also in order to capture the competitive aspect of knowledge acquisition, foreshadowing a series of questions on BEACON, best practices and confidence. Because only one type could be selected, the results provided a personal maximum category which can be aggregated; however, care should be taken in interpretation given that the level of each individual’s personal maximum may differ from one another. In summary, know-who problem mechanisms was most likely to be associated with compatibility, know-how implementation knowledge with observability, and by a small margin, know-what works best practice-type knowledge with relative (competitive) advantage.

Given the same framing, another group of knowledge questions asked the respondents to select one type in response to the question “If you were pressed for

Figure 16
time or money in making or supporting a strategic management decision, what type of knowledge would you want to collect FIRST [sic]?” Under these conditions, individuals selected know-what works knowledge ahead of other knowledge types, with know-how under a money constraint and know-about under a time constraint placing second. It is surprising that know-about knowledge would be highly cited under time constraints, given the technical character of this information. It could be that individuals felt that knowledge uncovering the mechanistic aspects of problems would be most likely to be rhetorically persuasive to colleagues under time pressures. It is also possible that individuals did not fully understand the wording of “know-about problems” in this question, thinking instead that this meant “be aware that a problem exists,” which would amenable to selection under time constraints. Nonetheless, a similar response did not occur in the question about being pressed for time, suggesting either that respondents were discriminating appropriately – or perhaps, that several individuals selected know-about because it was the first of five options in a drop-down selection box. A difficult decision was made to keep the ordering of the knowledge types consistent as about-what-why-how-who across all items throughout the survey, privileging clarity and consistency with the oft-provided definitions over the benefits of randomizing the knowledge types across questions. In either case, it was unfortunately not possible to gain additional insight on the respondents’ motivation due to data collection restrictions.

Using the responses for their top three trusted and frequently sought sources, respondents were asked to select the types of knowledge that this source provided to them when making or supporting strategic management decisions. The question was structured using check boxes, whereby the respondent could select any number of sources from zero to five. Individual responses were aggregated by source across knowledge types, converted into percentages and are reported in Table 10 below. The color highlights in the chart are based on the approximate interquartile ranges from another question in the survey (see below) asking individuals to weigh the importance of each knowledge type – these highlights (red: values ≥ 25%; yellow:
15% < values < 25%; green: 15% ≤ values) should act as a yardstick only, and not as a calculated measure of the importance of knowledge by source. Chi-squared and Fisher’s exact tests were conducted using count data for knowledge types by source against an expected one-fifth of each total N for source, based on a working assumption that each source was equally likely to provide any knowledge type. By this standard, no source provided a given knowledge type significantly over others, with only Quality Organizations marginally different from an even balance of types (p=0.07). As an aggregate, all sources provided more know-about and know-what and less know-who than an even balance of types (p=0.000). Again, the trade-off made to keep the knowledge types in order may have played a role by means of order effects attenuating responses to the final types, especially know-who. However, as in the situation above, it is not possible to determine whether this was a choice, an artifact, or an anomaly. Despite this, it is interesting to note that over 77% of the referenced sources feeding into this chart represented sources of evidence external to the organization of the focal respondent.

<table>
<thead>
<tr>
<th>Category</th>
<th>About %</th>
<th>What %</th>
<th>Why %</th>
<th>How %</th>
<th>Who %</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEACON</td>
<td>25.0</td>
<td>33.3</td>
<td>8.3</td>
<td>25.0</td>
<td>8.3</td>
<td>12</td>
</tr>
<tr>
<td>Colleagues</td>
<td>21.1</td>
<td>21.1</td>
<td>16.9</td>
<td>19.7</td>
<td>21.1</td>
<td>71</td>
</tr>
<tr>
<td>Consultants (Advisory Board/Sg2/etc)</td>
<td>18.2</td>
<td>27.3</td>
<td>21.8</td>
<td>23.6</td>
<td>9.1</td>
<td>55</td>
</tr>
<tr>
<td>Corporate data/report</td>
<td>23.9</td>
<td>23.9</td>
<td>15.2</td>
<td>17.4</td>
<td>19.6</td>
<td>46</td>
</tr>
<tr>
<td>Executives</td>
<td>16.7</td>
<td>22.2</td>
<td>22.2</td>
<td>22.2</td>
<td>16.7</td>
<td>18</td>
</tr>
<tr>
<td>Experts</td>
<td>18.2</td>
<td>27.3</td>
<td>18.2</td>
<td>27.3</td>
<td>9.1</td>
<td>11</td>
</tr>
<tr>
<td>Government (CMS/AHRQ/etc)</td>
<td>29.7</td>
<td>27.0</td>
<td>18.9</td>
<td>13.5</td>
<td>10.8</td>
<td>37</td>
</tr>
<tr>
<td>Literature</td>
<td>24.3</td>
<td>21.4</td>
<td>24.3</td>
<td>20.0</td>
<td>10.0</td>
<td>70</td>
</tr>
<tr>
<td>Other Hospitals</td>
<td>23.5</td>
<td>17.6</td>
<td>23.5</td>
<td>17.6</td>
<td>17.6</td>
<td>17</td>
</tr>
<tr>
<td>Professional/Advocacy</td>
<td>23.3</td>
<td>22.1</td>
<td>19.8</td>
<td>20.9</td>
<td>14.0</td>
<td>86</td>
</tr>
<tr>
<td>Quality Orgs (IHI/TJC/UHC/etc)</td>
<td>24.3</td>
<td>27.1</td>
<td>17.4</td>
<td>17.4</td>
<td>13.9</td>
<td>144</td>
</tr>
<tr>
<td>Web search</td>
<td>50.0</td>
<td>50.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2</td>
</tr>
<tr>
<td>Total N</td>
<td>132</td>
<td>139</td>
<td>108</td>
<td>110</td>
<td>80</td>
<td>569</td>
</tr>
</tbody>
</table>

The question on knowledge variety was asked in the following way: “Q18. Now, imagine that there were no barriers to collecting evidence of any type to support your organizational decision-making process. // Describe the ideal distribution of knowledge that would be reflected in this evidence. Use one or more sliding bars to allot 100 percent across the five knowledge types.” Consistent with the theoretical drivers of knowledge variety, the question was designed as a constant sum scale summing to 100 points. A hidden elapsed-time counter was embedded on this page of the web-based survey in order to determine whether individuals had spent an adequate amount of time on the page in order to read the definitions, or whether they had simply scanned over them and moved onto the next page. During the course of the survey period, an unknown technical glitch occurred
and invalidated the counter – an error which the Qualtrics support staff was unable to remedy. Therefore, an assumption is being made that the individuals read the knowledge definitions with enough detail in order to faithfully complete the subsequent questions.

The scale used for knowledge variety is known as a constant sum or ipsative scale because it uses “measurement based on intra-individual comparisons” (McLean and Chissom 1986) and because, “mathematically speaking, data are ipsative if a given set of responses always sum to the same total” (Meade 2004). In this case, the data are structured as multiplicative ipsative data in which the “variables are measured in relative proportions with respect to their sum” (Chan and Bentler 1998). Ipsative scales are intuitively useful, and are appropriate for individual diagnosis, but have present serious challenges in statistical analysis. First, derived from ipse, the Latin word for itself, ipsative scales produce within-person rankings that measure the relative strength of characteristics specific to a given respondent, and are not appropriate to compare across individuals as if they were absolute levels (Closs 1996, Fischer 2004). The constant sum constraint “does not yield utilities but probabilities. Although functionally related to utilities, they are not the same, because utilities pertain to single parties whereas the constraint that probabilities sum to 100% gives them a relational character” (van der Eijk et al 2006). Second, because the enforced inter-item correlation produces a non-invertible covariance matrix, the inclusion of ipsative variables in regression analyses is statistically problematic (Meade 2004, McLean and Chissom 1986). Third, confirmatory factor analysis is not appropriate in this case, for three reasons: a) because the knowledge types are theoretically related, the items composing the factor are correlated by design; b) the types are related by a constant sum scale design, so they fall on a simplex and factor analysis yields a boundary case; and c) related to the first two issues, factor analysis produces a factor loading with zero uniqueness (or 100 percent commonality). Some have argued that ipsatively-scaled constructs with at least 20-30 items retain valuable statistical properties when there is low average intercorrelation (Bartram 2006, Baron 1996), and others challenge this assertion (Eglert 2010). Despite these concerns, because “ipsative methodologies ultimately yield a rank ordering of values based on importance, they are believed to duplicate the way values are cognitively held by individuals” and are “believed to more closely represent an individual’s true values, rather than his or her public endorsements of socially desirable statements” (Meglino and Ravlin 1998). In any case, a key lesson in this study is to exercise caution using constant sum scales due to their ipsativity, perhaps exploring conjoint analyses as an alternative.

The box and whiskers plots below (Figure Group 18) summarize the findings of the ideal distribution of knowledge question. For the reasons outlined above, it is important to take these values with a grain of salt as an interpersonal comparison.
The sliding-bar interface of the question interface forced respondents to allot the full total of 100 points and did not allow more than 100 points to be allotted at any time. In retrospect, this design left a lot to be desired — several respondents complained in a survey feedback field that the bars did not function as expected, a technical glitch which did not arise in pre-testing and which may have accentuated response strategies, such as equal point allocation. Nevertheless, as expected, the median responses centered around 20 percent for each knowledge type, and there was moderate subgroup variation for interquartile and min/max ranges, shown here across Chief/Not-Chief level of position. These two indications were sufficient to suggest that knowledge variety was worth exploring in practice.

Instead of using individual items or a composite of knowledge types, the level of knowledge variety was assessed using an entropy measure determined for each respondent. Frequency counts are inappropriate to measure complexity when the relative distribution is important. Using an entropy-based measure, “the more uniformly spread the variation is amongst the different categories, the greater the measured complexity will be” (Dooley 2002). In this case, a Herfindahl–Hirschman Index (HHI) was created by taking the sum of the squared proportion of each of the knowledge categories in the five category typology, and a normalized HHI (HHI*) was created for use in analyses (see the equations below).

\[
HHI = \sum_{i=1}^{N} s_i^2 \quad \text{HHI}^* = \frac{(HHI - 1/N)}{1 - 1/N}
\]

The Herfindahl–Hirschman Index is an entropy measure used principally by economists to measure the distribution of company market share. When the normalized measure is 0, all companies in the space have an equal market share; as the normalized measure approaches 1, one company approaches a total monopoly.
In this study, when the normalized HHI is 0, each knowledge type has equal share – namely, 20 percent of the total space of five knowledge types; and as the normalized HHI approaches 1, one knowledge type become increasingly dominant over the others, towards a monopolistic share. When an individual displays a preference for knowledge variety, the normalized HHI will be near zero, indicating that individuals desire to retain complete type diversity in knowledge regulation. The normalized HHI was deemed to be more intuitive and also allowed interpretation according to percentage change.

The small sample size (N=93) and comparatively large number of organizations (N=38) in the dataset resulted in most organizations being represented by three or fewer respondents; only five organizations had more than three respondents, and the maximum number responding from any given hospital was eight (see Figure 19). Because of this low within-hospital response rate, two planned routes for analysis had to be abandoned. First, organization-level analyses using HHI scores aggregated across individuals within hospitals were not valuable because there was too little variation across composite scores. Second, because there were too few i-level responses within j-level categories, it proved inappropriate to use hierarchical models of individuals within hospitals (or further, j hospitals within k systems) using the Stata gllamm command or similar functions.

As expected, the normalized HHI measure was skewed right, an observation verified through normal probability and quantile plots. Using the ladder and gladder functions in Stata, a square-root transformation was selected and a transformed variable created for use in linear regression modeling. Figure Group 20 displays histograms for both transformations overlaid with a standard normal curve and a kernel density estimation curve. Statistical outliers were examined by individual respondent and a decision was made to keep them in the dataset, as they represented legitimate responses containing useful information and did not appear to be technical errors.
Bivariate OLS regressions of the square-root transformed normalized HHI measure on various predictors were conducted to examine potential associations. Given the small size of the dataset, it was not possible to control for organizational affiliation using fixed effects – adding nearly 40 fixed effects to the specification would have quickly used up degrees of freedom in analytic models. Instead, hospital-level clustering was addressed through post-estimation Huber-White heteroskedasticity-robust standard errors, using the Stata cluster(organization), robust command. Correlation analyses and post-estimation variance inflation factors were used to determine the appropriate control variables included. Analyses were considered indicative of, but not confirmation for, a potential association, and additional analyses were conducted. Throughout specification testing, an interesting potential relationship emerged between HHI and the number of years since an individual completed his or her most recent academic degree, as suggested by the regression output and the difference at the tails of the confidence intervals (Figure Group 21; regression, left; histogram of years since last degree, right).
Following McDowell and Cox (2004), a generalized linear model (glm) was further used to estimate individual-level relationships where knowledge variety was the dependent variable. They note that, when the dependent variable is a proportion, an option is “to estimate using glm with family(binomial), link(logit), and robust; this is the method proposed by Papke and Wooldridge (1996).” This manner of analysis is called a fractional logit model. A regression of normalized HHI on years since recent degree with organization-level clustered, robust standard errors was conducted and remained significant. A final model for this relationship including covariates was chosen by theoretical relevance, joint covariate testing, and minimizing the reported Akaike- and Bayesian Information Criteria. Controlling for age, sex and education level, years since recent degree was significant at the p=0.002 level.

