



Key Drivers of Successful Implementation of an Employee Suggestion-Driven Improvement Program

Citation

Tucker, Anita L., and Sara J. Singer. "Key Drivers of Successful Implementation of an Employee Suggestion-Driven Improvement Program." Harvard Business School Working Paper, No. 12-112, June 2012.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:9211198>

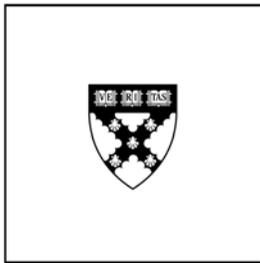
Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)



Key Drivers of Successful Implementation of an Employee Suggestion-Driven Improvement Program

Anita L. Tucker
Sara J. Singer

Working Paper

12-112

June 25, 2012

Copyright © 2012 by Anita L. Tucker and Sara J. Singer

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Key Drivers of Successful Implementation of an Employee Suggestion-Driven Improvement Program

Abstract

Service organizations frequently implement improvement programs to increase quality. These programs often rely on employees' suggestions about improvement opportunities. Organizations face a trade-off with such suggestion-driven improvement programs. On one hand, the improvement literature recommends that managers focus organizational resources on surfacing a large number of problems, prioritizing these, and selecting a small set of high priority ones for solution efforts. The theory is that soliciting a large number of ideas from employees will surface a set of higher priority problems than would have been identified with a less extensive search. Scarce organizational resources can be allocated to resolving the set of problems that provide the greatest improvement in performance. We call this an "analysis-oriented" approach. On the other hand, managers can allocate improvement resources to addressing problems raised by frontline staff, regardless of priority ranking. This "action-oriented" approach enables more resources to be spent on resolving problems because prioritization receives less attention. To our knowledge, this tradeoff between analysis and action in process improvement programs has not been empirically examined. To fill this gap, we randomly selected 20 hospitals to implement an 18-month long employee suggestion-driven improvement program; 58 work areas participated. Our study finds that an action-oriented approach was associated with higher perceived improvement in performance, while an analysis-oriented approach was not. Our study suggests that the analysis-oriented approach negatively impacted employees' perceptions of improvement because it solicited, but not act on, employees' ideas. We discuss the conditions under which this might be the case.

1. Introduction

Process improvement programs can create a competitive advantage for organizations (Loch et al., 2003). Organizations often utilize frontline staff's suggestions about improvement opportunities as an

input for process improvement programs (Imai, 1986, Parker et al., 1997). However, despite their intuitive appeal and popular following, employee suggestion-based improvement programs have produced mixed empirical results (Harlos, 2001). For example, three studies of the employee suggestion-based improvement program that we study in this paper found that the program led to positive change in organizational climate (Benning et al., 2011, Frankel et al., 2008, Thomas et al., 2005), while another two studies found that basically the same program had a negative impact on climate (Benning et al., 2011, Singer et al., forthcoming). This paper attempts to resolve these conflicting findings by identifying the conditions under which the program was most effective.

In particular, we examine a tradeoff in how managers implement the program. The tradeoff involves choosing whether to allocate resources to identify and prioritize problems with the goal of solving the highest priority problems versus to allocate resources to resolve problems with the goal of addressing as many as possible. The tradeoff arises because identifying and prioritizing problems requires resources, which otherwise could be used to address already known problems (Wachter, 2009). Furthermore, suggestion-based programs can surface more problems than the organization can solve with its limited human and financial resources (Frankel, et al., 2008, Reppenning and Sterman, 2002). The resource shortfall creates an allocation decision between resolving a smaller number of higher priority problems and resolving a larger number of lower priority problems.

Managers can focus resources on generating as many improvement ideas as possible, even if the organization doesn't have enough problem-solving capacity to address all of them. This approach may be beneficial because having a large number of ideas enables analysis of frequency and severity (Leape, 2002). Using these data, the most important problems can be identified and prioritized for solution efforts (Bagian et al., 2001). This analysis-oriented approach enables the organization's resources to be used for resolving the highest priority problems. Furthermore, this approach may increase managerial confidence in the program because problems that are addressed emerge from a large number of ideas submitted by a cross section of employees rather than from a potentially idiosyncratic concern raised by one employee.

In contrast, managers can focus resources on solving problems rather than on surfacing and prioritizing them. This can be accomplished by addressing issues as they emerge with little emphasis on using prioritization scores as a filtering mechanism (Johnson, 2003, Repenning and Serman, 2002). An action-oriented approach may be beneficial because prioritizing problems consumes substantial resources that could instead be diverted to addressing problems (Wachter, 2009). In addition to the direct impact, an action-orientation may have an added benefit of increasing employees' motivation. Prior research has found that frontline staff more willingly engage in the discretionary behaviors required for process improvement if they believe that managers will act on their ideas (Gandhi et al., 2005, Morrison and Phelps, 1999). Furthermore, research on accidents has found that small problems combine to cause major accidents, and that it is difficult to know in advance which problems will be involved in the next accident (Reason, 1990). This research calls into question the possibility of using analysis to accurately prioritize problems with regard to safety. Finally, solving problems as they are identified limits the number of unsolved, identified problems to align with available problem solving resources. This may prevent the unproductive firefighting syndrome where a long queue of unresolved issues creates pressure to patch problems rather than to remove underlying causes (Bohn, 2000). In summary, resources may be better spent solving known problems rather than identifying a large set of problems and limiting solution to the small subset with the highest expected impact (Johnson, 2003).

