 # 06-049

Implementing New Practices: An Empirical Study of Organizational Learning in Hospital Intensive Care Units

Anita L. Tucker
Ingrid M. Nembhard
Amy C. Edmondson

Implementing New Practices:

An Empirical Study of Organizational Learning in Hospital Intensive Care Units

April 19, 2006

Anita L. Tucker*, Ingrid M. Nembhard**, Amy C. Edmondson **

* Wharton School of Business, University of Pennsylvania

** Harvard Business School, Boston, MA 02163

Please send correspondence to the first author at 551 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104-6340; Tel: (215) 573-8742; e-mail: tuckera@wharton.upenn.edu

Acknowledgements

Dr. Jeffrey Horbar, from the Vermont Oxford Network, provided valuable advice and assistance throughout this study, as did Dr. Richard Bohmer and Laura Feldman at Harvard Business School. We gratefully acknowledge their help. We also thank the associate editor and three anonymous reviewers at *Management Science* for invaluable feedback that greatly improved this paper. Participants and reviewers supporting the 2004 INFORMS, 2005 POMS, and the 2005 Academy of Management conferences provided numerous helpful comments. In addition, "Labor lunch" seminar participants at the Wharton School offered many helpful suggestions. We are grateful to Dylan Small and Edward George for suggestions regarding statistical analysis. Jack Hershey, Rachel Croson, Lorin Hitt, and Joel Waldfogel also provided feedback on earlier drafts of this paper. The Harvard Business School Division of Research, the Operations and Information Management Department at Wharton, and the Fishman-Davidson Center for Service and Operations Management provided generous financial support.

Implementing New Practices:

An Empirical Study of Organizational Learning in Hospital Intensive Care Units

Abstract

This paper contributes to research on organizational learning by investigating specific learning activities undertaken by improvement project teams in hospital intensive care units and proposing an integrative model to explain implementation success. Organizational learning is important in this context because medical knowledge changes constantly, and hospital care units must learn if they are to provide high quality care. To develop a model of how improvement project teams promote essential organizational learning in health care, we draw from three streams of related research – best practice transfer (BPT), team learning (TL), and process change (PC). To test the model's hypotheses, we collected data from 23 neonatal intensive care units seeking to implement new or improved practices. We first analyzed the frequency of specific learning activities reported by improvement project participants and discovered two distinct factors: *learn-what* (activities that identify current best practices) and *learn-how* (activities that operationalize practices in a given setting). We then conducted general linear model analyses and found support for three of our four hypothesis. Specifically, a high level of supporting evidence for a unit's portfolio of improvement projects was associated with implementation success. Learn-how was positively associated with implementation success, but learn-what was not. Psychological safety was associated with learn-how, which was found to mediate between psychological safety and implementation success.

Key words: organizational learning, health care, best practice transfer, team learning, process change, practice implementation

1. Introduction

In recent years a series of seminal reports sponsored by the Institute of Medicine have documented significant quality problems in American health care (2001). Problems of overuse and underuse have resulted in misalignment of medical care resources with patients' needs (Chassin et al. 1998). Additionally, research has documented inconsistencies in health care practices that suggest that the quality of treatment received by different patients varies enormously (McGlynn et al. 2003, Wennberg 2002). In the wake of these findings, health care organizations face a growing imperative to identify and spread best practices (McGlynn et al. 2003). However, best practices in health care present an imprecise and moving target. Rapidly changing knowledge (Chassin et al. 1998), disagreement over the efficacy of recommended practices (Cabana et al. 1999), difficulty adjusting to new clinical, organizational and interpersonal routines (Edmondson et al. 2001), and the need to modify practices to be context appropriate (Berta and Baker 2004) hamper organizational attempts to change existing practices so as to reflect current state-of-the-art. Identifying and implementing improved practices is an organizational learning process that requires systematic study.

The aim of the present research is to understand how hospital units, as complex service organizations, successfully implement best practices, or – as is more often the case – "potentially better practices" (Horbar et al. 2001). We define a work practice as a set of interrelated work activities that is informed by a body of knowledge or expertise and repeatedly utilized by individuals or groups to achieve a specified goal (Flynn et al. 1995, Schon 1983, Szulanski 1996). Work practices are thus a subset of the more general concept of organizational processes. Practices necessarily involve people and knowledge; people must apply knowledge to particular situations, and so changing practices requires changing behavior. The ability to predict and foster implementation success (defined as employee commitment to and consistent use of new, better practices (Klein and Sorra 1996)) is vital for health care organizations.

In developing our arguments, we draw from three relevant literatures—best practice transfer, team learning, and process change. First, documented better health care practices often exist that should be transferred across organizations. Some better practices are clinical, such as the consistent use of beta-

blockers for heart attack patients or annual colorectal cancer screening for people over 50. Others are operational, for example, using multidisciplinary rounds that include pharmacists to reduce medication errors. Therefore, the literature on best practice transfer (BPT) informs our question. Second, the literature on team learning (TL) also can provide insight given that implementation of new practices is often led by a team and involves trial and error learning (Bohmer and Edmondson 2001). TL research examines collaborative efforts of cross-disciplinary groups to improve current knowledge through problem-solving and experimentation (Allen 1977, Edmondson et al. 2001, Sarin and McDermott 2003). This stream has emphasized the need for a supportive organizational context to enable behaviors required for learning, and described an iterative process of reflection and action that fuels team learning (Edmondson 2002, Schippers et al. 2003). Third, operations management scholars have modeled process change efforts (PC). This body of research focuses on changes in production processes made in response to technological advances (Carrillo and Gaimon 2000)—a similar phenomenon to healthcare organizations adjusting their practices to reflect advances in knowledge and technology. Scholars suggest that process change will be more successful if managers invest in deliberate—or "induced"—learning activities (Dorrah et al. 1994, Fine 1986, Mukherjee et al. 1998), and that higher levels of knowledge about causal links between process inputs and outputs enable implementation (Bohn 1994).

This paper addresses a question at the nexus of these three literatures: How do improvement project teams identify, adapt and spread new practices in their organizations (Argote et al. 2001)? To answer this question, we empirically examine organizations using improvement project teams to learn. For brevity, we will refer to such teams as "project teams". An organization has learned when its members uniformly modify a work practice to reflect new knowledge (Garvin 2000). In our context, hospital units learn when members consistently use a new practice. Project teams create organizational learning through activities that spread the new practice throughout the unit, such as disseminating research articles about the practice and trial runs. We incorporate an important variable seldom considered in organizational learning research – level of supporting evidence for the practices being learned – and we investigate which learning activities are most important to success (Argote et al. 2001, Mukherjee et al. 1998, Schippers et

al. 2003, Zellmer-Bruhn 2003). Two types of learning activities emerged in data analysis, which we refer to as *learn-what* and *learn-how*. Learn-what involves identifying existing knowledge, and learn-how engages people in knowledge discovery or modification, enabling it to be used in a new context.

In the next section, we further review research on BPT, TL and PC, and develop four hypotheses about implementation success in the health care context. To test these hypotheses, Section 3 describes the data we collected in 23 intensive care units. Section 4 presents the analyses and results, and Section 5 discusses the implications of our findings for future research and for health care and other managers.

2. Implementing New Practices: Theory and Hypotheses

2.1. Best Practice Transfer

Research on best practice transfer (BPT) focuses on organizational efforts to identify and transfer work practices that yield desired results in one location to another, presumably lower-performing, location (Baum and Ingram 1998, Darr et al. 1995, Szulanski 1996, Winter and Szulanski 2001, Zander and Kogut 1995). BPT research shows that when best practices are successfully transferred within and across organizations, performance improves at the recipient unit (Baum and Ingram 1998, Darr et al. 1995, O'Dell and Grayson Jr. 1998, Winter and Szulanski 2001). However, empirical research also reveals the difficulties of transferring knowledge across organizational borders (Argote 1999, Szulanski 1996).

Accordingly, the BPT literature emphasizes practice, organizational, and environmental characteristics that influence transfer success (Argote et al. 2001). Notably, practice codifiability and context-independence enable transfer (Argote 1999, Szulanski 1996, Zander and Kogut 1995). However, many best practices contain a large tacit component that cannot be codified or contain context-specific aspects. One estimate from the semiconductor industry assesses 80% of best practices intended for transfer as non-codifiable (O'Dell and Grayson Jr. 1998). Therefore, researchers have focused on identifying effective mechanisms for transferring tacit knowledge, such as rotating workers between units (Song et al. 2003) or embedding knowledge in technology (Epple et al. 1996). Exact replication also can help transfer tacit knowledge by including all elements of a practice (codified and tacit), even when the

rationale for each element is not understood (Winter and Szulanski 2001). Other research finds that organizational factors such as prior transfer experience (Zander and Kogut 1995) and absorptive capacity (Szulanski 1996) drive transfer success. Additionally, the strength of social ties between individuals in source and target sites are presumed to facilitate transfer by easing interpersonal interaction (Baum and Ingram 1998, Darr et al. 1995).