Marginal effects for parameters were calculated post-estimation, allowing interpretation in terms of the values of normalized HHI. In summary, a one-year increase in years since last degree is associated with an average 0.0034 percentage point decrease in the value of normalized HHI (se: 0.0011; 95% CI: [-0.0055, -0.0013]). This is a meaningful change, given that values of normalized HHI are strongly concentrated around 0.2, with an interquartile range spanning only 0.09 percentage points. A simultaneous quantile regression of normalized HHI was conducted to determine whether the effect was being driven by a large or a small number of years since graduation. Quantile regressions use least absolute deviations from the median, make no distribution assumptions, and are tolerant of outliers included in the model; and simultaneous quantile analyses in Stata use bootstrapped covariance matrices, allowing comparison of coefficients between quartiles. Results for years since recent degree were significant only for the top quartile, suggesting that those individuals furthest removed from formal education (>29 years) were driving the overall effect for HHI. However, this is not confirmable because significance testing between quartiles did not show significant differences (although marginal for the interquartile).

These results are interesting in terms of Cascio’s (2007) admonition that managers many years removed from formal academic training may summarily dismiss some research as “too academic” for practical application. In this sample, decision-makers further from formal degrees tended to privilege a wider range of knowledge than their colleagues. Moreover, this result was insensitive to the inclusion of an individual’s tenure in his or her current position. In light of this

39 In Stata 11: glm normhhi recentdegree_age sex edlev age, link(logit) family(binomial) cluster(org) robust. Education level was coded as associates, bachelors, masters, and doctorate level. Years since recent degree was continuous and age was coded in categories, as shown in demographics above.

40 In Stata 11: sqreg normhhi recentdegree_ago age sex edlev, quantile(.25 .5 .75) reps(100)
essay, it is possible to interpret Cascio’s suggestion in a different way. It might be that individuals removed from formal education do, in fact, dismiss some evidence as “too academic” – not necessarily because they detract from the value of one type of knowledge, such as know-about, but instead because they appreciate the full range of knowledge in evidence, including know-about and know-why. In other words, perhaps individuals with more time in the world of practice tend towards a more “generalist” perspective on knowledge than their counterparts. The survey data do not allow further inference, but is possible that this generalist perspective is generated by the experience that individuals accrue having been confronted with different kinds of decisions and having had the opportunity to personally observe feedback from these decisions over time.

Concluding the survey were several questions attempting to establish a link between knowledge variety and confidence. Using the respondent’s own definition of evidence-based management given earlier in the survey, these questions asked individuals to describe how confident they would be in a strategic management decision using that approach. The interface for these questions was a separate sliding bar ranging from zero (not at all confident) to 100% (fully confident) for each of several confidence categories, including clinical and organizational outcomes.41 These questions, placed at the end of the survey, suffered from even fewer responders (generally one-third less) than for earlier questions, making inference difficult. At first, it appeared that there was a link between an individual’s normalized HHI and his or her confidence that an evidence-based management approach would increase the organization’s legitimacy as a high quality provider. However, this relationship proved fleeting, as it failed to achieve significance under numerous specifications. Nonetheless, it is worth noting a potential link here, given the theoretical importance of legitimacy in driving isomorphic tendencies in organizational decision-making. In particular, there appeared to be a possible association between organizational legitimacy and an individual’s belief that collecting know-what was a priority under organizational funding pressures – although this association also failed under glm specifications with controls. Due to the analytic complexity in linking HHI and confidence, the UC-Berkeley Department of Statistics Statistical Consulting Service was consulted for assistance.42 In tandem with the service, a Bayesian model on a Dirichlet distribution was modeled to overcome problems caused by the constant sum simplex (Hijazi and Jernigan 2009). Despite this effort, the low number of respondents by

41 Other confidence categories referred to an individual’s confidence that the decision: a) will persist unchanged in the next budget cycle; b) will increase your organization’s legitimacy as a high quality provider; c) will yield the same results as those of comparison organizations; d) will not lead to surprising or unintended consequences; and, e) would have changed substantially given more information.

42 Thank you to Karl Rohe and his advisers in the UC-Berkeley Statistics Department for the insight and support on analysis
organization remained a substantial enough problem that significant results were unobtainable.

The survey also asked two open-ended questions regarding personal definitions of evidence-based management and best practices. These items were positioned in the survey flow before any mention of knowledge types so that the elicited perspectives would be baseline of any discussion of the central construct. The section read:

“Answer these questions based on your current understanding without consulting literature, organizational documentation or the opinion of others in your office. There is no right or wrong answer – these questions are to explore the characteristics that you stress as important.

- Q15. What is your personal definition of evidence-based management?
  What processes distinguish an organization that uses evidence from one that doesn’t?
- Q16. What is your personal definition of a best practice for management?
  What are the characteristics of a best practice compared to any other practice?”

Seventy-six responses were compiled in Excel and Stata, and analyzed for content, complexity and commonality. In response to the evidence-based management question, responses ranged from terse comments (“Review of literature” and “Based on what works”) to extensive, paragraph-length definitions. The range of the word count for responses was three to 365, with a median of 32 words, and the longest response contained multiple parts and literature citations. The best practice question was more compact, with a range from five to 91 and a median of 27.

An attempt was made to analyze these qualitative data through categorization and counts. For evidence-based management, two basic categories, “use data/systematic analysis” and “use proven results/know-what” emerged immediately. Similarly for best practice, the categories “use proven results” and “it is an organizational process” naturally emerged. In both cases, these categories covered many responses, but were unsatisfying in their inability to capture the full sense of the response. Definitions of evidence-based management clearly appealing to the use of codified data or systematic analysis accounted for roughly 35% of the responses, and those referring to proven results or observed practices accounted for roughly 15%. Definitions of best practices were split more evenly, with each category accounting for approximately 35% of the total. Nonetheless, the remainder of both questions consisted of miscellaneous response categories, including three appealing to a balance of evidence, about a dozen mentioning both structured

43 Three additional responses ("x", “not applicable”, and “I don’t have a ‘personal’ definition”) were dropped from the analysis
analysis and copying others (“mixed”), and an equal number that used language too vague to classify (“unclear”). Moreover, several responses were unique enough to warrant their own “other” classification, such as a long response that referred to multiple forms of knowledge focusing on “credibility” in each case, one that appealed entirely to internal data and benchmarking, and one that couched a response in terms of processes that would assist the organization in meeting accreditation requirements. In addition to these, a few responses were skeptical of the process at all – for example, two individuals described best practices as “typically a one-organization experience that has driven specific results” and “The distinguishing characteristic of a ‘best practice’ in management is that it is what the speaker wants others to do.” Of note, only three mentions of Lean/Six Sigma or Studer practices were mentioned across all responses to both questions.

The categorized data were examined across various subgroups within the dataset both statistically and for holistic patterns, but no significant relationships emerged, even under moderate degrees of aggregation. Instead, quotations yield valuable insight into the range of respondent definitions – a representative sample is presented in Tables 11 and 12 below. In summary, there was no better example of the wide range of definitions than the following two definitions of management best practices, both from Chief Human Resource Officers, who responded in terms of a data-driven organizational process and of copying proven external results, respectively:

- “Best practice management is knowing that management is not a perfect science. It is a verb, managing a situation, an individual, a team, an outcome, etc. It is gathering facts, identifying desired outcome, building sustainability and making the best decision with known facts.”
- “My personal definition is a tried and true process that has been used elsewhere, fully vetted and has been successful... A best practice company learns from others and takes advantage of great ideas/processes. If someone else has done it successfully, why not learn from that information.”

In an amusing twist, two individuals responded by directly citing definitions from a free online dictionary and from Wikipedia, and yet continued to define evidence-based management as deliberate, systematic appraisal of valid scientific research; about her response, one of these individuals, a Director of Quality Management, further remarked “…no I didn’t make it up”.

91
Table 11: Definition of Evidence-based Management

<table>
<thead>
<tr>
<th>Respondent title</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Director of Quality Mgmt</td>
<td>Evidence-based management is simply using evidence from literature to guide management decisions.</td>
</tr>
<tr>
<td>Director, Risk Management</td>
<td>Peer reviewed literature that has [strong] statistical evidence for the interventions.</td>
</tr>
<tr>
<td>Chief Executive Officer</td>
<td>Organization uses best-practices based on the highest quality studies and practices available in order to make decisions. This principle is used throughout the management of the organization. Practices are implemented using PDSA cycles, and tested [rigorously] and routinely. Decisions are supported by and continuously reinforced by process and outcomes data.</td>
</tr>
<tr>
<td>Chief Medical Officer</td>
<td>Management decisions based on literature in combination with personal experience … get results that last</td>
</tr>
<tr>
<td>Chief Financial Officer</td>
<td>Making decisions using a planned process and attempting to include various interested stakeholders.</td>
</tr>
<tr>
<td>Director, Risk Management</td>
<td>It brings knowledge and questions to the table, and allows others to participate, learn and share. Not based on one’s opinion, but based on many. Now a day [sic] almost all organizations use evidence-based practices.</td>
</tr>
<tr>
<td>Chief Quality Officer</td>
<td>Use principles published in well-respected journals to make management decisions.</td>
</tr>
<tr>
<td>Chief Operating Officer</td>
<td>Use of a management model that has been researched and found to be successful across a broad spectrum of organizations.</td>
</tr>
<tr>
<td>What distinguishes an organization that uses evidence?</td>
<td>A clear structure for assessing evidence, creating opportunities for thoughtful review/discussion of the evidence, a sense of priorities throughout the organization, and a decision-making process that is not stacked toward the capricious or political.</td>
</tr>
<tr>
<td>Director, Quality Management</td>
<td>Organizations that do not use evidence may “fly by the seat of their pants” e.g. make spontaneous decisions based on what one person or a group thinks is best, or how it has been done in the past. Organizations that do not use evidence may be non-system/process thinkers. Examination of the causal factors may not have occurred.</td>
</tr>
<tr>
<td>Chief Nursing Officer</td>
<td>A wide-spread cultural norm that includes asking for/providing the credible, objective data about what actually works which supports a proposed change or action. This should occur at all levels of the organization, frequently, and matter-of-factly.</td>
</tr>
<tr>
<td>Vice President, Risk Mgmt</td>
<td>Distinguishing characteristic of an organization that uses it would be a little slower and more methodical. To me it implies a more academic approach or setting.</td>
</tr>
</tbody>
</table>

Table 12: Definition of Management Best Practices (Survey Data)

<table>
<thead>
<tr>
<th>Respondent title</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Medical Officer</td>
<td>Best practice means finding benchmark for comparison of well defined measures of management.</td>
</tr>
<tr>
<td>Chief Executive Officer</td>
<td>Best practice is based on comparison of outcomes/metrics for financial and/or quality measures and choosing the facilities that have the best results.</td>
</tr>
<tr>
<td>Chief Executive Officer</td>
<td>Benchmarking to obtain consistent desired outcomes</td>
</tr>
<tr>
<td>Clinical Education Specialist</td>
<td>Based on what works and what most organizations have chosen.</td>
</tr>
<tr>
<td>Chief Quality/Patient Safety Officer</td>
<td>Consulting other hospitals that are successful</td>
</tr>
<tr>
<td>Chief Operating Officer</td>
<td>Using a practice that is supported by reliable and valid data from a trusted source.</td>
</tr>
<tr>
<td>Chief Operating Officer</td>
<td>The term “best practice” begs the question – compared to what? Best practices are typically transported from other organizations, and are “best” when compared to other organizations. What is best in one organization may not always be best in our organization.</td>
</tr>
<tr>
<td>Chief Nursing Officer</td>
<td>There are WAY too many practices labeled as “best practices” that are based on nothing more than the opinion of some leader about a new process that was successful in a “small test of change”.</td>
</tr>
</tbody>
</table>

What are the characteristics of a best practice compared to any other practice?

| Chief Financial Officer                  | A best practice either produces superior results using comparable resources or similar results using fewer resources.                          |
| Chief Medical Officer                    | Characteristics include well defined measure and statistics of spread of data, for example, quartile ranking.                            |
| Chief Quality Officer                    | Many organizations report it has been successful. A research study may not have been [conducted] to validate success.                    |

§4.4: Interview Site Selection: Details, justification and limitations

Building on the findings of the survey portion of the study, a series of interviews were conducted in select BEACON hospitals. Interviews are particularly good at eliciting the “analogue quality” of knowledge, in which “stories, metaphors, analogies, and shared experiences” help communicate tacit knowledge (Nonaka 1994). The goal was to elicit models of how evidence-based management and best practices operate, the extent to which different knowledge types were used, and how BEACON was seen as a knowledge intermediary.
In the original research plan, a preliminary analysis of the survey data along the knowledge type dimension was to have generated a list of hospitals ranking from high to low values of knowledge type variety. From this ranking, three exemplar hospitals were to have been selected, representing organizations using high, medium and low levels of interpersonal knowledge variety in strategic decision-making, and in-depth interviews conducted with top management at each hospital to illuminate decision-making across the continuum of knowledge variety. However, given the low within-hospital response rate, it proved difficult to select these hospitals based on this criterion. In many cases, a given hospital had only one or two respondents, making the intended collective hypothesis moot.