This paper tests hypotheses on this tradeoff between an analysis-oriented versus an action-oriented approach using data from service organizations' implementation of a popular employee suggestion program. The program is commonly referred to as "Management-By-Walking-Around" (MBWA) (Peters and Waterman, 1982). MBWA involves senior managers visiting their organization's frontlines to observe and talk with employees while they do their work. The intention is that managers and frontline staff will work together to identify and resolve obstacles to efficiency, quality or safety (Frankel et al., 2003, Luria and Morag, 2012, Peters and Waterman, 1982). The program has been adopted in many settings including semiconductor manufacturing (Luria and Morag, 2012), schools (Sagor and Barnett, 1994), and hospitals (Benning, et al., 2011, Benning, et al., 2011, Frankel et al., 2003). It is theorized to

improve performance through two levers. First, it is a visible sign of managers' commitment to the improvement program, which in turn motivates employee engagement in the discretionary activities required for improvement (Beer, 2003, Luria and Morag, 2012). Second, improvement occurs when problems raised by frontline employees are addressed (Frankel, et al., 2003, Pronovost et al., 2004).

We examine the tradeoff between analysis and action among 58 work groups in 20 organizations that were randomly selected to conduct an MBWA-based program. We exploit variation in their implementation and performance to better understand the approach that resulted in positive performance. In doing so, this paper answers the call for research to examine the drivers of successful (or unsuccessful) implementation of improvement programs (Dixon-Woods et al., 2011, Nembhard et al., 2009).

We used a perceptual measure of performance, *perceived improvement in performance (PIP)*, as our outcome variable. Four reasons support this choice. First, MBWA-based programs have been shown to impact organizational climate, which is an important driver of performance (Benning, et al., 2011). Second, perceptions of climate are an important outcome in their own right because they influence employee behaviors, which in turn impact objective measures (Zohar et al., 2007). Third, the objective of this MBWA-based program was safety, which is challenging to measure objectively because by definition safety implies the absence of problems (Gaba 2003). Employee perceptions are a reliable source of information about process quality because employees are close to the work and know if system failures are becoming less frequent. For example, research has found that nurses' perceptions of safety culture are associated with safety outcomes, such as mortality, readmissions, and other objective performance outcomes, such as length of stay (Hansen et al., 2010, Hofmann and Mark, 2006, Huang et al., 2010, Rogers et al., 2010, Singer et al., 2009)). Fourth, the use of a perceptual measure was necessary as the hospitals were unwilling to share confidential data about clinical outcomes or safety incidents with us.

We find that an action-oriented approach to the MBWA-based program was associated with improved PIP. The action-oriented approach was manifested by a higher percentage of solved problems that were considered "easy" to solve, which enabled more problems to be addressed with the same set of human and organizational resources. In contrast, the analysis-oriented approach, as characterized by identifying and

solving higher priority problems, was not associated with improved PIP. Our results thus call into question the general applicability of advice recommending an analysis-oriented approach where a large number of problems are identified, prioritized and only the highest priority ones are selected for resolution. We propose that using an analysis-oriented approach may not be beneficial when the problem landscape has many small to medium priority problems and few high priority ones, as we believe is the case in our study.

2. Analysis-Oriented versus Action-Oriented Approaches to Process Improvement

Process improvement programs can yield variable results, with some organizations showing success and others failure (Choi and Behling, 1997). This study examines variation in implementation success for one MBWA-based program. We examine the impact of an analysis-oriented versus action-oriented approach to implementing the MBWA-based program during each of two key steps in the process improvement process: problem identification and problem selection. We first consider problem identification.

2.1. The Effect of Implementation Approach during Problem Identification

According to innovation theory, an analysis-oriented approach to problem identification is preferable. Generating a larger set of ideas, it is argued, will result in a higher value for the best idea (Girotra et al., 2010). This outcome is beneficial in situations like product development tournaments where significant investment is required to bring an idea through the entire development process. Organizations typically do not have enough resources to simultaneously develop multiple ideas and therefore selecting the highest potential impact ideas is important. Furthermore, success can be achieved through one innovative idea that is forwarded through the process to become a new blockbuster product. Thus, in the product development context, success ultimately depends on finding and developing one or two blockbuster ideas (Terwiesch and Ulrich, 2009).