Applying lessons from the BPT literature, such as the benefit of prior experience, to healthcare organizations faces obvious limitations. BPT studies largely stem from high-volume, manufacturing organizations seeking to replicate high performing plants such that the company can produce identical products in multiple locations. Health care organizations, in contrast, are complex service organizations that face profound variability in patients, resources, staff mix, and facilities across organizations. Furthermore, Berta and Baker (2004) argued that many of the best practices needing to be transferred in health care are composed of tacit knowledge, suggesting a need for exact replication to ensure effective transfer. Yet, exact replication may not be beneficial when organizational culture or context varies significantly (Baum and Ingram 1998), as is often the case with different hospitals and hospital units (Berta and Baker 2004). Thus, for industries like health care delivery that may require substantial adaptation of practice to context, BPT advice may be inadequate, and possibly even counterproductive.

2.2. Team Learning

The team learning (TL) literature has examined how teams manage the interpersonal challenges of developing or implementing new practices within their work groups or organizations (Argyris and Schon 1978, Edmondson 2002, van der Vegt and Bunderson 2005). This work builds on the premise that teams are effective vehicles for organizational learning and performance improvement (Argote et al. 2001, Guzzo and Dickson 1996, Leonard 1995, Nonaka and Takeuchi 1995, Senge 1990). Health care organizations increasingly use multidisciplinary teams to improve unit-wide clinical practices (Horbar et al. 2001, Uhlig et al. 2002). Thus, we examine how project teams within neonatal intensive care units (NICUs) identify, adapt and implement best practices unit-wide (organizational learning).

TL research describes what constitutes successful team processes and which organizational variables

support good team processes (Cohen and Bailey 1997, Guzzo and Dickson 1996). Effective team learning processes include trial-and-error experimentation and collaborative problem-solving (Argyris and Schon 1978, Bunderson and Sutcliffe 2002, Edmondson 1999). When project teams take action to experiment and reflect on their actions—including mistakes that might have been made—to improve future problem-solving cycles, they can be said to be learning (Edmondson 2002, Schippers et al. 2003). Research suggests that willingness to engage in this process depends on a supportive organizational context and, more specifically, a supportive interpersonal climate. Prior studies noted the positive effects of facilitative leadership (Edmondson et al. 2001, Sarin and McDermott 2003), psychological safety (Edmondson 1999) and collective team identification (van der Vegt and Bunderson 2005).

TL research thus has shed light on the process of adapting work practices and identified social and organizational antecedents, but paid less attention to the specific learning activities that promote successful implementation, and ignored variance in the state of knowledge underlying a given set of work practices —a factor of central conceptual and practical importance in health care.

2.3. Process Change

Operations management researchers have investigated process change (PC) within organizations using mathematical modeling and empirical data, primarily from manufacturing settings. The underlying premise is that managers decide how much effort to divert from current production into deliberate quality-enhancing/cost-reducing efforts (Carrillo and Gaimon 2000, Dorrah et al. 1994, Fine 1986). This "induced learning" differs from autonomous learning – the beneficial side-effect of cumulative production volume identified in learning curve studies (Dorrah et al. 1994, Fine 1986, Li and Rajagopalan 1998).

Managers are motivated to allocate resources to *induced learning* when they face a gap between desired quality levels and actual quality levels (Li and Rajagopalan 1998) or face an industry-wide technology change (Carrillo and Gaimon 2000). Researchers assert that even with high levels of induced learning, process change in organizations is difficult—especially across locations and shifts of employees (Tyre and Hauptman 1992). Engaging in induced learning before implementing change, through employee training and off-line experimentation with the new technology, can improve performance

(Carrillo and Gaimon 2000, Tyre and Hauptman 1992). Furthermore, understanding scientific knowledge about inputs and outputs ("conceptual knowledge") and knowing to obtain desired results ("operational knowledge") leads to the highest levels of process change success (Bohn 1994, Mukherjee et al. 1998). Although this research has focused on manufacturing firms (e.g. Carrillo and Gaimon 2000, Li and Rajagopalan 1998, Tyre and Hauptman 1992), identifying which induced learning activities promote process change in health care may extend its relevance to complex service organizations

2.4. An Integrative Theory of Practice Implementation in Health Care

Collectively BPT, TL, and PC research point to key categories of factors to consider in predicting the success of new practice implementation in health care: (1) knowledge properties of work practices, (2) social and organizational factors and (3) induced learning efforts. Within these categories, we identify a set of specific variables, the level of evidence supporting new practices; psychological safety; and two types of learning activities – learn-what and learn-how. Below we develop justification for these four variables and hypotheses related to each.

Evidence: A knowledge-related property of health care practices. The BPT and PC literatures highlight the importance of knowledge-related properties of practices on implementation success (Argote 1999, Bohn 1994). The foremost concern related to potentially better practices in health care is the *level of published evidence* proving that the portfolio of recommended practices will result in better outcomes for patients. Interest in this knowledge-related characteristic often begins with medical training which emphasizes using practices supported by strong research evidence (Institute of Medicine 2001). The quality of research evidence varies, as evidenced by a classification system developed by Muir Gray (1997). At the strong end of the scale, meta-analyses of randomized clinical trials (RCTs) are regarded as the highest level of evidence. At the weak end, anecdotal opinion is considered the lowest form of evidence. In between these extremes, in order of decreasing evidence are a single RCT, non-experimental studies, and qualitative studies.

We propose a link between level of evidence and implementation success. When the evidence for practices are high (i.e. meta-analysis of RCTs), health care professionals can have greater confidence that

the portfolio of practices will produce desired results without exposing patients to inappropriate risks. At an extreme, strong evidence provides nearly irrefutable justification for implementation, and failure to do so would be medical negligence. Similarly, research support produces greater confidence in the proposed new practices, which in turn should stimulate greater implementation motivation. Motivation can also be sparked by the performance benchmark available when other organizations—as shown in published research papers—have better outcomes with the new practices than the organization considering implementation has with its current method (Meyer and Goes 1988). Such motivation in turn fosters implementation success (Szulanski 1996, Zander and Kogut 1995), perhaps because the knowledge that others have succeeded with the new practices encourages persistence with the difficult task of changing behaviors (Meyer and Goes 1988). In contrast, when the level of evidence for a bundle of practices is low, the legitimacy is easily questioned and motivation to comply is low. A literature review of the barriers to implementation of recommended clinical practices found that lack of agreement about the purported benefits was the second-largest reason for physician non-compliance (Cabana et al. 1999). In summary, a portfolio of new practices (defined as the set of practices being implemented concurrently in an organization) will be implemented more successfully if it is supported by high levels of evidence.

HYPOTHESIS 1. *The level of evidence supporting a unit's portfolio of new practices will be positively related to the implementation success of the new practices.*

Learning activities. The activities used by a project team to a) find better practices; b) modify these practices to fit its context; and c) disseminate the better practices within the unit should also influence the implementation success of its portfolio of better practices. Prior work shows that project teams vary in their learning activities (Bunderson and Sutcliffe 2002, Edmondson 1999, van der Vegt and Bunderson 2005), even within the same organization (Edmondson 2002). Moreover, this process variance predicts performance. Teams that iteratively engage in reflection and action perform best (Edmondson 2002, Schippers et al. 2003).

These types of activities have been called reconstruction and external relations/transfer processes (Argote et al. 2001); learning processes (Mukherjee et al. 1998); and internal and external processes

(Cohen and Bailey 1997). We refer to them as simply "learning activities", which indicates that the actions can include both internal (e.g. project group meetings) and external process (e.g. site visits to other NICUs), and can be performed by the project team (e.g. literature reviews) and in conjunction with the larger unit (e.g. disseminating articles or pilot runs). Bundles of learning activities work together to create different types of knowledge, which Garud (1997) has termed know-what, know-how, and know-why. To reflect the fact that hospitals did not "know" a priori what to do, but instead engaged in learning activities to discover and modify best practices, we follow the convention set by Edmondson and Moingeon (1996) and define the terms "learn-what" and "learn-how." Learn-what describes a bundle of learning activities that seek to *identify best practices*. We use learn-how to refer to a bundle of learning activities aimed at discovering the underlying science of a better practice so as to *operationalize the practice* in a target organization. In the present study, therefore, the categories of know-how and know-why are both seen as outcomes of learn-how, because these two types of knowledge are tightly related in the context of health care practices. As illustrated later in this paper, learn-how tends to involve deep understanding of why a practice works, as well as how to carry it out. Even when separable, prior research has shown that know-how is a stronger predictor of goal achievement than know-why (Mukherjee et al. 1998).