Given the statistical uncertainty around the organizational-level aggregate of variety, a decision was made to select a new criterion for interview site selection. The best potential candidate was Question #10 on the survey, which asked the respondent to select from among a drop-down list “the BEACON member that you most admire for its capacity and ability to make strategic decisions.” This question was widely answered as it appeared early in the survey (i.e., the third question after the demographic section), was simple to select, and was asked before the items and priming on variety, a choice originally done in order to capture the respondent’s general sense of admiration while naïve to the specific concept of variety as an issue that might drive their level of admiration.

Although this question did not encompass the theoretical depth of the variety question, it proved to be the best option given the timeline and access limitations facing the study at the time. Using the choice of admired hospital allowed the study to capture decision-making exemplars based on the quality and robustness of their strategic decision-making. Moreover, using a highly-admired source retained the study framing around a) the external perception and evaluation of a source for best practices and b) the appendant network-like inter-organizational links of influence. Nonetheless, because it obviously deviated from the research protocol, exclusive focus on variety hypothesis in the interview period became untenable.

From a total of 78 responses, two hospital sites stood out, having over 10 mentions each and accounting for approximately 30% of all responses. These hospitals have been labeled as Private NFP (i.e., private not-for-profit) and Teaching. Both admired hospitals had very good Leapfrog scores in the past three years, one was a HealthGrades Patient Safety Award winner, and the other a multiyear HealthGrades Clinical Excellence Award winner. A counter-example, labeled Safety Net, was chosen from among the group of hospitals with low or no mentions as an admired hospital. Because it was not asked directly, it is difficult to select a hospital that was explicitly not admired for their capacities. However, Safety Net proved to be an interesting site for four reasons: first, as the label
implies, it is a safety net hospital, serving the indigent and uninsured; second, it had invested heavily in the activities of the BEACON Collaborative over several years, and was one of the most well-represented hospitals in their membership; third, in that year, it had no major safety accolades and had received a major fine for a recent error; and finally, its only mention as an admired hospital was made by an employee of the same facility. All three of these hospitals were at the BEACON kick-off meeting in June 2005, and had attended over 75 meetings of BEACON activities over the course of six years of membership. Each one also had also subscribed to the BEACON sepsis initiative, reporting a desired 25-30% decrease in sepsis rates for the course of the project. Furthermore, all three sites were large, ranging from 400-600 beds, but represented distinct geographic location and patient populations, clinical and fiscal challenges, and current operational foci. A series of charts and graphs are presented below to summarize; attendance data was compiled from internal BEACON sources, and complaints/ incidents/ enforcement actions were compiled using publicly-available data from the California Department of Public Health, Health Facilities Database.

<table>
<thead>
<tr>
<th>Table 13</th>
<th>All Hospitals in BEACON</th>
<th>BEACON Hospitals that attended &gt;75 meetings</th>
<th>Total Meetings Attended</th>
<th>Total Attendance at Meetings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>50.6</td>
<td>85.2</td>
<td>99</td>
<td>292</td>
</tr>
<tr>
<td>SD</td>
<td>24.5</td>
<td>7.2</td>
<td>96</td>
<td>731</td>
</tr>
<tr>
<td>N</td>
<td>45</td>
<td>9</td>
<td>81</td>
<td>354</td>
</tr>
<tr>
<td>Min</td>
<td>1</td>
<td>80</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>99</td>
<td>99</td>
<td>731</td>
<td></td>
</tr>
</tbody>
</table>

Figure Group 22
Table 14: Executive Participation in BEACON Site Visits

<table>
<thead>
<tr>
<th>Site</th>
<th>Executive Suite Attendance</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private NFP</td>
<td>CEO</td>
<td>CEO</td>
</tr>
<tr>
<td>Teaching</td>
<td>CEO, CMO, Asst CMO, CNO</td>
<td>CEO, CMO, Asst CMO, CNO</td>
</tr>
<tr>
<td>Safety Net</td>
<td>CEO, CNO, Chief of Staff</td>
<td>CEO, CNO, Chief of Staff</td>
</tr>
</tbody>
</table>

| 2009          |                                                  |               |
|---------------|                                                  |               |
| Private NFP   | CMO, CNO, Vice-Chief Clinical Operations         |               |
| Teaching      | No formal site visit conducted at this site      |               |
| Safety Net    | CEO, CNO, Chief of Staff                         |               |

| 2010          |                                                  |               |
|---------------|                                                  |               |
| Private NFP   | CEO, CMO, CNO, Dir Quality/Safety                |               |
| Teaching      | No formal site visit conducted at this site      |               |
| Safety Net    | CEO, COO, CMO, CNO, CQO                         |               |

| 2011          |                                                  |               |
|---------------|                                                  |               |
| Private NFP   | No formal site visit conducted at this site      |               |
| Teaching      | No formal site visit conducted at this site      |               |
| Safety Net    | No formal site visit conducted at this site      |               |
All three hospitals had mission and vision statement publicly and prominently available; a summary is presented in the table below.

<table>
<thead>
<tr>
<th>Table 15</th>
<th>Mission</th>
<th>Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private NFP</td>
<td>Innovation, local control, competitive on cost and quality</td>
<td>Known as high quality, compassionate, and responsible resource manager</td>
</tr>
<tr>
<td>Safety Net</td>
<td>Compassion, respect</td>
<td>Patient safety, staff retention, culturally responsive, efficient resource manager</td>
</tr>
<tr>
<td>Teaching</td>
<td>Healing, teaching, discovering</td>
<td>Be best provider and employer</td>
</tr>
</tbody>
</table>

§4.4.1: Interview Sampling

At the three selected hospital sites, interviews were requested with the entire top management team and upper management in quality, safety and risk positions who were either included in the survey panel or were specifically identified as individuals possessing perspectives worthwhile to elicit by BEACON staff, other study participants, or executive administrative assistants at each site. As much of this sample was interviewed as possible. On the whole, the sites were very supportive, and individuals were candid in sharing their opinions – in only one case (at Safety Net), was an explicitly requested interview rejected.

Interviews were semi-structured, following the same layout as the survey, but asking more in-depth, open-ended versions of questions. An interview guide was
developed, focusing on fine-grained aspects of knowledge types that were not fully captured in the survey portion of the research. After a few interviews, it became apparent that extensive questions about knowledge types proved too oblique for most interviewees. Instead, individuals were prone to focus on sources of evidence, and it was difficult to adequately correct this bias within the time permitted. Therefore, subsequent interviews did not directly address the issue of knowledge types, assuming that individuals were more likely to disclose their perspectives as a byproduct of discussing their strategic challenges and evidentiary sources. That individuals did not answer these questions should not be taken as an indication that they were unable to comprehend the knowledge types, but rather that they were not able to sufficiently separate knowledge types from knowledge sources within the scope of the interview.

All interviews took place in the respondent’s workplace, either in his or her office or in an adjacent conference room. With explicit permission of the respondents, all interviews were audiotaped and detailed meeting notes were taken using a Livescribe Pulse Smartpen. Together with the Smartpen-coded notebook, spreadsheets were developed to sort and order data by respondent, providing the requisite broad-view perspective for inter-respondent comparisons. Because of the utility of the detailed meeting notes, a decision was made to maintain working quotation files, but not to fully transcribe the interviews. Transcription is most useful in a research context when the researcher either intends to analyze the data using qualitative software such as N*Vivo or Atlas.ti, or to count and compare mentions of a given specific word or concept in spreadsheets like Excel. Neither motivation for transcription proved valuable in satisfying the exploratory, narrative goals of this project. Dedicated qualitative software packages are most valuable when conducting predominantly grounded theory processes, in which categories are unknown at the start of analysis and emerge through coding, or when there might be temporal structure to the qualitative data that could be captured and analyzed. Because the interviews were structured into sections and most stories conveyed were discrete and brief, temporality in the interview was not a major concern. Moreover, because there were formal ex ante categories being investigated, the interviews explicitly prioritized these items and a grounded theory process would be inappropriate, as it would reproduce these categories in the theoretical limit.

The major analytic limitation of the qualitative analysis is that there was only one rater of the data collected, making the study quasi-ethnographic in nature.

---

44 The Smartpen is a device which contains a small audio recorder and camera facing towards special micro-dot lined paper (i.e., the camera is only usable on the notebook; no video recording of the interview was made). When used with special paper provided, software in the pen time-stamps the audio of the interview against the notes taken, allowing the researcher to easily find the audio corresponding to a given note taken or note made at a given time during the interview.

45 This is my personal reflection based on several projects using each of these software packages separately and in tandem.
In multi-rater analyses, there is the possibility to calculate the degree of correspondence between raters on the categories created and frequencies of given mentions using inter-rater reliability measures, such as a kappa statistic. As the correspondence of the various raters nears 100 percent, the data are presumed to be very structured around the given categories elucidated. Time, money and the requirement to amend the existing IRB proposal to include a new study team member prevented hiring a research assistant. As the interviews were already structured across theoretical categories of interest and tied to the survey portion of the study, it is hoped that the additional correspondence of a second or third rater would not have significantly changed the findings. Nonetheless, the findings of this section are presented in a conservative way to avoid overstating observations.

§4.4.2: Interview Results

Twenty-four individuals were interviewed in 23 sessions across the three hospital sites, resulting in 22 hours of audio data collected. There were 11 interview sessions at Teaching, seven at Private NFP, and five at Safety Net, accounting for a large proportion of top management at each site. The overall average time per interview session was 57 minutes and 24 seconds; the maximum interview took one hour and 19 minutes and the minimum interview took 41 minutes. Interviews were structured into four basic sections:

- Respondent characteristics, including educational and employment history, and an ordered list of his or her strategic priorities for the organization;
- Definitions of evidence-based management, its operation in the organization, and emphasis on types of knowledge, where possible;
- Definition of best practices and where they are found; and,
- The utility and substitutability of the BEACON Collaborative in organizational innovation decisions around patient safety and quality.

The interviewees represented the highest-level decision makers in their organizations, responsible for strategic management and organization-wide innovation decisions. The Chief Executive, Chief Medical and Chief Nursing Officers at all three sites were interviewed, together with a range of other top management positions whose input and influence affect strategic decisions in Table 16 below. Nearly all respondents had an advanced degree; there were four MDs, four MBAs, four MSNs, and a handful of other degrees, including a law degree and several masters in public health and public administration. Organizational tenure ranged from less than one year to over 35 years. The former case was an anomaly – an effort was made to schedule an interview with an outgoing top manager, but a meeting could only be arranged with his replacement, an individual who was a recent addition to the organization, but otherwise had a long tenure in comparison.
organizations. Roughly half of the interviewees had organizational tenures over 10 years, and several of those with shorter tenures had careers in other organizations before moving to the focal organization.

<table>
<thead>
<tr>
<th>Title</th>
<th>Private NFP</th>
<th>Safety Net</th>
<th>Teaching</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Executive</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Chief Financial</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chief Human Resources</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Chief/Asst Chief Medical</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Chief/Asst Chief Nursing</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Chief Operational</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Chief Quality</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Chief Strategy</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Director, Pt Safety/Quality</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Director, Risk</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Other (Librarian)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7</strong></td>
<td><strong>5</strong></td>
<td><strong>11</strong></td>
<td><strong>23</strong></td>
</tr>
</tbody>
</table>

There were a wide range of strategic priorities reported by interviewees, but no clear commonality across individuals within organizations. In the case of Safety Net, establishing and extending a Medicare 1115 Waiver and attaining Magnet status for the organization\(^{46}\) were both mentioned twice prominently, but other priorities were varied. At both Teaching and Private NFP, individuals cited an array of strategic priorities. Some respondents were specific in conveying their priorities, listing topics such as improving information technology to better manage computerized physician order entry (CPOE), reducing patient readmissions and transitions across inpatient and outpatient settings, and managing external agreements for practice referrals. Others were less so, providing unelaborated descriptions such as improving value for money, developing “top notch talent”, and managing culture change to support participation in quality initiatives. A subset is presented in Table 17 below.