Health care organizations face the challenges of both learning what better practices exist and how to implement them. Knowledge in the health care environment fluctuates rapidly, making it difficult to keep abreast of all potentially better practices. Project teams who engage in learn-what activities, such as literature reviews and site visits to other similar organizations, are likely to discover both better practices and implementation tips that increase the likelihood of successful knowledge transfer. Research conducted outside of health care supports this hypothesis. Darr and colleagues (1995) observed that pizza store project teams who conducted site visits to other pizza stores within their franchise learned better ways of organizing their inventory, which they implemented with great success. Therefore, we propose that learn-what fosters implementation success in health care.

HYPOTHESIS 2. *Learn-what will be positively associated with the implementation success of a*

unit's portfolio of new practices.

Transferring best practices across organizational boundaries is not a simple process (Argote et al. 2001). Exact replication is often not possible because of structural or operational differences between organizations (Spear 2005). As a result, successful implementation often requires experimentation with the practice to customize it for the target organization—in other words, learn-how (Klein et al. 2001). Spear's (2005) research on implementing the Toyota Production System in a network of Pittsburgh hospitals found that contextual work differences between two hospitals prevented a project team from one hospital from applying specific fixes developed at the second hospital. Instead, the adopting hospital's project team needed to engage in learn-how to find and tailor its own solutions to the problem. Similarly, prior research on implementation of computerized production support systems (Klein et al. 2001), experimentation (Allen 1977) and operational learning (Mukherjee et al. 1998) shows that project teams were more successful when they engaged in high levels of these activities. Thus, we expect that the implementation success of a practice portfolio will correlate with learn-how.

HYPOTHESIS 3. *Learn-how will be positively associated with the implementation success of a unit's portfolio of new practices.*

Organizational context. Prior research has shown that learn-how, which requires experimentation and collaborative problem-solving, occurs more frequently in supportive organizational contexts (Edmondson 1999, Klein et al. 2001). A supportive organizational context includes several, related aspects, such as managers modeling cooperative problem-solving behavior (Edmondson 1999, Sarin and McDermott 2003), organizational support (e.g. financial and non-financial resources and rewards to accomplish the goals) (Hackman 1990, Klein et al. 2001), and psychological safety (Edmondson 1999). Psychological safety refers to a supportive work unit in which members believe that they can question existing practices, express concerns or dissent, and admit mistakes without suffering ridicule or punishment. As prior research has found that manager coaching and psychological safety are both highly correlated and theoretically causally related (e.g. Edmondson 1999), we consider only psychological safety – the causally proximal variable – in this paper to minimize multicollinearity. Furthermore, studies of practice

change in health care organizations have recognized the importance of psychological safety, despite strong industry-specific factors that inhibit open discussion of errors, such as a rigid, profession-based hierarchy and a tendency to blame individuals for system-generated mishaps (Institute of Medicine 2001).

Project teams conduct their activities—such as dry runs and pilot tests—within the unit context. Therefore, we propose that the level of psychological safety in the unit will impact the project teams' willingness to use learn-how. In summary, psychological safety in the unit should enable learn-how.

HYPOTHESIS 4. *Learn-how mediates the effect of psychological safety on implementation success.*

3. Methods

3.1. Research Context

The research sites for this study, NICUs, are highly specialized hospital units that provide care for premature infants and critically ill newborns. Approximately 9% of newborn babies are admitted annually to a NICU, where they receive care from a multidisciplinary group of neonatologists, nurses, respiratory therapists and other allied health professionals. In recent years, substantial clinical and technological progress has led to higher survival rates as well as to the viability of increasingly smaller and sicker infants. The constant advancements have created a salient need for NICUs to learn what new practices have been discovered, *and* how to implement better practices locally.

To address these learning challenges, in 2002, 44 NICUs in the United States and Canada joined a collaborative to facilitate best practice transfer, and invited us to study their efforts. As a foundation for conducting their improvement work, project teams consisting of 3 to 7 people from these NICUs attended collaborative-wide meetings every six months where industry consultants taught them strategies for driving organizational change, such as rapid cycle improvement, root cause analysis and systems thinking (Horbar et al. 2001). The project teams typically consisted of a neonatologist, nurse practitioner, nurse, nursing unit director, and a respiratory therapist.

At the meetings, project teams from different NICUS worked together to discover better practices related to seven targeted areas. Thirty-four NICUs participated in only one target area, while ten other NICUs had two teams working on two separate target areas. The seven areas were as follows: pain management and sedation (12 NICUs), infection control (7 NICUs), respiratory care (16), maternal and newborn departmental collaboration (5), staff retention (7), family-centered care (3), and discharge planning (5). Multiple better practices were identified for each area, with a total of 93. (Table 1 of our online supplement provides more details on the seven target areas and related practices.)

In addition to formal, collaborative participation in the assigned target area, project teams could also decide to implement best practices from other, non-assigned target areas. Each NICU, on average, implemented or was working on implementing 36.2 practices, with a standard deviation of 10.9. For clarity, we refer to the bundle of practices implemented within a target area as an improvement project. All NICUs engaged in several concurrent improvement projects, with a mean of 3 areas. Three of the NICUs' project teams met weekly, four met every other week, 11 met monthly, and 5 met every other month. Meetings ranged from 30 to 90 minutes.

3.2. Data Collection Process

Our unit of analysis was the NICU. We collected data from the NICUs in the collaborative in three phases. First, we visited four NICUs, where we observed unit functioning and interviewed 23 project team members to better understand NICUs and project teams' improvement efforts. This phase was conducted to facilitate the design of a survey that would be applicable across the full range of hospitals in the collaborative. We thus selected a sample of four NICUs from the population of 44 that included differences in teaching status, size, past collaborative participation and improvement projects (e.g., infection control versus respiratory care). Table 2 of our online supplement provides details about the visited sites. At each site, we toured the unit and interviewed 5 to 7 project team members representing a cross-section of roles using open-ended questions about their improvement projects. Interviews were tape-recorded and lasted 30 to 90 minutes. Upon our return home, we created detailed transcripts of our visits, combining both tape-recorded and

handwritten notes. We extracted from the transcripts a list of learning activities mentioned by project team members during the interviews. We used this list of inductively-generated items in our survey to gather data about project teams' learning activities.

In the second phase, from September 2002 to June 2003, we created and administered a two-part organizational assessment survey. Each hospital's Internal Review Board approved our study before administering the survey. The survey items came from a prior survey completed by the previous NICU collaboratives (Baker et al. 2003), to which we added constructs from BPT and TL literatures, as well as the list of learning activities from our site visit data. Part I of the survey, which gathered information about respondent demographics, psychological safety, and the general success of NICU improvement projects, was completed by everyone. At the end of Part I, the survey wording instructed continuing on to PART II ONLY for individuals who were project team members for one of the projects shown in the customized list in the NICU's cover letter. Part II gathered information related to a single project, which the respondent selected from the customized list and wrote on a line on the survey. We pilot tested the survey with the four NICUs we visited. Descriptive statistics and psychometric tests indicated no need to alter the survey and results triangulated our site visit observations, lending credibility to the validity of the survey instrument. We did not include pilot data in further analyses.

We then invited the 40 remaining, non-pilot site NICUs in the collaborative to participate in the survey. Twenty-three agreed, for a NICU response rate of 58%. We compared participating sites to non-participating sites, and found no significant differences between the two groups on a variety of structural, clinical, and patient acuity measures. For example, there were no differences based on hospital ownership type (not-for-profit, for-profit, government), teaching status, level of severity of the care provided in the NICU, volume of extremely low birth weight babies (ELBW, less than or equal to 1000 grams), number of times the site participated in prior collaboratives (0, 1 or 2), length of stay, percentage of ELBW babies with Apgar scores ≤ 3 one minute after birth, percentage of babies transported from another hospital ("outborn"), average birth weight and gestational age. Table 2 of our online supplement summarizes this information for the pilot, participating and non-participating NICUs.

Between July 2003 and May 2004, we administered the survey to 3130 healthcare professionals in the 23 participating NICUs, receiving a total of 1440 responses for an individual response rate of 46%. The individuals surveyed, a mean of 63 per NICU, represented a wide range of professions including neonatologists, nurses, respiratory therapists, pharmacists, social workers and nursing unit managers. Of the 1440 respondents, 265 individuals also completed Part II of the survey. We refer to the 1175 individuals who completed only Part I of the survey as "non-project" and those 265 who completed Part II as "project" respondents. Because NICUs conducted multiple improvement projects, we deliberately solicited information about multiple projects per NICU to gain a more accurate picture of unit-level learning. We received data about 69 NICU-project combinations and of these, 44 (64%) had more than one respondent. We had data from an average of 3 projects per NICU (st dev = 1.5).

3.3 Study Variables

3.3.1. Perceived Implementation Success

We used three items to measure the extent to which staff perceived that new practices improved NICU care. These items, drawn from Baker and his colleagues (2003), included "The quality improvement projects currently underway in this unit are making a difference in how we do things", "Members of this unit have changed their behaviors to match the practices recommended by the quality improvement projects", and "Our work on these projects has improved the care we give to our neonates." Respondents indicated their agreement with these statements on a seven-point scale (1 = strongly disagree, 7 = strongly agree). Cronbach's alpha was 0.87, surpassing the .70 threshold for reliability suggested by (Nunnally 1967).

3.3.2. Level of Evidence

We used the collaborative's ratings of level of evidence for the different projects. Project teams from different NICUs working together on a target area, in conjunction with a subject-matter expert, reviewed medical literature for each potentially better practice in that area and collectively assigned a Muir Gray (MG) score to each practice (Muir Gray 1997). The MG score indicated the level of scientific evidence linking the proposed practice with desired results, ranging from 1 (a high level of evidence) to 5 (little

evidence). Table 3 in our online supplement provides details on the criteria for MG scores. We averaged the MG scores for all of the practices within each target area to create an average MG score. The average scores from highest evidence base to lowest, were 1.9 for pain and sedation (10 practices), 2.3 for infection control (7 practices), 2.5 for respiratory care (15 practices), 4.0 for staffing (5 practices) and NICU-maternal collaboration (6 practices), 4.8 for family-centered care (27 practices) and 5.0 for discharge planning (23 practices, all scored as 5).

3.3.3. Learning Activities

As previously mentioned, from our interviews we identified 12 learning activities that project teams performed for their improvement projects. The activities included: site visits to other NICUs, literature reviews, distribution of articles to staff, project team meetings, solicitation of staff ideas, use of workbooks of potentially better practices compiled by previous collaboratives, opportunities for staff to provide feedback prior to full implementation, pilot-runs of practices, education sessions with staff, conference calls with other NICUs focused on the same improvement area, and Plan-Do-Study-Act problem-solving cycles. We did not supplement this list with other activities mentioned in the improvement literature such as the use of external consultants who teach improvement techniques (Wysocki 2004) and site visits to successful non-health care organizations (Spear 2005) not mentioned in the interviews with project team members, because the list was developed to reflect the experiences of the study sample.

We listed the 12 learning activities in Part II of the survey and asked project respondents to rate how frequently their project team used each learning activity during the course of the single project they had identified (1 = did not use the practice at all, 5 = used the practice to a very great extent). We conducted an exploratory factor analysis of the 12 learning activities. First, we performed principal component analysis with varimax rotation to identify the number of latent factors. Applying the Kaiser rule of only retaining factors with eigenvalues greater than one, we arrived at a three-factor solution (eigenvalues 5.88, 1.16, 1.03), explaining 67.3% of variance in the survey data. We then performed principal axis factoring stipulating a three-factor solution and used an oblique, promax rotation, in recognition that different learning

activities were likely correlated. Retaining only items with factor loadings of .45 or greater on only one factor in the pattern matrix (which is more easily interpreted than the structure matrix), we found 6 activities loaded uniquely on the first factor, 3 on the second and 1 on the third. Two items failed these criteria and were dropped from subsequent analysis. We reran principal axis factoring with an oblique, promax rotation using only the ten items. In this analysis, the first factor had an eigenvalue of 5.26, the second 1.07, and the third .76. The low eigenvalue of the third factor coupled with the fact that only a single item, project group meetings, loaded on it, motivated us to discard this factor and rerun the analysis with the ten items specifying a two-factor solution. The solution explained 63% of the variance in our data.

The first factor, which we termed *learn-how*, included six activities that allowed for the experiential learning required to adapt better practices to each unit: staff feedback prior to implementation, solicitation of staff ideas, education sessions with staff, project team meetings, pilot-runs (limited short time implementations of a new practice), and the use of problem solving cycles (Plan-Do-Study-Act cycles). Cronbach's alpha for this scale was 0.88. The second factor, which we termed *learn-what*, included four activities for discovering existing best practice information: distributing articles to staff, conducting literature reviews, site visits to other hospitals, and conference calls with other hospitals. Cronbach's alpha was .75. Table 1 lists the items and their factor loadings.

Insert Table 1 about here

3.3.4. Psychological Safety

We used three items modified from Edmondson's (1999) scale to assess psychological safety: "People in this unit are comfortable checking with each other if they have questions about the right way to do something", "The people in our unit value others' unique skills and talents" and "Members of this NICU are able to bring up problems and tough issues." Respondents indicated their agreement with these items on a seven-point scale (1= strongly disagree, 7 = strongly agree). Cronbach's alpha was .74.

3.3.5. Control Variables

Other variables outside of our model could have influenced implementation success. For example, units

with greater staff tenure (number of worked years in the unit) might enjoy greater organizational power and ability to gather necessary resources for improvement projects (Bunderson and Sutcliffe 2002), thereby accelerating the pace and success of their projects. Units whose employees on average worked more hours per week (hours) might have better teamwork and greater motivation to improve unit practices, which increased their efforts and opportunities for success.

To control for alternate hypotheses, we ran a general linear model to test the influence of the following control variables on our dependent measure: tenure, hours, hospital ownership type (e.g. government, not-for-profit, for-profit); hospital teaching status; highest level of NICU care (no major surgery, no cardiac surgery, all surgery); volume of extremely low birth weight (ELBW) babies; number of times participated in a prior collaborative (0, 1, or 2); and the number of potentially better practices from the 2002 collaborative that the unit implemented or was in process of implementing (mean 36.2, std. dev. 10.9). Our model also included validated risk adjusters from the NICU patient population to control for the possible influence of patient acuity: length of stay (LOS); percentage of ELBW babies who had an Apgar score $\leq$ 3 one minute after birth (Apgar); percentage of babies transported into the NICU from another hospital (outborn); average birth weight and gestational age (Richardson et al. 2001). Given our small sample size (N=23), we could not include all of the control variables in our models, and therefore retained only those variables that explained at least 5% of the variance, as measured by the partial eta-squared value. Our final control variables were staff tenure, weekly hours, level of NICU care, number of prior collaboratives, number of practices, outborn, and LOS.

3.4. Use of Independent Data Sources

We designed the study to mitigate the problem of inflated correlations among self-reported variables. First, we linked *non-project nurses'* measures of psychological safety and perceived implementation success with learning activity data (learn-what and learn-how) gathered *only from project team members*, regardless of profession. The one exception was when psychological safety and implementation success were in the same model. In that case, we used project team members' report of psychological safety and nurses' view of

success. We excluded physicians' measures of psychological safety and perceived implementation success because prior research has shown that nurses provide more accurate and conservative reports of organizational climate and performance (Baggs et al. 1999, Leonard et al. 2004). Our analysis replicated the earlier studies. We found that physicians, on average, had higher scores for psychological safety and perceived implementation success than did nurses (6.2 versus 5.6 for psychological safety; and 6.2 versus 5.7 for implementation success; both variables significant at the alpha <.001 level). Although other professional groups, such as respiratory therapists, did not differ significantly from nurses on either variable (5.3 for psychological safety; 5.6 implementation success), we chose to rely only on nurses' reports of psychological safety and implementation success because having multiple professional groups lowered interrater reliability and nurses were the larger population in our sample (n=846 versus 108 therapists). Second, we used a measure for level of evidence that was independent of our survey data. This use of multiple, distinct respondents and sources helped minimize inflated correlations between measures of level of evidence, learning activity, psychological safety and implementation success that might otherwise occur with single respondents (Spector 1994).