\(^{46}\) The ANCC Magnet Recognition Program is a run by the American Nurses Credentialing Center to recognize “quality patient care, nursing excellence and innovations in professional nursing practice.”
Table 17: Examples of Strategic Priorities

<table>
<thead>
<tr>
<th>Respondent title</th>
<th>Organization</th>
<th>Priority (Summarized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Medical</td>
<td>Teaching</td>
<td>Improving patient experience</td>
</tr>
<tr>
<td>Director, Risk</td>
<td>Teaching</td>
<td>Maintaining integrity of medical records system</td>
</tr>
<tr>
<td>Chief Quality</td>
<td>Safety Net</td>
<td>Increase staff alignment with vision and values</td>
</tr>
<tr>
<td>Chief Operational</td>
<td>Safety Net</td>
<td>Establish and extend Medicare 1115 Waiver</td>
</tr>
<tr>
<td>Vice-Chief, Clinical</td>
<td>Private NFP</td>
<td>Avoiding readmissions</td>
</tr>
<tr>
<td>Operations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Director, Quality</td>
<td>Private NFP</td>
<td>Building team-based and collaborative organization</td>
</tr>
</tbody>
</table>

Among other priorities, one CMO described his top strategic priority as improving the patient experience, describing it as an “important competitive feature, and increasingly now almost a regulatory feature..., where there is now both payment and shame applied to ratings of the patient experience.” He noted this as a priority despite his observation that the evidence for a correlation between patient experience and ability to grow “is somewhat lacking when compared to the standards for, say, other forms of scientific evidence, but nevertheless... most of the decisions we make aren’t ‘evidence-based’ either.”

There were almost as many conceptions of evidence-based management as there were interviews: a few individuals were unfamiliar with the term, many responded with panoptic definitions (“National management best practices and literature to support high level tactics that an organization can take that will improve the performance of whatever it is that you’re trying to do with the management”), and some described details of the decision-making process, from convening stakeholders to commissioning literature searches. This mirrored the case of Safety Net, where some interviewees presented well-designed personal processes that captured knowledge across the typology, and others responded: “Is there a systematic approach to how you approach every problem? ... For the most part, no. I mean I’ve got to be honest with you. And I think this is where it’s different from... medical practice, which is much more systematic, organized and evidence-based.” Even at the two hospitals admired for their strategic decision-making capability, there remained no formal organizational sets of procedures to coordinate management decisions. One individual at Private NFP described how she employs multidisciplinary work groups and literature searches conducted by the hospital’s in-house librarians to assess the scope of a problem, identify gaps, and develop interventions. Though “we’re getting there,” she suggested that other executives in her hospital have their own models and that there is no standard
practice for the application of management evidence: “We know the concepts, [the issue is] employing them in day to day work.”

An example of two individuals from Teaching best displays the differences in approach and lack of organizational standardized practices for management decisions. One executive described the challenges of evidence-based management in quality and safety decisions, drawing a distinction between clinical and management decision-making. For strictly clinical decisions, the decision-making process was clearly structured: first, assess the peer-reviewed literature; next, consult comparable organizations nationally; and then, seek the guidance of expert organizations, such as the Institute for Healthcare Improvement (IHI), the Joint Commission, or the Association of American Medical Colleges (AAMC). For management decisions, the process became more complicated. Using handwashing as an example, the executive described the approach to address a problem which grew from “somewhat of an organizational imperative, but without much in the way of progress” into a “major initiative.”

The core of the approach, which was described as “believable comparisons,” stemmed from the acknowledgment of a deeply-ingrained organizational culture around internally-generated knowledge and debate – to the point where “virtually every statement is challenged” and “doing what you’re told is not generally a major motivator.” Benchmarking against comparison organizations suggested that Teaching was underperforming, and that it was possible to “leverage the [natural] competitiveness of the workforce.” An internal team was assigned to explore options, and after technological solutions, such as RFID tags and color-changing soap, were rejected as inappropriate, a decision was made to “look outside, because inside didn’t work.” At this point in the interview, the respondent posed the rhetorical, but salient question: “How long do you fail with internal solutions before you look outside?” The team chose to turn to the Joint Commission because it was considered an expert organization, and because it was the “only organization” at the time addressing problems requiring this type of culture change. The Joint Commission partnered with the internal team in a semi-consulting relationship, including making a number of on-site visits and suggesting that Teaching collects a “giant number of measures and publish[es] them widely.” Internally-generated knowledge contributed to the process and message that would be relayed; for example, interventions including patient involvement were investigated, but failed because patients didn’t adequately question doctors. A final handwashing plan was composed of an “honest” message to staff, resources for rollout, consistent data and frequent reporting designed to generate intra-professional and intra-unit competition, and inclusion of handwashing measures as part of the standard employee incentive program.

47 Inappropriate because they were perceived as not technologically ready for investment
The respondent noted that while there was no set protocol for evidence-based management, there was an “unwritten standard for major decisions.” To be a rhetorically persuasive tool at Teaching, evidence presented to others “must be in written form,” such as a report or PowerPoint presentation. While the content itself is not standardized, persuasive argument must contain both a preamble (“here’s the situation we’re in”) and a rationale with “at least something outside of us.” The question of why an initiative is important to pursue “always has to be accompanied by ‘because they’re doing it’ or ‘because no one’s doing it and we need to be first.’”

At the end of the interview, the respondent summarized: “We’re slaves to evidence in certain areas; others, not so much.”

Also at Teaching, a second executive noted several components of an evidence-based management process: 1) using internal data extensively; 2) acting/waiting for results, and reversing the decision, if necessary; and, 3) finding out “who else has done this and what decision have they made?” To this she added “let’s not reinvent the wheel,” noting that “sometimes as an organization, we’re very insulated, and not necessarily [prone to say] ‘well let’s go out and see what everybody else has done, or have they done anything at all, or do they even talk about this subject?’” As an example, the respondent cited a recent decision to require all of the hospital’s nurses to wear a common navy blue uniform. Although it wasn’t a “burning platform” issue, the decision was made based on three major drivers. First, the respondent described visiting a nearby children’s hospital which had made the same change a year earlier “for the purposes of branding and security,” and speaking with nurses and others who helped frame the decision about “what worked well... and what were the pitfalls.” This evidence proved rhetorically important to obtain the “buy-in” of staff leadership before committing resources over other priorities. Second, the respondent cited evidence from personal observation of watching staff arrive and leave work, noting that “people just didn’t look as professionally crisp as they could be... and even though we have a dress-code, the dress-code was just ignored.” This “visual cue” was also validated by others external to the organization, and was taken as a “call to action to sit down with executive team and discuss the pros and cons.” Third, justification was drawn from a survey which was coincidentally conducted by the University Health Consortium (UHC) during this decision-making process: “[My leadership group and I] had actually done a survey as we were making the decision. And it was interesting how, of the UHC... about a good 20% of those organizations are getting ready to go into ‘apparel’, as they call it, or they already have standardized uniforms.” Additional impetus was given in terms of diffusing innovation. The respondent noted, “everybody else is going the same way, so when you’re in a sort of ‘wave’ of change, do you want to be an early adopter or do you want to be at the end of that wave? And I don’t think we were early, because there were many organizations around this country that had decided to go back to standard uniforms for nurses.”
Convening a staff advisory group of nurses to “talk about it and lay out how we would do it”, the respondent was faced with resistance and had “a dickens of a time” getting the staff to accept the proposals. The initial meetings were marked by “gnashing of teeth of why, and the evidence, and oh my gosh and whatever, but there were a few people in the room, and I give them credit, who gave stories of how they had gone to other medical facilities and people were disheveled and they had a lack of confidence in those people based upon how they looked. So people started listening because it’s their peers. So then we talked about it and talked about it, and had people try on uniforms... It was a good 6-8 month process to select a vendor.” The respondent suggested that resistance was caused by a few reasons: first, that the “big picture” of patient preference and professionalism was rejected by some nurses who preferred variety in uniform color and style; second, that there was concern over “poor body image” from ill-fitting uniforms; third, that there was a slippery slope implication in regulating appearance; and finally, that there was insufficient debate over whether the change would have clinical benefit beyond “this looks nice”.

After initiating the uniform changes, the respondent found that, “lo and behold, 99.9% of the nurses... said that it was the right decision to do this.” This change in attitude was attributed to the uniform’s capacity to provide an identity by which nurses can find and bond with colleagues, and supported by compliments from patients who thought the uniforms looked professional. An “amnesty day” was held to allow staff with personally-purchased “rogue uniforms” to turn them in for organizationally-approved uniforms. Subsequently, an organizational decision was made for all staff to wear identity uniforms, providing a “crispness and sharpness across the organization... which tells everyone what I do in terms of my job.”

Citing Peters, Kotter, Collins and Studer, the respondent confirmed that management decisions were made in a “holistic” way, different than clinical decisions, and noted that “at the end of the day, you have two choices... you can wait for someone to say, here’s how you’re going to do it, or you take charge and say ‘Okay, let’s take a look at the options and this is the way we’re going to do it. We’ll make these recommendations.’ And I opted for number two.”

Despite reporting frequent interaction as a top management team, these two organizational colleagues demonstrated the divergence in approaches and understanding of evidence-based management. Both cases drew on benchmarking against comparison organizations and both leveraged the rhetorically persuasive “they’re doing it” argument. The former example engaged a knowledge broker to develop a strategy that included frequent data reporting to encourage competition, coupled with an employee incentive. The latter connected with an admired organization which had been successful and used bold leadership and internal
champions to support the decision. Both the handwashing and uniform examples drew on multiple types of knowledge. On the one hand, the handwashing example offered a more structured approach appealing to the “know-about” mechanisms around competition provided a convincing rationalist approach. On the other hand, the uniforms example leveraged “know-why” to invoke social proof in a way which was salient and persuasive enough to overcome resistance. It was unclear how their other colleagues approached these respective decisions, so it is difficult to interpret how the group’s knowledge diversity or approaches to evidence-based management helped or hindered the decisions. In any event, both decisions succeeded in being implemented, and the uniform decision even expanded to encompass other roles across the hospital.

Other organizations were not as marked in their examples, but nonetheless faced similar divergence of process across interviewees. As a consequence, subgroup analysis of the data by organization produced only one emergent commonality. Supporting previous findings on the barriers to organizational evidence use, financial resources appeared to be important in evidence-based management processes. Several individuals at Private NFP noted that their in-house library was a major source for supporting literature, and a tour of the library confirmed the range of resources available. Nonetheless, interviews with the chief librarians suggested that when management requested a literature search on a given topic, it rarely explained the context for which the evidence was being compiled, and that management neither returned for additional information after a preliminary report, nor gave feedback on the utility of the evidence reported. The COO of Safety Net confirmed the other side of the proposition, noting that there is “no strategic planning here,” and that externs from nearby universities (e.g., MPH students) are often relied upon as “internal consultants” to produce literature searches of peer-reviewed practice. This strategy was not lost on Private NFP either, which likewise reported that it sometimes leveraged nurses conducting research for Magnet certification to compile evidence reports.

Dividing the data by other subgroup dimensions, including gender and organizational title, failed to return any significant patterns. This also held true in the analysis of interviewees based on the years elapsed since their last degree, an item of interest in the survey data. For this analysis, individuals were split into quartiles of time, and the first (degree date 1970-1979) and fourth (degree date 1993-2007) quartiles were compared across definitions of evidence-based management and attitudes, the utility of membership in the BEACON Collaborative. As across gender and role, both quartiles of six individuals displayed a range of responses, including different levels of importance placed on literature/codified or stakeholders/tacit knowledge and divergent attitudes towards the significance of BEACON in strategic decision-making.

105
Only one interesting association by subgroup emerged – namely, Chief Nursing Officers, and nurses in general, tended to differ from Chief Medical Officers in their attitudes towards the BEACON Collaborative. Where the CNOs praised the work done by BEACON and believed that membership was highly valuable to their organization, the CMOs were more measured, critical and agnostic to the value of participation. While all of the physicians thought that the Collaborative was good for “networking and friendly competition...[with] the kids on our block,” and that “the structure of BEACON made it safe to share,” all reported “limited interaction” with it in their roles as CMO. One noted that the “energy of BEACON is infectious,” but that “[we] could have gotten more out of it, if we put more into it.” Two respondents were critical especially of BEACON’s lack of rigor in data collection and reporting, particularly with respect to risk adjustment. One individual noted that he would not consider BEACON an expert organization in league with the Institute for Healthcare Improvement or the Joint Commission, and another thought highly of the organization, but noted that he was left with the problem of what “we do with the tools they provide.” In contrast, nurses, thought that the Collaborative “filled in a big gap locally,” providing a “real collaborative” in which there was “local enthusiasm and motivation to participate, to be the best.” They were markedly less critical about data quality: one individual appreciated the outcomes quantified in terms of lives saved, another praised the sessions for frank feedback of data, and no individual specifically highlighted data problems. Instead, they pointed to the benefits of belonging to a community of similar individuals, noting the benefits of hearing “what other teams are doing,” being recognized by peers for successes, and knowing that “you are not alone in solving the problem.” One CNO proudly noted sending one of her staff to present and collect feedback on a small test of change using SBAR implementation, and another described how her staff became “pollinators.” They praised the benefits of sharing “stories, anecdotes and best practices” and being able to hear from national experts, because attending a Collaborative meeting “brings the IHI to me.” One nurse suggested that BEACON helps executives determine priorities by demonstrating “what is most applicable to the environment” – a point reflected by the CEO of another organization, who noted that the Collaborative “took a lot of variables off the table.” Another individual considered her organization’s membership in BEACON to be a “great recruiting tool” for new nurses and a signal of the hospital’s appreciation of nursing roles.

The distinction between nursing and medicine on this point is not unexpected. Having attended several BEACON events, it was clear that the principal attendees, by a large margin, were nurses in both nursing and quality or patient safety roles. On the one hand, much of BEACON’s programming was on operational activities in quality and safety, for instance on medication management.

---

48 SBAR stands for “Situation-Background-Assessment-Recommendation”, a communication tool developed for healthcare by Kaiser Permanente and disseminated widely through the Institute for Healthcare Improvement.
– a level which might predominantly draw individuals with nursing, rather than medical, degrees. It might also be that more of these operational-level individuals were able to attend, either by total number or by number who could leave work to participate in quality improvement or professional development activities. On the other hand, in addition to high-level, system-focused programs (such as a lecture on human factors engineering at one quarterly meeting), the Collaborative also made special efforts to develop sessions addressing and empowering the role of nurses in the quality and safety process – for example, one keynote speaker at an annual meeting spoke about doctor-nurse communication, emphasizing the perspective and plight of nurses in the face of physicians who often belittle them. One nurse interviewee simply described BEACON as “nurse-centric.” In both cases, it is not surprising that the interviewed physicians either did not attend or did not consider the meetings useful, and suggested the general observation that in “nursing care there was a more productive relationship.”

Nonetheless, nurses were also critical of the Collaborative – in almost every case, nurses mentioned the importance of INLP (the Integrated Nurse Leadership Program, also a program funded by the Gordon and Betty Moore Foundation), the Magnet Recognition Program, IHI, UHC, and the SNI (Safety Net Institute) as similar, sometimes substitutable, sources for the Collaborative. One nurse pointed to the benefits of “collegial relationships... [but noted that there was] very little new to learn,” and another proposed that after several meetings, attendance became more of a “social event.” One individual noted that membership in BEACON (and INLP) helped her colleagues develop basic skills, after which they were able to rely on BEACON less and do more quality and safety work internally. It is important to note here that this subgroup association did not hold for non-physician or non-nurse executives, with some describing the Collaborative in glowing terms and others reporting that they had never seen a BEACON-related item cross their desks.

Despite the lack of relationships by subgroups, there were a number of common insights shared across interviewees. The first was the mention of the Advisory Board and Sg2, healthcare consultancies and knowledge intermediaries, by virtually every respondent regardless of organizational affiliation. Opinions on these sources varied, but they were generally viewed as “key organization(s) to synthesize” information and give “tips and tricks” on strategic management topics – leading one CMO to reflect that “What some people say is that if you read everything you can from the Advisory Board, then every hospital ends up with basically the same strategic plan. And actually, it’s kind of true. If you look at the emphasis that most hospitals have, their emphasis on growth is in the high-margin areas – so all hospitals focus on cancer, orthopedics, heart vascular, and women’s services, and some focus on neuroscience and neurosurgery.” The second was a reliance on benchmarking as a decision-making driver that might be used in identifying best practices. One individual noted that benchmarking is “not
necessarily to find the place that has solved it, but to recognize that we need help, so we use benchmarking as a way to signal to ourselves that help is needed.” The third was the acknowledgement that it was acceptable and desirable to copy practices where possible. To wit, one COO mentioned that a decision always includes asking, “Are there other hospitals that are willing to share? ... We have no shame, we’re great at not reinventing the wheel;” a Chief Human Resources Officer suggested that one should look for proven benchmarks and not “reinvent the wheel;” a CNO mentioned that there is “no magic bullet” and executives should “replicate with awareness of how things work here;” a CMO suggested that not everyone has to “recreate the wheel,” but rather, that it is preferable to find and customize strategies; and a CEO simply commented that he was “a big believer in stealing.”

As with evidence-based management, there were a range of definitions for best practices (see Table 18). For most respondents, there was a clear indication that best practices conveyed and were drawn upon principally as a source of “know-what works.” These practices generally derived from respected organizations or benchmarked peers; as the CEO of Teaching noted, “when organizations are identified as leaders, we look to them.” On the whole, the respondents did not appear to view best practices as the inflexible packages or set bundles that might derive from regulatory requirements or participation in standardized initiatives, but as recommendations, starting points, or a marker of “what other hospitals are doing.” A CNO noted the importance of “knowing you can’t just lift it from one institution to the next, because I don’t think it’s quite as easy as that, or we’d all be doing it.” One particularly sophisticated Chief Strategy Officer described the best practice process as observing a “pattern of the way people approach things and some people seem to get it better, and you begin to see who those organizations are like ThedaCare, Cleveland Clinic, Mayo [Clinic] certainly, Group Health Cooperative of Puget Sound, [Intermountain Health (named later)], and a few other organizations are just renowned for being able to translate theory into practice... and those are the ones that I kind of seek out, and watch the literature carefully for what they’re doing. And they’re the bellweathers for us.” Despite the recommendations of these organizations and of knowledge intermediaries,49 he notes that “I don’t end here writing our plan. I want all the perseveration and all of the knitted brows to be ours, not theirs.” Several individuals maintained this healthy skepticism of best practices – one COO noted that the “term is used a lot [and there is] no good answer; personally it means, keep eyes open and learn from others – they have to prove it.”

49 He notes the Advisory Board, Sg2, Medical Group Management Association, Leadership Institute, American College of Healthcare Executives, and the Hospital Council.
<table>
<thead>
<tr>
<th>Table 18: Definition of Management Best Practices (Interview Data)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respondent</strong></td>
</tr>
<tr>
<td>CEO</td>
</tr>
<tr>
<td>COO</td>
</tr>
<tr>
<td>CQO</td>
</tr>
<tr>
<td>CEO</td>
</tr>
<tr>
<td>CFO</td>
</tr>
<tr>
<td>CMO</td>
</tr>
<tr>
<td>CSO</td>
</tr>
<tr>
<td>Chief HR Officer</td>
</tr>
</tbody>
</table>

Nevertheless, despite circumspection on best practices, these organizations were not immune to pursuing management trends. For instance, while the CNO of one hospital described Lean Management as the “choice of the moment” and the CMO of another described it as a “buzz word,” all three organizations were currently pursuing the incorporation of a Lean management plan into their hospital strategy. One individual at Safety Net described pursuing Lean in this way: “we’ve got fires from all directions and we need to get more efficient, there’s no way we can tackle all these problems... well, okay, here’s a fix, folks have figured out to be more efficient... that matches what our needs are... it sounds like a good fit.” Another described seeking an efficiency fix which was “proven like Lean.” Several individuals pointed to the successes of other hospitals in Lean implementation (e.g., “We’ve seen places like Virginia Mason do some very impressive things”). A Director of Quality illustrated the process by which her organization decided on Lean, starting with the desire to pursue “a new leadership structure.” She remarked how “there are healthcare bundles which you should do, and then there are management bundles... so we’ve looked at Lean as something that we want to explore.” Asked how her management team knew about Lean, she noted that “you can’t pick up anything on healthcare without reading something about Lean and healthcare today and how many organizations are moving to it, and Joint Commission is talking about it now...” After taking a group of executives, including top management and board members, to site visits at Boeing and Virginia Mason Hospital, they confirmed their interest in pursuing an initiative and interviewed
three consulting firms to help develop a plan to bring Lean into the organization. Although the project is now underway, she expressed some “reservations about how we’re going to get that to work here because we’ve had issues [related to adopting management best practices] in the past.” With this exception, no individual commented on personal reservations or potential unintended consequences of adopting Lean in his or her hospital. In fact, it is interesting to note that most mentions of Lean omitted citing specific results or studies demonstrating success, and rather appeared to take for granted its association with superior performance and the likelihood of reproducing those results. Only one executive in the interview sample cautioned that “I don’t hear a whole lot of specifics of what it has done for [my colleagues].” A typical response cited some internal efficiency impetus to pursue Lean, but proceeded only in general terms. For example, in response to a question about a decision to pursue Lean that recapitulated the study focus on strategic decisions related to quality and safety, a Chief Strategy Officer noted:

“For us, we know it’s Lean.... We’re going to head into Lean now, because we know that ultimately, we’re going to have to produce more with less. And it’s going to have to be better; we’re going to have to measure an outcome that is... at least as good as or better than we’ve had in the past, and it’s going to have to be less costly or we won’t be able to survive. So that means we need to apply Lean technology to all of our processes. And whether it’s business process or care process, we’ve got to improve our quality, so it means getting variation out and using checklists and all the tools [available].”

The equivocation between “flexible and benchmarkable” best practices and “standardized, initiative- or regulatory-based” best practices was particularly noticeable when overadoption of bundled practices led to innovation failures. At Safety Net, a respondent described how organizational participation in the IHI Transforming Care at the Bedside (TCAB) initiative included the standardized feature of “nursing hourly rounding,” which had previously been successful in private not-for-profit hospitals, but had not yet spread into public hospitals. Because the nurses hated it, she notes, “they bastardized the whole concept into something that is not hourly rounding, and there’s no essence of it.” At Private NFP, an individual mentioned bringing in the standardized Studer management bundle for patient satisfaction, which promptly failed because top management “didn’t walk the walk” (she meant this both figuratively and literally, as in, they didn’t round on the hospital floors) and because staff refused to communicate via scripts such as “Is there anything else I can help you with today?” At Teaching, a respondent recounted an attempt to copy a teamwork training package from a comparison organization; although “it was customized, ... that to me is an example of it didn’t come from the core group that needed the training. It came from external, imposed upon them as, boy, this is a really good idea that will help you in a lot of ways. And the adoption was never complete. They went through the motions, but the concepts never were ingrained and adopted.”
Finally, in the course of the interviews, an unexpected and interesting issue emerged about pursuing best practice through quality improvement. Several individuals mentioned the difficulty of participating in multiple quality projects, collaboratives, and reporting initiatives at the same time, especially if these projects required reporting on different sets of measures, or worse, according to different definitions of similar measures. One Chief Operating Officer noted how there are “so many people measuring so many things. [I wish we] didn’t have 50 different institutions measuring six different things. No one knows what to focus on. ... I think at this point there are [too many measures].” This point was echoed by a Director of Patient Safety who noted that there are “so many well-intended collaboratives,” and that the “desire to be a good citizen” leads to “overcommitment.” Citing her colleague on the skyrocketing number of initiatives, the Director of Risk Management at the same organization noted “how did we ever do this when we didn’t have electronic systems? I mean, some of this is we can do it, so therefore... we are going to do it, even though we don’t do an analysis of ‘should we do it’ or ‘do we need to do it.’” A Chief Quality Officer commented that “it’s a little distracting when you have multiple collaboratives focusing on the same initiative,” and that “it’s either a good thing about being in the Bay Area or bad thing, but there are a lot of people interested in driving improvement about a lot of similar issues.” The staff of the BEACON Collaborative was aware of and sensitive to this conundrum, calling it “initiative burnout.” A Chief Human Resources Officer agreed with the description, “because it’s work, frankly – these all require reporting.” Better prioritization by executives and serial participation can help; one CEO noted that “we try to prioritize what we’re going to do... because we only have so many internal resources,” citing an example in which the organization chose to pursue the INLP sepsis initiative over the BEACON initiative after they could not secure the necessary buy-in from physicians for the BEACON version of the project.