4. Analysis and Results

4.1. Aggregation of data

Our dependent variable, perceived implementation success, was conceptually meaningful at the unit level, necessitating aggregation of individual responses to the NICU level (N=23). To ensure aggregation was appropriate, we calculated within-group agreement measures for our composite variables of perceived implementation success, psychological safety, learn-what and learn-how (Bliese 2000, Klein and Kozlowski 2000). Following Bliese's (2000) suggestion, we calculated an interrater agreement score (r_{WG}) for each NICU's measures of each composite variable using the method outlined by James and colleagues (James et al. 1993). The mean r_{WG} was 0.76 for implementation success, 0.73 for psychological safety, 0.66 for learn-what, and 0.72 for learn-how, suggesting adequate agreement for aggregation at the NICU-level (Glick 1985, Zellmer-Bruhn 2003). In addition, we calculated intraclass correlations (ICC[1] and [2]) to test

convergence within NICUs. All ICC[1] values were significant, greater than zero, and had significant ANOVA F-statistics (implementation success ICC[1]=.42, F =8.26, p<. 001; Psy. safety ICC[1]=.30, F=5.33, p<.001, Learn-how ICC[1]=.11, F=1.90, p=.01, Learn-what ICC[1]=.11, F=1.87, p=.01), indicating that aggregation to the NICU level was appropriate (Kenny and LaVoie 1985). Finally, ICC[2] values also suggested reliability of NICU means (.88 for implementation success, .80 for psychological safety, .47 for learn-how and learn-what) (Klein and Kozlowski 2000). Collectively, these measures supported averaging individual responses to create NICU-level variables.

4.2. Findings

Table 2 reports descriptive statistics and correlations. On average, respondents worked 35 hours per week and had been employed in their NICU for 5-9 years. Project teams conducted more activities to create learn-how (3.5 on a 5-point scale) than learn-what (3.2). A paired t-test comparing the means was significant at the .001 level. In addition, learn-what and learn-how were positively correlated (r=.54, p=.009), suggesting that project teams that pursued more learn-what also conducted more learn-how activities. There were significant, positive correlations between psychological safety and learn-how activities (r=.36, p=.090 and learn-what and learn-how with perceived implementation success (learn-what r=.35, p =.099; learn-how r=.35, p=.101). As stated earlier, only non-project nurses provided data for psychological safety and perceived implementation success, while project respondents provided learn-how and learn-what data. Therefore, the correlation between psychological safety and learn-how; as well as learn-how's and learn-what's links to perceived implementation success were not inflated from being from the same data source (Spector 1994).

Insert Table 2 about here

We tested our hypotheses using generalized linear models (GLM) at the unit level (N=23). As mentioned earlier, we controlled for past collaborative participation, level of care provided by the NICU, average years worked in the unit, average hours worked per week, number of implemented best practices,

percentage of outborn babies, and average length of stay. In addition, to control for the alternative hypothesis that organizational support for improvement projects impacted implementation success (Klein et al. 2001), we included an "organizational support" survey measure gathered from project respondents. It consisted of two items related to financial and non-financial support for the improvement projects. Model 1 in Table 3 shows our tests of Hypotheses 1-3 (F= 5.64, p=.007; adjusted R^2 = .73). The model shows support for Hypothesis 1. Higher levels of evidence for a unit's portfolio of improvement projects increased implementation success (B=.-.38, std. error =.13, p=.02; Recall that lower Muir-Gray scores reflected higher levels of evidence). The model also shows that, contrary to Hypothesis 2, engaging in more learn-what activities —such as literature reviews— was not associated with higher levels of implementation success (B= -.19, std. error =.24, p=.46). However, consistent with Hypothesis 3, use of more learn-how activities—such as pilot tests and dry runs—did contribute to greater implementation success (B = .99, std. error = .25, p=.003).

Models 2 through 5 in Table 3 show the results from the test of Hypothesis 4, which predicted that learn-how would mediate the impact of psychological safety on implementation success. Following the three-step procedure for testing a mediating variable (Baron and Kenny 1986, Shaver 2005), we first tested whether psychological safety influenced implementation success. To reduce the correlation due to common respondent, we regressed *non*-project nurses' ratings of implementation success (Y) on improvement project participants' ratings of psychological safety, X (r_{WG} =.76, ICC[1]= .16, F= 2.36, p<.001, ICC[2]=.58). The equation is shown below.

$$Y = \alpha_0 + cX + \varepsilon_0 \quad (1)$$

As Model 2 shows, psychological safety did predict perceived implementation success (F=4.41, B= .64, p=.008). Thus, in Model 3, we continued to test for mediation by regressing learn-how (M) on *non*-project nurses' rating of psychological safety (X′).[1] The regression equation is shown below.

[1] Recall that only improvement project staff provided learn-how data. Therefore, to reduce correlation due to common respondents, we used non-improvement project nurses' ratings of psychological safety.

$$M = \alpha_1 + aX' + \varepsilon_1 \quad (2)$$

With all of the control variables included, the model was not significant (F=2.59, p= 0.072) and had a poor fit (adjusted R^2 = -.067). Therefore, we ran a second model, Model 4, removing control variables with significance levels greater than 0.3. This model fit better (F= 2.59, p=.072, adjusted R^2 = .22), and psychological safety was significant (B=.49, std. error=.24, p=.05), fulfilling mediation test requirements. In the final step, shown in Model 5 we regressed perceived implementation success (Y) on both psychological safety (X′) and learn-how (M). The regression equation is shown below.

$$Y = \alpha_2 + c'X' + bM + \varepsilon_2 \quad (3)$$

To prove mediation, *b* in Equation 3 needed to be statistically significant and *c′* needed to decrease in significance. As shown in Model 7, not only was the model significant (F=6.91, p = .003, adjusted R^2 = .78), but also *b*, the coefficient for learn-how, was significant (B= .78, std. error = .20, p=.004). Psychological safety, *c′*, on the other hand, became insignificant (B= .42, std. error = .27, p=.15), indicating complete mediation. These results therefore support Hypothesis 4 that learn-how mediates psychological safety's effect on implementation success. Moreover, because both psychological safety and implementation success were rated by non-project team nurses, the results provide a conservative test of mediation - given the possible overestimated impact of psychological safety on implementation success.[2]

[2] Recent research by Shaver (2005) suggests that two-stage least squares (2SLS) is a more appropriate method for testing a mediation hypothesis than the method outlined by Baron and Kenney (1986). Shaver's method reduces the problem of correlated error terms in equations 2 and 3 by replacing the actual mediator data—in our case, project team members' rating of learn-how—with either (1) an instrumental variable (a variable that can be a surrogate for the mediator variable, but is theoretically uncorrelated with the outcome variable) or (2) predicted values of the mediator estimated from Equation 2. However, as stated in Kennedy (2003) the challenge of using instrumental variables is finding a suitable instrument. We were unable to find an indicator of learn-how that could not be presumed to also influence perceived implementation success because such activities (for example meetings and/or communications with NICU members) would be visible to staff members thereby potentially influencing measures of implementation success and therefore were not superior to the learn-how measure from our survey data. In addition, as we had different sources of data for learn-how and psychological safety, our measure of learn-how was *less* correlated with psychological safety than a predicted value of learn-how based on psychological safety would be (using project members' data r = .36, two-tailed p=.09 versus predicted values' r =.63, p=.001). Therefore, we concluded that given our data, it was more appropriate to

In summary, we found that success at implementing improvement projects was positively associated with higher levels of supporting evidence for a unit's portfolio of projects and higher levels of learn-how. Our results are also consistent with the notion that learn-how—a trial and error process carried out through dry runs and structured problem solving cycles—is fostered by psychological safety. Learn-how mediated the relationship between psychological safety and implementation success. Learn-what activities had no direct effect on success in this sample.

Insert Table 3 about here

5. Discussion and Conclusions

This research examined factors associated with successful implementation of new practices in health care organizations. We sought to better understand a particular approach to making this happen – the use of improvement project teams as organizational change agents. This approach is consistent with prior theorizing on the importance of teams generally for organizational learning (Edmondson 1999, Senge 1990), but offers a more specific conceptualization of how project teams can be used for deliberate or induced learning. Project teams are useful for this purpose because even well motivated individuals working alone are likely to find it difficult to effect organizational change. Project teams can help overcome barriers to change because they introduce commitment to results and promote engagement in learning by individuals directly responsible for carrying out the organization's work.