§4.5: Concluding Thoughts on Practice

The qualitative portion of the application study provided several interesting reflections on the state of evidence-based management and best practices in these admired hospitals, but fell short of meeting the initial goals of the application part of this project. In particular, the main proposition of interest was unsupportable for two critical reasons. First, the study hoped to examine the customization of an adopted practice. For reasons described above, sepsis had been selected as an example to examine the extent of overadoption or customization of the “best practices” which were suggested through BEACON Collaborative meetings, resources and training sessions. During the course of the interviews, it became apparent that an association between evidence and sepsis customization would be difficult because of differential starting points and confounding pressures. A
Director of Quality noted that “we did not adopt any of the sepsis bundle initiatives due to our staff already doing the care – we were ahead of the curve. Our mortality rate had already decreased [before starting the BEACON project].” Another CMO noted that BEACON was two years behind cutting-edge standards in their protocol, which cost them reputation in his opinion. Both a CMO and a CQO noted that the bundle significantly overlapped with other pressures, particularly initiatives by INLP, the Safety Net Institute, and the Medi-Cal (California Medicaid) CMS 1115 Waiver Delivery System Reform Incentive Program (DSRP); the CQO added that the BEACON and INLP initiatives “chose to collect their data in different ways, which is very frustrating.” Second, the qualitative interviews failed to display significant enough patterns emergent at the organizational level in order to establish a convergent or divergent shared mental model of knowledge variety for the top management team. In fact, given the inability to discuss knowledge types in any significant detail, the only useable knowledge variety data at the individual level was derived from respondent descriptions of sources of data, which were heavily skewed towards codified know-about and know-what sources – and which were applied in different decisions across different settings. Clearly, these issues were problematic to demonstrating a strong association between knowledge variety and customization.

Nonetheless, in three interviews, the discussion became sufficiently reflective that it was possible to explain the knowledge-types in a fashion similar to the survey instructions and example. In an exploration/exploitation line of argument, one individual suggested that there may be a “point of diminishing returns for learning about how to think about a problem” and that it might be more important to “synthesize intangibles and make decisions better next time.” A second interviewee observed that “all those elements [knowledge types] are happening concurrently, so you have to keep balancing who to involve, when to involve, how do you draw in your know-how, when to tap into someone else’s know-how, when do you go to the literature, and how do you interpret that and draw it into execution.” He continued that “the element that I saw myself getting lost with, in all candor, is the ‘knowing-who’ to bring in [when in a new environment]. I totally get how to do that in [a familiar environment], but it would be hard for me to talk about it. I think it just becomes innate after a certain point, you just know how it all works. But I found myself struggling mightily... [so] I reached out to my colleague... [to understand] the dynamics of the organization and who the informal leaders were and who the formal leaders were. And, all it takes is a couple of gaps [in understanding] and you start to figure out ‘I won’t do that again.’” The third individual, a Chief Strategy Officer, drew a distinction between approaches that were aleatory in nature (“definable” and “oft-used” issues, like benchmarking on cardiology) and those that were epistemic in nature (“esoteric issues”); the former could be addressed using sources like Sg2 or the Advisory Board, but the latter
often required “personal calls” to colleagues to figure out assumptions. About the use of knowledge variety in his organization, he continued:

“Attempts here to sort of teach the organization how to make decisions in an evidence-based fashion are kind of hit and miss. ... I think senior management has demanded that a lot of decisions be made in an evidence-based way, but has it taught the organization the information well about how to do that? No. ... There's not much of a culture around doing [evidence-based management]. You can't get away with seat-of-the-pants arguments, and no facts, and bluster your way through things. You know, people just don't get away with that. But do we help people learn how to be better managers and bring all the kinds of knowledge to bear that you described? Not really.”

Both the survey and interview captured interesting relationships, but neither was robust enough to answer the deep practice questions posed by the theoretical section. Future research designs should consider either strongly limiting a first-stage survey to a focused few questions, or reversing the order such that the survey acts as a confirmatory second stage for interview data. Surveys confronting hierarchical effects should endeavor to power site samples and pursue minimal response numbers accordingly. Interviews are powerful tools, but time limitations restrict the amount of theoretical background that can be given; therefore, it is inadvisable to structure interview questions around topics that are difficult to immediately comprehend. Capturing mental models around knowledge variety start at interviews, but would have been better accomplished through longer-term ethnography, observing the evidence accumulation process and strategy meetings around a given topic with a given top management team; Gasson (2005) provides a nice model for this sort of data collection. If interviews stand alone, there is a tradeoff between allowing individuals to speak their mind and reveal personal mental models through a semi-structured interview format, and structuring interviews more closely in order to allow for better inter-individual comparisons. In cases where team mental models are desired, the latter is preferable. Generally speaking, partnering with organizations takes time and effort. Introductions into the organization are important, but gaining the approval and support of administrative staff separately is critical.

It is difficult to determine whether individuals displayed knowledge variety in the interview section. Respondents clearly attempted to use multiple sources of information and at times cited sources used for mechanisms, justification, previously-successful practices, partners, and practical application, yet it was not possible to determine their predilection for these knowledge types relative to each other. No one asserted that a given knowledge type was clearly dominant, and no particular knowledge type was cited as unnecessary in strategic decisions. There were too few standardized decisions and too little in-depth discussion around them in order to allow counting the relative amounts on a given strategic decision.
Nonetheless, the application section provided some general observations. At the individual level, there appears to be a significant association between the number of years since an individual’s most recent academic degree and his or her propensity towards using a variety of knowledge in strategic decision-making. This effect remained after controlling for age and tenure in one’s current organizational position, suggesting that the number of years in the world of practice, and away from the academy, alters an individual’s preference for knowledge. To the extent that this generalism leads to more robust organizational decision-making, it is an important relationship. The empirical portion of the study does not allow further speculation on the mechanisms by which this operates. However, theory would suggest that because these individuals are steeped in the practice setting, the “patterns of conceptual organization that render data intelligible [to them]” are “subjectively tainted” by their practice experiences (Mattesich 1978). In other words, the more time an individual has to confront problems, take action on them, or personally observe feedback in a practice setting, the more likely his or her “conceptual organization” is to incorporate knowledge beyond the mechanistic data that is conventional to academia. In particular, pattern recognition around the social aspects of problem-solving is likely to create a condition by which individuals give preference to interpretations based on a broad spectrum of knowledge. While no direct links could be drawn between knowledge variety and confidence outcomes, the overall quantitative picture suggests that these proximal outcomes might be worth further exploration. Finally, there appears to be a strong preference among individuals towards a strategic decision-making stopping rule based on the balance of all evidence collected, rather than on a critical piece of confirming or disconfirming evidence. This emphasis on balance might have been partially affected by the nature of the survey or the framing of the question, but nonetheless suggests an interesting development in attitudes towards evidence-based management processes.

One critical change over the five years between since the MATCH project is that the traditional concept of evidence-based management appears to be more widely appreciated by top management in hospitals. In other words, there is a developing zeitgeist that a structured, literature-based analysis is warranted for management topics. The survey was clearer on this point than the interviews. Comparing the total mentions of evidence sources consulted in each study (see Figure 24), quality organizations retained their primacy in importance, but academic literature, which had previously composed only one mention, now jumped to the second most cited source. While interviews revealed that some individuals

50 In the BEACON study, the previous knowledge broker/experts category was divided into consultants (as intermediaries) and experts. While the composite source slightly exceeded the MATCH percentage, experts represented only a fraction of total mentions. Whether or not this is a change between periods, the fact that experts ranked low is notable.
retained a more idiosyncratic, local view of evidence-based management, eschewing academic literature searches for observation, experience and stakeholder input, the general sentiment among interviewees was that there can and should be a more structured process, and that academic management literature could be more frequently incorporated into strategic decision-making. In other words, consistent with MATCH, most respondents described best evidence in terms of academic literature, but proceeded to describe how they use less than they should. As one interviewee noted, management was no longer “just what you were taught in school.”

![MATCH Study](chart1.png) ![BEACON Study](chart2.png)

Figure 24

The finding that emerged most clearly from the study is that there is no standardized procedure for evidence-based management. A critical goal of the evidence-based management movement should be to help organizations develop and maintain a common, or at least commonly-understood, map for strategic decision-making. This finding supports that of Rundall et al (2007) in promoting decision-making standardization, but emphasizes the common role of knowledge across individuals, rather than stopping at the existing model of the “four A’s”. Under knowledge variety, the standardized procedure is to maintain a common set of equally-weighted knowledge types across the individuals involved in decision-making. This common set is maintained through sensemaking processes, which might implemented by means of a consistent script or evidence report structure. This could be part of what a Chief Quality Officer described as a clear structure for assessing + thoughtful discussion + priorities + “not capricious” decisionmaking, or what a Director of Quality described as thinking “in a systems way.” No matter the
strategy around evidence or its particular implementation, as one respondent put it, “the time is right” for organizations to adopt a standardized tool.51

Another general finding was that quality organizations still play a large role as knowledge intermediaries. Prominent sources included the Institute for Healthcare Improvement, the University Healthcare Consortium, the Joint Commission, the Leapfrog Group, and the National Quality Foundation; and use of these sources ranged from passive citations, to benchmarking against national standards or admired organizations, to active participation in initiatives and site consulting arrangements. Despite the prominence of non-profit quality organizations, nearly all of the interviewees also mentioned using consulting groups, such as the Advisory Board and Sg2, to assist in compiling evidence, developing alternatives, or managing implementation. Together with professional (e.g., American College of Healthcare Executives) and governmental (e.g., AHRQ) organizations, these knowledge brokers accounted for a large part of the evidence accumulation and application process. Intermediaries and benchmarked organizations also functioned as sources of best practices. Best practices were defined in many ways, but there appeared to be a difference between “flexible and benchmarkable” best practices and “standardized, initiative- or regulatory-based” best practices. Where organizations participated in the implementation of standardized “management bundles,” there was a risk that they would fall into overadoption and innovation failures.

Furthermore, collaborations are now a part of healthcare management life, and are most useful for learning what other local hospitals are doing. Most participants in the BEACON Collaborative were positive about its contribution to the organization’s approach to quality and safety decisions – but this was especially true among nursing staff, which was specifically targeted by BEACON and reported the most benefit. Each of the interviewed sites was involved in multiple quality improvement collaboratives, which sometimes led them to participate on overlapping initiatives. As one individual noted, “often times is it the case that you end up bumping into other activities.” Where the overlaps led to “negative complementarity,” organizations often struggled through multiple implementation requirements and measure sets, especially when initiative were tied to reimbursement. In some cases, the organization was able to opt for the better fitting intervention, but in others they were forced to “divide the baby... [and] back off both tactics.” As a result, there is a possibility that quality and safety staff might experience “initiative burnout” as they try to meet multiple implementation requirements.

51 She also noted that she had not heard of the Informed Decisions Toolbox, but that she would be interested in learning about it.
In addition to these findings, a number of complementary observations emerged. The first was a need to return to evidence after the initial decision-making process. As the librarians at Private NFP pointed out, top management at their hospital neither returned for additional information, nor gave feedback on the usefulness of evidence collected. This finding suggests two potential issues. First, the breakdown in feedback communication between management and the librarians might stunt organizational learning on how to conduct future evidence searches. Second, the failure to return for additional information might endanger the effectiveness of decisions that face protracted implementation, or those with long-term or high-impact consequences. A Chief Human Resource Officer observed the importance of being vigilant about evidence in terms of assumptions, noting “we have to continue to revisit this, guys, because once we are live all of our assumptions are going to be tested. And what I can commit to you is that, give me six months after we go live, to validate and recalibrate if necessary.” A second observation was the use of externs or Magnet-certifying nurses as levers to collect and synthesize evidence in the face of tight resource constraints. To the extent that these individuals have less practical experience or are closer to academia, they might influence the organizational attention processes around strategic decision-making. Finally, one individual suggested an item worthy of mention: “Hospital administrators and business people don’t have a continuing education requirement.” Despite certification programs available from the American College of Healthcare Executives (ACHE) which require continuing education for recertification, there remains room for improvement in the professionalization of the management field overall and in the reach of management education across operational roles within existing programs.

§5.0: Discussion and Conclusion

This essay established the theoretical importance of knowledge variety as a mechanism for aligning knowledge used in evidence-based management. It also demonstrated how difficult it is to observe knowledge variety in practice, especially as a shared mental model in strategic decision-making.

While the main proposition remains inconclusive, some insight emerged on the mechanism by which individuals draw more equally on types of knowledge. As individuals spend more time in practice settings, they become further immersed in craft knowledge (or, techne). Consequently, this craft knowledge develops practical wisdom (or, phronesis) about decision-making in a given practice setting. In application, an individual’s practical wisdom results in a more “generalist” orientation towards knowledge and evidence. The extent to which decision-making

52 This comment was made by a high-level nurse administrator responsible for clinical operations; she did not refer to ACHE certification during the interview.
groups display preference for a narrow or broad understanding of knowledge might have consequences for the degree to which inter-individual sensemaking occurs. In other words, high or low phronesis, measurable through knowledge variety, might provide a cognitive explanation for how organizations approach innovation decisions. Consistent with the proposal of Greenhalgh and Wieringa (2011), phronesis may be a more fitting way to understand the operation of the translation of research into practice.

In a practitioner response to one of the MATCH publications (Rundall et al 2007), Peter Butler, Executive Vice President and Chief Operating Officer of Rush University Medical Center, offered the following commentary:

“The article references ‘evidence-informed medicine’ as a catalyst to strengthening ‘evidence-informed managerial decision making’ as if these efforts are not connected. Much of managerial decision making, whether led by administrators or clinicians, is all about evidence-informed medicine.”

One way to interpret Butler’s comments is in terms of immersion in the practice setting. As noted above, healthcare is both medicine and management intertwined, and while these two areas differ in terms of professionalization and the validity of environment, they share a common arena of practice for individuals to develop craft knowledge. Evidence-based medicine is by nature more “rational” than the inherently socio-behavioral evidence-based management, but perhaps they meet at a generalist approach – at the pragmatic union of mechanistic and value-laden knowledge. Understanding the cognitive drivers of knowledge and translation may be a viable route to unraveling this perceived “lack of fit of most research evidence with [healthcare managers’] decision-making needs” (Rundall et al 2009).

At the organizational level, retaining knowledge variety in sensemaking might assist the organization in avoiding “errors of closure” in the interpretation of complex environments, thereby allowing easier reconciliation of the errors of underspecification (Schulman 2012). In this way, organizational reliability can be understood as a knowledge problem, and knowledge variety as a way of providing “reliability at the speed of judgment” (ibid.). This is the link between evidence-based management and the study of behavioral risk and reliability. In this sample, healthcare organizations expressed a strong desire to copy the successes of others – but copying presents overadoption and isomorphic risks that these organizations may be unable to comprehend. While theory suggests that specification robustness provides reliability in the face of these dangers, the empirical portion of this study was unable to establish a relationship to either performance directly or to the confidence proxy. It therefore remains an empirical question open to further study.

As a reminder from §3.1, validity refers to the stability of relationships between “objectively identifiable cues and subsequent events or between cues and the outcomes of possible actions.”
One particularly sophisticated survey respondent contested the application of knowledge variety in the practice setting. This individual, a Chief Quality Officer, reported a strong preference for know-what knowledge, resulting in less overall personal preference for knowledge variety. The point is well-taken:

“The five core knowledge types were not something I found of much value in considering the evidence base for good practices. It may be that certain of these are so contextual in the day-to-day operations of a quality department (e.g. know why we would implement) that we just don’t think of them separately. We tend to think of a) what are the evidence based practices we need to implement; b) what are the barriers; c) who owns the process; d) what needs to happen to operationalize; and e) how will we measure success or failure. This is all simplified through the use of PDSA [plan-do-study-act process], so we don’t really discuss epistemological typologies very much, if at all.”

In response, there are two points. First, the fact that some knowledge types are “so contextual... that we just don’t think of them separately” is both reflective of practical wisdom and is a reason why knowledge variety is important. Every unit and every person maintains its own assumptions about what is too deeply understood and shared to specify; the dangers to organizational attention accrue when mental models are oversimplified such that assumptions are not shared in inter-personal decision-making. Facing epistemic uncertainty, an inquiry system using common knowledge categories helps the organization elaborate these assumptions. Evidence-based management that oversimplifies is not robust to the uncertainties of strategic decision-making. Second, knowledge variety is not inconsistent with a PDSA concept, and may be complementary to it at a fundamental level. Deming described a System of Profound Knowledge composed of 1) appreciation of a system, 2) a theory of knowledge, 3) psychology, and 4) understanding variation, which provided a foundation for the PDSA, or Shewart Cycle, process (Deming 2000). Gareth Parry, Senior Scientist at IHI, suggested that knowledge variety might map onto the System of Profound Knowledge by providing the epistemological foundation behind each element of the system (Parry 2011). Knowledge variety, like the System of Profound Knowledge, is designed to provide the sense by which systems approach and understand complex problems. In this way, knowledge variety is not an alternative to the PDSA cycle, but a way of elaborating it, making it more robust, and providing a feedback mechanism for learning processes.

It is difficult to discern the character of the coming “evidence-based global society” and the benefits or unintended consequences that might accrue. What is certain, however, is that the outcome will depend on what counts as evidence and what it means to be evidence-based. If evidence-based management is the map by which we steer toward success or failure, then it will be necessary to revisit the
fidelity of the mapping and consider the consequences of pursuing different approaches. This essay has argued that when organizations draw upon a broad range of knowledge in a systematic way, they are more likely to make decisions that map the increasing complexity of the environment and regulate the uncertainty it poses. To the extent that knowledge variety improves sensemaking and attention, it should also provide the framework for organizations to make strategic decisions that improve performance related to safety and quality. Further exploration on the character of knowledge and judgment in organizational decision-making is recommended, particularly with respect to avoiding overadoption and maintaining reliability over time.
REFERENCES


## APPENDIX

### Analytic Sample (Adjusted Response Rate) by Hospital

<table>
<thead>
<tr>
<th>Hospital</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alameda County Medical Center</td>
<td>0%</td>
<td>10%</td>
<td>20%</td>
<td></td>
<td>20.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alameda Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td>22.2%</td>
<td></td>
<td>42.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alta Bates Summit Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td>10.0%</td>
<td></td>
<td>28.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California Pacific Medical Center</td>
<td>0%</td>
<td></td>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td>42.9%</td>
<td></td>
</tr>
<tr>
<td>Catholic Healthcare West</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Hospital of the Monterey Peninsula</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daughters of Charity Health System</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctor’s Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td>14.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eden Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>El Camino Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td>15.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good Samaritan Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital Corporation of America</td>
<td>0%</td>
<td></td>
<td></td>
<td>22.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Northern CA Region</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Antioch</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Hayward/Fremont</td>
<td>0%</td>
<td></td>
<td></td>
<td>11.1%</td>
<td></td>
<td>27.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Berkeley</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Oakland/Richmond</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Redwood City</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP San Francisco</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP San Jose/Santa Teresa</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP San Rafael</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Santa Clara</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Santa Rosa</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP South San Francisco</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KP Walnut Creek</td>
<td>0%</td>
<td></td>
<td></td>
<td>16.7%</td>
<td></td>
<td>40.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marin General Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mills-Peninsula Health Services</td>
<td>0%</td>
<td></td>
<td></td>
<td>11.1%</td>
<td></td>
<td>30.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novato Community Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Connor Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional Medical Center of San Jose</td>
<td>0%</td>
<td></td>
<td></td>
<td>9.1%</td>
<td></td>
<td>27.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saint Francis Memorial Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saint Louise Regional Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salinas Valley Memorial Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Francisco General Hospital and Trauma Center</td>
<td>0%</td>
<td></td>
<td></td>
<td>20.0%</td>
<td></td>
<td>36.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Mateo Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santa Clara Valley Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequoia Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td>25.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seton Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Mary’s Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td>22.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL Rose Hospital</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford Hospital and Clinics</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sutter Regional Offices</td>
<td>0%</td>
<td></td>
<td></td>
<td>16.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCSF Medical Center</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA Medical Center San Francisco</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA Palo Alto Health Care System</td>
<td>0%</td>
<td></td>
<td></td>
<td>7.1%</td>
<td></td>
<td>23.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ValleyCare Health System</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington Hospital Healthcare System</td>
<td>0%</td>
<td></td>
<td></td>
<td>10.0%</td>
<td></td>
<td>22.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unattributable Sample</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

139
Introduction and Purpose
My name is Peter Martelli. I am a graduate student at the University of California, Berkeley, working with my faculty advisor, Professor Stephen Shortell, Dean of the School of Public Health. I would like to invite you to take part in my research study, which explores how people understand and use evidence-based management processes, the sources and nature of evidence they compile, and the role that knowledge sharing collaboratives, such as the BEACON Collaborative, play in organizational decision-making.

Procedures
If you agree to participate in my research, I will ask you to complete the attached online survey/questionnaire. The survey will involve questions about your role and attitudes on evidence-based management, evidence sources, best practices, and the BEACON Collaborative. It should take between 15 and 20 minutes to complete, and you can return to it anytime using the same link.

In the future, it is possible that a longitudinal data collection for this study may be conducted. If so, I will contact you to request your participation in the follow-up survey. As always, you may decline to participate at that future time.

Benefits
There is no direct benefit to you anticipated from participating in this study. However, it is hoped that the information gained from the study will help the research community better understand how organizational decision-makers select and use evidence.

Risks/Discomforts
As with all research, there is a chance that confidentiality could be compromised; however, precautions are in place to minimize this risk. (See below for more detail.)

You are always free to decline to answer questions you don’t wish to answer, or to stop the survey at any time for any reason.

Confidentiality
Your study data will be handled as confidentially as possible. If results of this study are published or presented, individual names and other personally identifiable information will not be used.

To minimize the risks to confidentiality, your responses on this online survey are being transmitted under encryption and stored in accordance with the survey platform data management policy. At the end of the survey portion of this study, data from this survey will be downloaded, encrypted and password-protected on the study investigator's digital media. Data stored online at the survey site will then be deleted. Access to personally identifiable files will be limited to investigators listed on the UC-Berkeley Committee for Protection of Human Subjects protocol. No individually-identifiable data will be shared with BEACON, its affiliates or sponsors.

When this study is completed, I will save your data for use in future research done by myself or others. These data will be retained in an encrypted, password-protected indefinitely after the study is over. The same confidentiality measures described above will be taken to protect the study data for any future use.

Compensation
You will not be paid for taking part in this study.

Rights
Participation in research is completely voluntary. You are free to decline to take part in the project. You can decline to answer any questions and are free to stop taking part in the project at any time.
Whether or not you choose to participate, to answer any particular question, or to continue participating in the project, there will be no penalty to you or loss of benefits to which you are otherwise entitled. Whether or not you choose to participate will have no bearing on your membership in BEACON, or the level or quality services you receive from the Collaborative.

Questions
If you have any questions about this research, please feel free to contact me. I can be reached through the UC-Berkeley School of Public Health's Dean's Office at (510) 643-8451 or directly at my email address: martellj@berkeley.edu.

If you have any questions about your rights or treatment as a research participant in this study, please contact the University of California at Berkeley's Committee for Protection of Human Subjects at 510-642-7461, or e-mail subjects@berkeley.edu.

If you agree to take part in the research, please save or print a copy of this page for your future reference, then click on the "Accept" button below to start the survey.

Accept

☐

2. Demographics

Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

- The following survey is intended for its initial recipient only.
  Please complete it based on your own opinions and not on behalf of your institution.

- It consists of a total of 30 questions on 10 pages.

- It will ask you about your decision-making related to strategic management decisions.
  Strategic management decisions involve:
  - Determining long-term goals that produce business value;
  - Aligning competencies with environmental opportunities and constraints;
  - Allocating resources to meet goals;
  - Implementing managerial courses of action to meet goals; and
  - Evaluating implemented courses of action.

- Management decisions focus on outcomes like productivity, absenteeism, and employee or patient satisfaction. Please differentiate these from clinical decisions, which focus on outcomes like fewer nosocomial infections and safer care transitions.

- Your answers are very important from an organizational point of view, so please answer these questions even if you think you don't have direct influence over decisions like these.

First, please answer a few personal demographic questions
(On this page: Questions 1-7 of 30)

Q1. Your organization
If you hold a position representing multiple hospitals, please select your health system at the bottom of this list.

Q2. Your primary job title
☐ Chief Executive Officer
☐ Chief Administrative Officer
☐ Chief Operating Officer
☐ Chief Compliance Officer
☐ Chief Quality/Patient Safety Officer
☐ Senior Legal Counsel
Q3. Years of tenure in your current position

Years of tenure

Q4. Your gender
- Male
- Female

Q5. Your current age range
- <35
- 36-45
- 46-55
- 56-66
- >65

Q6. Your academic education (check all that apply)
- BABS or similar
- BA
- MBA
- MSN
- MD/DO
- PhD/DxPH or similar
- Other degree(s)

Q7. Your most recent degree and the year of completion

Most recent degree

Year of completion (YYYY)

Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

3. Decision Making
Now, a few organizational questions
(On this page: Questions 8-10 of 30)

Q8. Please select the BEACON member that you consider to be most broadly comparable to your organization
Q9. How frequently do you personally consult this organization (by phone, email or mail) as a source of evidence to support strategic management decisions?

- Almost never
- Infrequently
- About half the time
- Frequently
- Almost always

Q10. Please select the BEACON member that you most admire for its capacity and ability to make strategic management decisions

**Timing**

This page timer will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

And a few questions related to your decision-making

(On this page: Questions 11-14 of 30)

Q11. Do you feel that your personal decision-making style is a good fit with the culture of your organization?
- Excellent fit
- Good fit
- Poor fit

Q12. When involved in a strategic management decision, how frequently do you interact with your coworkers?

- Very infrequently
- Infrequently
- About half the time
- Frequently
- Very frequently

Q13. Which of the approaches indicated on the continuum below do you generally prefer when making or supporting strategic management decisions?