The health care organizations studied here explicitly created project teams as temporary learning systems, supplementing current operations. More generally, this approach helps solve a problem identified in prior research: busy health care professionals, consumed by their regular jobs, rarely act upon opportunities to engage in process improvement (Tucker and Edmondson 2003). A similar challenge faces other complex service organizations: when employees face variable and unpredictable customer demands, their ability as individual service providers to figure out how to improve work processes is

use our survey measure of learn-how than to use 2SLS due to the high correlation between psychological safety and predicted learn-how values in 2SLS.

limited by lack of accountability and lack of perspective on the full set of work processes.

At the same time, project teams by themselves do not ensure successful implementation of new practices. Many teams are ineffective (Hackman 1990). Employees may resist changes imposed by project teams. Scholars theorize that resistance derives in part from the "not-invented-here" syndrome, skepticism toward external sources and over reliance on local knowledge (Katz and Allen 1982). Yet, the suspicion that typically surrounds imported practices, particularly for health care professionals trained to be critical, may be countered by evidence. Our results showed that NICUs with practices supported by high levels of scientific evidence experienced greater implementation success. Although our data cannot prove causality, the motivating effects of legitimacy gained through evidence could explain this result. Similarly, Szulanski's (1996) research showed that best practice transfer occurred with greater ease when the recipient regarded the source as trustworthy and reputable. Szulanski inferred that recipients' belief that the source had valuable information motivated their transfer efforts, reducing stickiness. Our findings also support a motivation-based explanation of the relationship between level of evidence and implementation success. In interviews, informants referred to health care providers' drive to do the best for their patients. Therefore, when a new practice comes with a strong evidence base, most caregivers will be motivated to adopt it.

However, motivation is seldom enough. Believing that it is a good idea to implement a new practice is different from knowing how to use the practice and how to use it in your context (Kilo 1999). This requires opportunities to try out and possibly modify the practice. Klein and her colleagues (2001) found that a broad set of policies and procedures—including time to experiment with a new technology—supported the implementation of a new technology. Similarly, our results suggest that the opportunity to become familiar with a portfolio of new practices was a second critical factor for implementation success. We extend Klein et al's (2001) work by identifying the purpose of different policies and practices (learn-what versus learn-how), and teasing apart their distinct roles in implementation success. We found a strong positive relationship between learn-how and implementation success. With activities ranging from seeking staff feedback to pilot tests and dry runs, learn-how appears to facilitate success in three ways. It allows practices to be (1) modified to fit the context, and provides opportunities for staff to (2) experiment with new

practices and to (3) have a role in shaping the practices. Prior research suggests that first-hand experience with changes increases engagement with and commitment to changes, enabling implementation efforts to succeed (Uhlig et al. 2002).

The advantages of learn-how do not ensure it will be used. Learn-how activities such as experimentation and collaborative problem-solving involve interpersonal risk (Edmondson 1999), and so a climate of psychological safety may play an important role in implementing improved practices. We found that in NICUs where staff felt it was safe to ask questions and raise difficult issues, project teams used more learn-how, a result that echoes other research showing that interpersonal climate influences learning and change efforts (Edmondson 1999, Larson et al. 2000, Uhlig et al. 2002). It also extends that research by revealing how interpersonal climate affects practice change—specifically, by enabling learning activities.

This research suggests that psychological safety, learn-how and evidence-based practices support project teams' ability to help their organizations learn – that is, to modify and improve how the organization's work is done. Project teams organize and lead this learning process. Evidence persuades others to join. Learn-how makes new practices work in a specific context, and psychological safety makes willingness to engage in this disruptive process possible. In this way, three factors work together to promote project team effectiveness in an organizational learning process.

5.1. Implications for Research

Our results provide preliminary support for an integrative theory of new practice implementation in hospitals and other complex service organizations. We show that constructs from separate research streams can work together to explain more variance than any of them separately. Drawing upon three literatures enabled us to provide a fuller description of practice implementation in settings where practice transfer requires modification – a description and model that explained considerable variance in implementation success. These findings suggest that other theories may benefit from the incorporation of constructs outside the usual domain of study, and that the variables explored in this research – evidence, learning activities,

and psychological safety – should be considered in future research on new practice implementation. This collection of variables suggests a framework for exploring practice change. Scholars may begin by asking: what are the relevant technical (e.g. knowledge-related), behavioral (e.g. learning approaches) *and* organizational influences (e.g. psychological safety) on the practice under study?

We paid particular attention in this study to examining learning behaviors, and in a relatively novel way. To understand what project teams actually *do* to improve work practices at the unit level, we measured use of specific learning activities (e.g., literature reviews, dry runs) rather than *general* learning behavior (e.g., reflection), which to date has been the dominant approach in the TL literature (Edmondson 1999, van der Vegt and Bunderson 2005). To our knowledge, this is one of few studies to take this approach (e.g., see Mukherjee et al. 1998 for another) and to include so many learning activities (12). In addition, the paper breaks new ground by studying learning activities in the health care context and by empirically distinguishing between activities that constitute learn-what versus learn-how. As discussed above, this research revealed the central role of learn-how – but not learn what – to implementation success. However, we caution against concluding that learn-what is unimportant for practice improvement. The effect of learn-what may have been undetectable in the small sample size (N=23) of organizations. In addition, through their involvement with the collaborative, the project teams in our study had already engaged in considerable learn-what activities at the twice-yearly meetings. It is possible that a larger sample of NICUs, including those not involved with a collaborative, might find learn-what to be a significant predictor of implementation success. Research should also bolster the construct of learn-how and examine if the sequence of learning activities influences implementation success. A lack of longitudinal data prevented investigation of whether it is necessary to engage in learn-what before learn-how. Research to answer this question would certainly aid organizations and project teams seeking to enhance current practices.

In addition to advancing empirical understanding and methodological approaches to the study of learn-what and learn-how, this research contributes a new variable to the literature on organizational learning: evidence-based support for practices. The significance of this construct suggests that future research on organizational learning should consider this knowledge-based attribute of practices in

addition to tacitness (Nonaka and Takeuchi 1995). Finally, our study results suggest that future research may need to consider psychological safety at the organizational level, which was positively associated with implementation success in this study. Most prior work on learning has measured psychological safety at the team level to predict team outcomes, with a study conducted by Baer and Frese (2003) that found organizational-level psychological safety to predict firm innovation as a notable exception. We measured psychological safety at the NICU-level and found that it predicted learn-how in project teams nested in the unit, which mediated a positive relationship between psychological safety and implementation success. However, future studies will have to tease apart whether psychological safety should be modeled as an enabler of learn-how or as a result of successfully implemented learn-how.

5.2 Implications for Practice

This study offers a starting point for health care organizations seeking to improve work practices. We start with the expectation that such organizations must set up project teams to investigate and implement new practices. We then propose that selecting a portfolio of practices supported by evidence is likely to prove motivating to both project team members and others. Further, project teams should provide opportunities for other staff to learn about the new practices through learn-how activities. The success of these steps depends on the cultivation of a climate of psychological safety within the organization.

More generally, we sought to understand factors that facilitate deliberate efforts to improve work practices, thus overcoming barriers to organizational learning. For example, the "not-invented-here" syndrome can lead to resistance of better practices developed outside an organization. Yet scientific evidence that a new practice is better than current practices may help motivate change – at least in a context where human life is at stake. A second barrier is a lack of knowledge or skill in carrying out a potentially better new practice. But, learn-how activities may help overcome this. A third barrier, the interpersonal risk of trying out new behaviors, can be partially reduced by a climate of psychological safety.

5.3. Limitations of the Study

Our results are inarguably most relevant for the health care delivery context; and we cannot assume they

generalize to other industries. Medical care is characterized by readily available knowledge of best practices in the medical literature. In addition, non-profit collaborations create unique opportunities to leverage these existing best practices (e.g., Horbar et al. 2001). Also NICU patients often remain in the unit for several months, and so the organization faces a long-term ongoing, rather than short-term, task. Thus, our results may have the most relevance for other long-term, collaborative work contexts, such as adult intensive care units, rehabilitation facilities, nursing homes, education, child care and social services. Settings that vary on these dimensions, such as temporary organizations (e.g. humanitarian relief efforts) or short-term assignments (e.g. emergency response units) might provide an informative contrast.

The self-selection of NICUs into the collaborative creates another limitation of this study. Other NICUs might not possess the same level of commitment and resources to improve their outcomes, and more dramatic triggers may be required to produce organizational learning. Again, future research could assess how practice improvement works in less committed work units, as well as advancing the model presented in this paper using a larger sample size. Finally, although perceived implementation success has the advantage of being applicable across projects with different outcome measures (with varying measurement units), the absence of objective measures of implementation success is a limitation of this study.