- Strongly prefer to copy "tried and true" models
- Strongly prefer to experiment from scratch

Q14. Which of the following items more often leads you to decide in favor of implementing a program or policy new to your organization?

- A critical piece of evidence collected
- The balance of all evidence collected

**Timing**

This page timer will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
4. Evidence Definitions

Answer these questions based on your current understanding without consulting literature, organizational documentation or the opinion of others in your office. There is no right or wrong answer – these questions are to explore the characteristics that you stress as important.

Q15. What is your personal definition of evidence-based management? What processes distinguish an organization that uses evidence from one that doesn’t?

Q16. What is your personal definition of a best practice for management? What are the characteristics of a best practice compared to any other practice?

Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

5. Knowledge Types

Consider a decision to invest resources in a program or policy that is new to your organization. You are committed to making this decision in an “evidence-based” way. But what evidence? Where does it come from? And what does it imply for your organization?

Scholars of knowledge management suggest that evidence conveys five core knowledge types:

- **Know about problems** - related to the nature, formulation, natural history and interrelations of social problems
  ex. definition of safety culture and the mechanisms by which it affects communication in groups

- **Know what works** - describing particular policies, strategies or specific interventions that have been successful
  ex. existing safety culture interventions, such as training sessions, that have produced desired outcomes

- **Know why you would implement** - explaining the relationship between values and policy directions
  ex. symbolic, emotional, ethical and cultural meaning of enacting a safety culture intervention

- **Know how to put into practice** - pragmatic, tacit or expert knowledge about program implementation
  ex. how to practically implement and evaluate an effective safety culture-focused intervention

- **Know who to involve** - what individuals or positions to involve and how to build alliances for action
  ex. internal and external collaborators to advise and support a given safety culture intervention
Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Q17. In your opinion, which knowledge type best suits your organization for each of the following:

You may pick the same type for multiple answers.
If you would like to see the knowledge definitions again, click here.

<table>
<thead>
<tr>
<th>Knowledge type</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>For relative advantage over your competitors</td>
<td></td>
</tr>
<tr>
<td>For compatibility with the decision-making style of your organization</td>
<td></td>
</tr>
<tr>
<td>For making &quot;trialable&quot; decisions, that can be tested out and modified</td>
<td></td>
</tr>
<tr>
<td>For making decisions with easily observable and measurable outcomes</td>
<td></td>
</tr>
</tbody>
</table>

Q18. Now, imagine that there were no barriers to collecting evidence of any type to support your organizational decision-making process.

Describe the ideal distribution of knowledge that would be reflected in this evidence. Use one or more sliding bars to allot 100 percent across the five knowledge types.

Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
6_Sources

This section focuses on the evidence sources you use
(Next 2 pages: Questions 19-23 of 30)

Q19. How much time and money does your organization currently spend to collect and process the evidence necessary to make informed strategic management decisions?

<table>
<thead>
<tr>
<th></th>
<th>Very low</th>
<th>Low</th>
<th>About right</th>
<th>High</th>
<th>Very high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of time</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
<tr>
<td>Amount of money</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
</tbody>
</table>

Q20. Name up to three sources of evidence that you especially trust and frequently seek when making or supporting strategic management decisions. Please list them in order of your preference (1 = highest).

Examples of a source include a particular website, journal, person, or professional society.

Source #1

Source #2

Source #3

Timing

This page timer will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Q21. What kind(s) of knowledge does this source generally provide you?

You may pick one or more types per answer.
If you would like to see the knowledge definitions again, click here

<table>
<thead>
<tr>
<th>Know about problems</th>
<th>Know what works</th>
<th>Know why you would implement</th>
<th>Know how to put into practice</th>
<th>Know who to involve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
<tr>
<td></td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
<tr>
<td></td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
</tbody>
</table>

Q23. If you were pressed for time or money in making or supporting a strategic management decision, what type of knowledge would you want to collect FIRST?

<table>
<thead>
<tr>
<th>Know about problems</th>
<th>Know what works</th>
<th>Know why you would implement</th>
<th>Know how to put into practice</th>
<th>Know who to involve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
<tr>
<td></td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
<td>◯</td>
</tr>
</tbody>
</table>

Timing

This page timer will not be displayed to the recipient.
7. BEACON

Now, let’s make the general decision above more specific.

Consider the management of SEPSIS in your hospital.

Think back to the last strategic management decision made by your organization which involved implementing a new program or policy around sepsis.

Please limit your focus to management decisions, like reorganizing a work process or adding staff, rather than clinical decisions, such as using chlorhexidine wipes as opposed to traditional disinfection methods.

Using the focus above, please answer a few questions about BEACON
(On this page: Questions 24-28 of 30)

Q24. For decisions like this, how central would you say the BEACON collaborative is to your decision process?

- Critical
- Significant
- Equal in importance to other sources
- Minor
- Peripheral

Q25. Do you think your organization uses BEACON as a source of evidence for innovation decisions more, less or the same than your comparison organization, "{(q:\/QID13/ChoiceGroup/SelectedChoices)}"?

- We use it a lot more
- We use it a little more
- We use it about the same
- They use it a little more
- They use it a lot more

Q26. What source or combination of sources do you think complement the knowledge you get from the BEACON Collaborative?

- Academic Literature
- Advocates/ Advocacy Groups
- Databases/ Internal Feedback
- Knowledge Brokers/ Consultants/ Experts
- Information on the Internet/ Media
- Other Conferences/ Professional Exchanges
- Regulations/ Case Law
- Validating Organizations/ External Standards
- Other

- KP Santa Clara Medical Center
- KP South San Francisco Medical Center
- KP Walnut Creek Medical Center
- Marin General Hospital
- Mills-Peninsula Health Services
- Novato Community Hospital
- O’Connor Hospital
- Regional Medical Center of San Jose
- Saint Francis Memorial
- Saint Louise Regional Hospital
- Saint Luke’s
- Saint Mary’s Medical Center
- Saint Rose Hospital
- Salinas Valley Memorial Healthcare System
- San Francisco General Hospital
Only the hospital selected as “admired” in Question #10 appeared among the hospital choices in Questions #26 and #27. No other options marked with a “>>” appeared as selections in the live survey.

| □ Doctors Medical Center, San Pablo          | □ San Mateo Medical Center                  |
| □ Eden Medical Center, San Leandro Hospital | □ Santa Clara Valley Medical Center         |
| □ El Camino Hospital                        | □ Sequoie Hospital                         |
| □ Good Samaritan Hospital                   | □ Seton Medical Center                      |
| □ KP Antioch Medical Center                | □ Stanford Hospital and Clinics             |
| □ KP Hayward Fremont Medical Centers        | □ UCSF Medical Center                       |
| □ KP Oakland Medical Center                | □ VA Palo Alto Health Care System           |
| □ KP Redwood City Medical Center           | □ VA San Francisco Medical Center           |
| □ KP San Francisco Medical Center          | □ Valley Care Health System                |
| □ KP San Jose Medical Center               | □ Washington Hospital Healthcare System    |
| □ KP San Rafael Medical Center             |                                             |

Q27. If your organization didn’t have access to the BEACON Collaborative, what sources or combination of sources do you think would provide the knowledge you get from the BEACON Collaborative?

| □ Academic Literature                       | □ KP Santa Clara Medical Center            |
| □ Advocates/Advocacy Groups                 | □ KP South San Francisco Medical Center   |
| □ Databases/Internal Feedback              | □ KP Walnut Creek Medical Center           |
| □ Knowledge Brokers/Consultants/Experts     | □ Marin General Hospital                   |
| □ Information on the Internet/Media         | □ Mills-Peninsula Health Services         |
| □ Other Conferences/Professional Exchanges  | □ Novato Community Hospital               |
| □ Regulations/Case Law                      | □ O’Connor Hospital                       |
| □ Validating Organizations/External Standards | □ Regional Medical Center of San Jose   |
| □ Other                                    | □ Saint Francis Memorial                  |
| □ Alameda County Medical Center            | □ Saint Louise Regional Hospital          |
| □ Alameda Hospital                          | □ Saint Luke’s                            |
| □ Alta Bates Summit Medical Center         | □ Saint Mary’s Medical Center             |
| □ California Pacific Medical Center        | □ Saint Rose Hospital                     |
| □ Chinese Hospital                         | □ Salinas Valley Memorial Healthcare System |
| □ Community Hospital of the Monterey Peninsula | □ San Francisco General Hospital        |
| □ Doctors Medical Center, San Pablo        | □ San Mateo Medical Center                |
| □ Eden Medical Center, San Leandro Hospital | □ Santa Clara Valley Medical Center       |
| □ El Camino Hospital                        | □ Sequoie Hospital                        |
| □ Good Samaritan Hospital                  | □ Seton Medical Center                     |
| □ KP Antioch Medical Center                | □ Stanford Hospital and Clinics           |
| □ KP Hayward Fremont Medical Centers        | □ UCSF Medical Center                      |
| □ KP Oakland Medical Center                | □ VA Palo Alto Health Care System         |
| □ KP Redwood City Medical Center           | □ VA San Francisco Medical Center         |
| □ KP San Francisco Medical Center          | □ Valley Care Health System               |
| □ KP San Jose Medical Center               | □ Washington Hospital Healthcare System   |
| □ KP San Rafael Medical Center             |                                             |
Q28. Earlier, you described a best practice as: "$\{q://QID25/ChoiceTextEntryValue\}"

Recall the most recent best practice on sepsis that your organization derived from the BEACON Collaborative.

How closely did your organization keep to the original program/policy design when developing its own management strategy?

<table>
<thead>
<tr>
<th></th>
<th>Very Low</th>
<th>Low</th>
<th>High</th>
<th>Very High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance of changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Timing
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

8. Outcomes

This last page asks you about your confidence in decisions.
(On this page: Questions 28-30 of 30)

In your words, evidence-based management entailed: "$\{q://QID24/ChoiceTextEntryValue\}"

For the following dimensions, please describe how confident you would be in a strategic management decision made using this evidence-based management approach. Use the sliding bars to select your level of confidence on a scale of 0 to 100.

Q29. How confident would you be that the decision will lead to beneficial clinical and organizational outcomes?

<table>
<thead>
<tr>
<th></th>
<th>Not at all confident</th>
<th>Less than confident</th>
<th>Don't have strong feelings either way</th>
<th>Generally confident</th>
<th>Fully confident</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
</tbody>
</table>

Confidence in CLINICAL outcomes

Confidence in ORGANIZATIONAL outcomes
Q30. How confident would you be that the decision:

<table>
<thead>
<tr>
<th>_CONFIDENCE RATING</th>
<th>Not at all confident</th>
<th>Less than confident</th>
<th>Don't have strong feelings either way</th>
<th>Generally confident</th>
<th>Fully confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will persist unchanged in the next budget cycle</td>
<td><img src="image1" alt="Confidence Level" /></td>
<td><img src="image2" alt="Confidence Level" /></td>
<td><img src="image3" alt="Confidence Level" /></td>
<td><img src="image4" alt="Confidence Level" /></td>
<td><img src="image5" alt="Confidence Level" /></td>
</tr>
<tr>
<td>Will increase your organization's legitimacy as a high quality provider</td>
<td><img src="image6" alt="Confidence Level" /></td>
<td><img src="image7" alt="Confidence Level" /></td>
<td><img src="image8" alt="Confidence Level" /></td>
<td><img src="image9" alt="Confidence Level" /></td>
<td><img src="image10" alt="Confidence Level" /></td>
</tr>
<tr>
<td>Will yield the same results as those of comparison organizations</td>
<td><img src="image11" alt="Confidence Level" /></td>
<td><img src="image12" alt="Confidence Level" /></td>
<td><img src="image13" alt="Confidence Level" /></td>
<td><img src="image14" alt="Confidence Level" /></td>
<td><img src="image15" alt="Confidence Level" /></td>
</tr>
<tr>
<td>Will not lead to surprising or unneeded outcomes</td>
<td><img src="image16" alt="Confidence Level" /></td>
<td><img src="image17" alt="Confidence Level" /></td>
<td><img src="image18" alt="Confidence Level" /></td>
<td><img src="image19" alt="Confidence Level" /></td>
<td><img src="image20" alt="Confidence Level" /></td>
</tr>
<tr>
<td>Would have changed substantially given more information</td>
<td><img src="image21" alt="Confidence Level" /></td>
<td><img src="image22" alt="Confidence Level" /></td>
<td><img src="image23" alt="Confidence Level" /></td>
<td><img src="image24" alt="Confidence Level" /></td>
<td><img src="image25" alt="Confidence Level" /></td>
</tr>
</tbody>
</table>

**Timing**
This page timer will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Thank you very much for completing this survey.

If you have any questions or suggestions, please contact Peter Martelli at martelli@berkeley.edu or type them in the space below.