5.4 Conclusion

This paper offered a novel conceptualization of organizational learning in complex service organizations in which implementation of new practices was carried out by project teams, each focused on specific set of practices as a target for learning. These project teams function as engines of learning for the organization. They provide accountability – thereby helping their organizations overcome well-known barriers to learning. Success is by no means assured with this approach, however. We thus also offered preliminary evidence of three factors that may help promote the success of such efforts. The organizations studied here appeared to have greater success when they implemented practices supported by extensive evidence and when project team members engaged in learning activities designed to promote engagement and understanding throughout their unit. Finally, organizational learning appeared more likely in

psychologically safe climates. This paper aims to facilitate future research by identifying and measuring micro-activities of organizational learning, and especially by demonstrating their role in addressing health care's quality improvement challenge.

References

Allen, T. J. 1977. *Managing the flow of technology: Technology transfer and dissemination of technological information within the R&D organization*. MIT Press, Cambridge, MA.

Argote, L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic Publishers, Boston, MA.

Argote, L., D. Gruenfeld, C. Naquin. 2001. *Group learning in organizations*. In Turner, M. E. (ed.), *Groups at Work: Advances in Theory and Research*: 369-411. Erlbaum, New York, 369-411.

Argyris, C., D. Schon. 1978. *Organizational Learning: a theory of action perspective*. Addison-Wesley Publishing Company, Reading, MA.

Baer, M., M. Frese. 2003. Innovation is not enough: Climates for initiative and psychological safety, process innovations, and firm performance. *J. Organizational Beh.* **24** 45-68.

Baggs, J. G., M. H. Schmitt, A. I. Mushlin, P. H. Mitchell, D. H. Eldredge, D. Oakes. 1999. The association between nurse-physician collaboration and patient outcomes in three intensive care units. *Critical Care Medicine* **27** 1991-1998.

Baker, G. R., H. King, J. L. MacDonald, J. D. Horbar. 2003. Using organizational assessment surveys for improvement in neonatal intensive care. *Pediatrics* **111** e419-425.

Baron, R. M., D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51** 1173-1183.

Baum, J. A. C., P. Ingram. 1998. Survival-enhancing learning in the Manhattan hotel industry, 1898-1980. *Management Sci.* **44**(7) 996-1016.

Berta, W. B., G. R. Baker. 2004. Factors that Impact the Transfer and Retention of Best Practices for Reducing Error in Hospitals. *Health Care Management Review* **29**(2) 90-97.

Bliese, P. D. 2000. *Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis*. In Klein, K. J., S. W. J. Kozlowski (eds.), *Multilevel theory, research, and methods in organizations*. Jossey-Bass, San Francisco.

Bohmer, R. M. J., A. C. Edmondson. 2001. Organizational learning in health care. *Health Forum* 32-35.

Bohn, R. E. 1994. Measuring and managing technological knowledge. *Sloan Management Rev.* **36**(1) 61-73.

Bunderson, J. S., K. M. Sutcliffe. 2002. Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Acad. Management J.* **45**(5) 875.

Cabana, M. D., C. S. Rand, N. R. Powe, A. W. Wu, M. H. Wilson, P. A. Abboud, H. R. Rubin. 1999. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* **282**(15) 1458-1465.

Carrillo, J. E., C. Gaimon. 2000. Improving manufacturing performance through process change and knowledge creation. *Management Sci.* **46**(2) 265-288.

Chassin, M. R., R. W. Galvin, The National Roundtable on Quality. 1998. The urgent need to improve health care quality. *Journal of the American Medical Association* **280**(11) 1000-1005.

Cohen, S. G., D. E. Bailey. 1997. What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management* **23**(3) 239-290.

Darr, E. D., L. Argote, D. Epple. 1995. The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Sci.* **41**(11) 1750-1762.

Dorrah, J. R., T. R. Gulledge, N. K. Womer. 1994. Investment in knowledge: A generalization of learning by experience. *Management Sci.* **40**(8) 947-958.

Edmondson, A. C., B. Moingeon. 1996. *When to learn how; and when to learn why: Appropriate organizational learning processes as a source of competitive advantage*. In Moingeon, B., A. C. Edmondson (eds.), *Organizational Learning and Competitive Advantage*. Sage Publications, Thousand Oaks, CA.

Edmondson, A. C. 1999. Psychological safety and learning behavior in work teams. *Admin. Sci. Quart.* **44**(2) 350-383.

Edmondson, A. C., R. Bohmer, G. P. Pisano. 2001. Disrupted routines: Team learning and new technology adaptation. *Admin. Sci. Quart.* **46** 685-716.

Edmondson, A. C. 2002. The local and variegated nature of learning in organizations: A group-level perspective. *Organ. Sci.* **13**(2) 128-146.

Epple, D., L. Argote, K. Murphy. 1996. An empirical investigation of the microstructures of knowledge acquisition and transfer through learning by doing. *Operations Research* **44**(1) 77-86.

Fine, C. H. 1986. Quality improvement and learning in productive systems. *Management Sci.* **32**(10) 1301-1315.

Flynn, B. B., S. Sakakibara, R. G. Schroeder. 1995. Relationship between JIT and TQM: Practices and performance. *Acad. Management J.* **38**(5) 1325-1360.

Garud, R. 1997. *On the distinction between know-how, know-why, and know-what.* In Huff, A., J. Walsh (eds.), *Advances in Strategic Management*: *Vol. 14*: 81-101. JAI Press, Greenwich, CT, 81-101.

Garvin, D. A. 2000. *Learning in Action*. Harvard Business School Press, Boston.

Glick, W. H. 1985. Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research. *Acad. Management J.* **10**(3) 601-616.

Guzzo, R. A., M. W. Dickson. 1996. Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology* **47** 307-338.

Hackman, J. R. 1990. *Groups that work (and those that don't).* Jossey-Bass, San Francisco, CA.

Horbar, J. D., J. Rogowski, P. Plesk, P. Delmor, W. H. Edwards, J. Hocker, A. D. Kantak, P. Lewallen, W. Lewis, E. Lewit, C. J. McCarroll, D. Mujsce, N. R. Payne, P. Shiono, R. F. Soll, K. Leahy, J. H. Carpenter. 2001. Collaborative quality improvement for neonatal intensive care. *Pediatrics* **107**(1) 14-22.

Institute of Medicine. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academy Press, Washington, D.C.

James, L. R., R. G. Demaree, G. Wolf. 1993. Rwg: An assessment of within group interrater agreement. *Journal of Applied Psychology* **78**(2) 306-339.

Katz, R., T. J. Allen. 1982. Investigating the Not Invented Here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R&D project groups. *R & D Management* **12**(1) 7-19.

Kennedy, P. 2003. *A guide to econometrics* (5th ed.). Blackwell Publishing, Malden, MA.

Kenny, D. A., L. LaVoie. 1985. Separating individual and group effects. *Journal of Personality and Social Psychology* **48**(2) 339-448.

Kilo, C. M. 1999. Improving care through collaboration. *Pediatrics* **103**(supp) 203-214.

Klein, K. J., J. S. Sorra. 1996. The challenge of innovation implementation. *Acad. Management J.* **21** 1055-1080.

Klein, K. J., S. W. J. Kozlowski. 2000. From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods* **3**(3) 211-236.

Klein, K. J., A. B. Conn, J. S. Sorra. 2001. Implementing computerized technology: An organizational analysis. *Journal of Applied Psychology* **86**(5) 811-824.

Larson, E. L., E. Early, P. Cloonan, S. Sugrue, M. Parides. 2000. An organizational climate intervention associated with increased handwashing and decreased nosocomial infections. *Behavioral Medicine* **26**(1) 14-22.

Leonard, D. 1995. *Wellsprings of knowledge: Building and sustaining the sources of innovation*. Harvard Business School Press.

Leonard, M., A. Frankel, T. Simmonds, K. Vega. 2004. *Achieving Safe and Reliable Healthcare: Strategies and Solutions*. Health Administration Press, Chicago, Il.

Li, G., S. Rajagopalan. 1998. Process improvement, quality, and learning effects. *Management Sci.* **44**(11) 1517-1532.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, E. A. Kerr. 2003. The quality of health care delivered to adults in the United States. *New England J. Medicine* **348**(26) 2635-2645.

Meyer, A. D., J. B. Goes. 1988. Organizational assimilation of innovations: A multilevel contextual analysis. *Acad. Management J.* **31**(4) 897-923.

Muir Gray, J. A. 1997. *Evidence-Based Healthcare*. Churchill Livingstone, New York.

Mukherjee, A., M. Lapre, L. V. Wassenhove. 1998. Knowledge driven quality improvement. *Management Sci.* **44**(11) S35-S49.

Nonaka, I., H. Takeuchi. 1995. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, Cambridge, UK.

Nunnally, J. C. 1967. *Psychometric Theory*. McGraw-Hill, New York.

O'Dell, C., C. J. Grayson Jr. 1998. *If only we knew what we know: The transfer of international knowledge and best practices*. The Free Press.

Richardson, D. K., J. D. Corcoran, G. J. Escobar, S. K. Lee. 2001. SNAP-II and SNAPPE-II: Simplified newborn illness severity and mortality risk scores. *J. Pediatrics* **138**(1) 92-100.

Sarin, S., C. McDermott. 2003. Effect of team leader characteristics on learning, knowledge application, and performance of cross-functional new product development teams. *Decision Sci.* **34**(4) 707-739.

Schippers, M. C., D. N. Den Hartog, P. L. Koopman, J. A. Wienk. 2003. Reflexivity and diversity in teams: The moderating effects of outcome interdependence and group longevity. *J. Organizational Beh.* **24** 779-802.

Schon, D. A. 1983. *How professionals think in action*. Basic Books, New York, NY.

Senge, P. M. 1990. *The fifth discipline: The art and practice of the learning organization*`. Doubleday Currency, New York.

Shaver, J. M. 2005. Testing for mediating variables in management research: Concerns, implications and alternative strategies. *Journal of Management* **31**(3) 330-353.

Song, J., P. Almeida, G. Wu. 2003. Learning by hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Sci.* **49**(4) 351-365.

Spear, S. J. 2005. Fixing Health Care from the Inside, Today. *Harvard Bus. Rev.* **83**(9) 78-91.

Spector, P. E. 1994. Using self-report questionnaires in OB research: A comment on the use of a controversial method. *J. Organizational Beh.* **15** 385-392.

Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management J.* **17**(Winter Special Issue) 27-43.

Tucker, A. L., A. C. Edmondson. 2003. Why hospitals don't learn from failures: Organizational and psychological dynamics that inhibit system change. *California Management Rev.* **45**(2) 1-18.

Tyre, M. J., O. Hauptman. 1992. Effectiveness of organizational responses to technological change in the production process. *Organ. Sci.* **3**(3) 301-320.

Uhlig, P. N., J. Brown, A. K. Nason, A. Camelio, E. Kendall. 2002. System Innovation: Concord Hospital. *Joint Commission Journal of Quality Improvement* **28**(12) 666-672.

van der Vegt, G. S., J. S. Bunderson. 2005. Learning and performance in multi-functional teams: the importance of collective team identification. *Acad. Management J.* **48**(3) 532-547.

Wennberg, J. E. 2002. Unwarranted variations in healthcare delivery: Implications for academic medical centres. *British Medical Journal* **325**(7370) 961-964.

Winter, S. G., G. Szulanski. 2001. Replication as strategy. *Organ. Sci.* **12**(6) 730-743.

Wysocki, B. J. 2004. To fix health care, hospitals take tips from factory floor, The Wall Street Journal: A1. New York, NY.

Zander, U., B. Kogut. 1995. Knowledge and the speed of transfer and imitation of organizational capabilities: An empirical test. *Organ. Sci.* **6**(1) 76 –92.

Zellmer-Bruhn, M. E. 2003. Interruptive events and team knowledge acquisition. *Management Sci.* **49**(4) 514-528.

TABLE 1. Principal Axis Factor Analysis (Oblique Promax Rotation) of Activities (N=265)

Variables	Learn-how	Learn-what
Opportunity for staff feedback before implementation	**.97**	-.17
Solicitation of staff ideas from staff	**.74**	.07
Education session with staff	**.73**	.02
Project team meetings	**.68**	.07
Limited time pilot project	**.68**	.03
Problem solving cycle (PDSA)	**.60**	.16
Distribution of articles to staff	-.02	**.80**
Literature Reviews	.08	**.72**
Site visits to other NICUs	-.07	**.49**
Conference calls with other NICUs working on similar projects	.32	**.45**
Variance explained	5.3	1.1
Variance explained (%)	52.6	10.7
Reliability (Cronbach's Alpha)	.88	.75

TABLE 2 Means, Standard Deviations, and Pearson Correlations for Study Variables (N=23)

Construct	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12
1. Psy Safety (7-pt)	5.6	0.4												
2. Learn-how (5-pt)	3.5	0.5	.36^											
3. Learn-what (5-pt)	3.2	0.4	.01	.54**										
4. Organizational Support (5-pt)	3.8	0.6	.09	.05	.14									
5. Level of Evidence (1=high, 5=low)	3.2	0.9	.32	.16	-.29	-.29								
6. Perceived Imp-lamentation Success	5.6	0.7	.60**	.35^	.35^	-.11	-.12							
7. Years exp. in NICU	5-9	0.5	.06	-.35^	-.35	-.09	.18	.13						
8. Ave hrs. worked/week	35.0	6.8	-.33	.11	-.05	.10	-.36	-.43*	-.04					
9. No. of prior collaboratives	0.8	0.7	.01	.14	.09	.32	.01	.02	.07	-.18				
10. No. of practices	36.2	10.9	.34	.32	.08	-.44*	.39	.22	-.04	-.03	-.04			
11. Level of NICU care	2.4	0.7	-.02	-.11	.12	.23	-.32	.02	-.11	.17	.07	.15		
12. % of babies outborn	29.5	33.9	.33	-.01	.12	.27	-.07	.31	-.15	-.30	.09	.02	.32	
13. Length of stay (days)	74.2	12.5	-.22	-.09	-.08	-.56**	-.11	-.10	-.08	.42*	-.24	.32	.25	-.32

^ = p<=.10; * = p < 0.05 (two-tailed); ** = p < .01 (two-tailed); *** = p < .001 (two-tailed)

TABLE 3. Generalized Linear Model (GLM) Regression Results. (B values with std. error in parentheses)

Model	1	2	3	4	5
Dependent Variable	Imp. Success	Imp. Success	Learn-how	Learn-how	Imp. Success
Intercept	2.7 (2.0)	2.98 (2.16)	2.99 (2.81)	1.08 (1.65)	1.18 (1.89)
Control Variables					
Past collaboratives [a]	F=3.15^	F=.87	F=.21	N/A	F=1.84
None	.62 (.27)*	.34 (.32)	-.21 (.44)		.38 (.28)
One	.30 (.26)	.12 (.30)	-.07 (.41)		.09 (.26)
Level of NICU care [b]	F=2.53	F=3.37^	F=.12	N/A	1.25
No major surgeries	-.58 (.48)	-.65 (.55)	-.15 (.72)		-.35 (.46)
No cardiac surgeries	.35 (.2)	.47 (.22)^	.13 (.30)		.27 (.19)
Ave. years in this unit	.65(.17)**	.40 (.17)*	-.41 (.22)^	-.37 (.18)*	.56 (.16)**
Ave. hours work/week	-.08(.02)***	-.05(.02)*	.03 (.02)	.02 (.02)	-.06 (.02)**
No. of new practices	.02 (.01)	.01 (.01)	.01 (.02)	N/A	.01 (.01)
Pct. of babies outborn	.004 (.003)	.004 (.003)	-.002 (.004)	N/A	.004 (.003)
Ave. length of stay	.000 (.01)	-.003 (.02)	-.02 (.02)	N/A	.002 (.01)
Predictors					
Organization Support	.08 (.21)	-.11 (.25)	-.11 (.33)	N/A	-.06 (.21)
Level of evidence	-.38 (.13)*	-.27 (.13)^	.08 (.18)	.11 (.11)	-.38 (.11)**
Learn-what (LW)	-.19 (.24)				
Learn-how (LH)	.99 (.25)**				.79 (.20)**
Mediator					
Psychological safety		.64 (.19)**	.42 (.40)	.49 (.24)*	.42 (.27)
Degrees of freedom error	9	10	10	18	9
F statistic	5.64**	4.41*	.86	2.59	6.91
p-value	.007	.013	.59	.07	.003
R^2	.89	.84	.52	.37	.91
Adjusted R^2	.73	.65	-.07	.22	.78

^ < 0.10, * p < 0.05; ** p < .01; *** p < .001

[a] Past collaborative participation: indicator variable with the referent of participation in two prior collaboratives.

[b] Level of care provided by the NICU: indicator variable with the referent of all types of surgeries performed.