Similarly, service organizations have limited resources for process improvement. Identifying the problems whose resolution offers the highest potential benefit and assigning these problems the highest priority is a key recommendation in the improvement literature (Juran and Gyrna, 1980). The rationale is that organizations benefit from surfacing a large number of problems and identifying higher value problems that, if selected and resolved, would yield the largest performance increases (Bagian et al., 1999). We propose that a work area's ability to identify high priority problems, a characteristic of an analysis-oriented approach, positively impacts PIP.

Hypothesis 1a (H1a): The greater the value of the highest-valued identified problem, the greater a work area's perceived improvement in performance.

An action-oriented approach emphasizes solving multiple identified problems. The underlying assumption is that organizations benefit by solving as many problems as possible, no matter how small. This approach thus encourages identification of problems that are easy-to-solve because this enables the resolution of more problems given the same level of organizational resources (Bohn, 2000). Organizations have been able to improve performance by identifying simple problems that can be easily solved (Thompson et al., 2003), which suggests that this is a viable approach to process improvement. An action-oriented approach is typical of lean production systems, which encourage employees to bring issues that interfere with production to their manager's attention by pulling an andon cord, even if the employee and manager can quickly resolve the issue (Liker, 2004). Lean organizations value identification of problems, even those that are easy-to-solve, because managers believe that solving them leads to improvement. Thus, we propose that a work area's ability to identify easy-to-solve problems, a characteristic of an action-oriented approach to process improvement, positively impacts PIP.

Hypothesis 1b (H1b): The greater the percentage of identified problems that are easy-to-solve, the greater a work areas' perceived improvement in performance.

2.2. The Effect of Implementation Approach during Problem Selection

We now develop prescriptive hypotheses about which types of problems should be selected for resolution efforts if the goal is to increase PIP. A significant challenge of suggestion-based programs is that they can surface more problems than an organization can solve given its limited human and financial resources (Bohn, 2000, Frankel, et al., 2008, Repenning and Sterman, 2002). Organizations face a tradeoff between spending resources on the process improvement task of selecting problems (e.g., the activities to decide which problems to address) versus on the task of solving problems. On one hand, managers can focus resources on prioritizing among large numbers of improvement ideas with the goal of solving the highest priority problems. For example, a priority score for each problem can be determined by multiplying its potential severity by its frequency of occurrence (Bagian, et al., 1999, Frankel et al., 2005). Using these calculations, higher priority problems can be detected and selected for solution efforts (Bagian, et al., 2001). Selection continues until problem solving resources are depleted or there are no more unsolved problems that scored high enough to warrant solution. This is characteristic of an analysis-oriented approach. This approach is similar to innovation tournaments, in which success is driven by a few high impact ideas that get through the development process and become blockbuster products (Terwiesch and Ulrich, 2009).

Service organizations commonly use an analysis-oriented approach to problem selection, as illustrated by the widespread adoption of incident reporting systems as a primary tool for improving safety in aviation and healthcare (Milch et al., 2006). The Aviation Safety Reporting System operated by NASA and the safety improvements it has provoked has been credited in part for the 65% reduction in fatal airplane crashes to one in 4.5 million in the decade ending 2007. Incident reporting systems are a key component of hospitals' patient safety systems, and have been endorsed by the Institute of Medicine (1999) and the Department of Health and Human Services (Levinson, 2010). A survey of 2,050 US

hospitals found that 98% had incident reporting systems (Farley et al., 2008). In incident reporting systems, hospitals encourage staff to report situations that did or could have led to patient harm. These reports are analyzed for trends and the most significant issues are resolved (Milch, et al., 2006, Nuckels et al., 2007). The program's focus tends toward extensive identification, analysis and prioritization; not all problems receive resolution efforts (Ramanujam et al., 2008, Wachter, 2009).

The primary benefit of an analysis-oriented approach is that the process ensures that scarce problem solving resources will be allocated to the set of problems predicted to yield the greatest impact. A consequence is that problems with low priority scores are identified, but not resolved. This may be beneficial because discarding low priority problems prevents the queue of unsolved problems from growing unmanageably long (Bohn, 2000). We hypothesize that an analysis-oriented approach to problem selection, as indicated by solving the highest priority problems, is associated with improved PIP.

Hypothesis 2a (H2a): Work areas that resolve a higher percentage of high priority problems will have more perceived improvement in performance than work areas that solve a lower percentage of high priority problems.

On the other hand, managers can maximize resources allocated to resolving problems by addressing problems as they are identified rather than first conducting prioritization analysis. An action-oriented approach to selecting problems to resolve is characterized by spending the majority of improvement resources addressing problems rather than analyzing them. The benefit of an action-oriented approach stems from the cumulative payoff of solving many small-scale problems. It is akin to the long tail principle that derives from studies of the retail sector. Scholars suggest that on-line retailers can achieve substantial profits by selling a large number of niche items that have small sales volume individually, but collectively add up to a large cumulative sales volume (Anderson, 2004, Brynjolfsson et al., 2011). The organizational learning literature discusses a similar tradeoff between devoting resources to exploring

new opportunities that have a high potential payoff versus exploiting existing, known opportunities with incremental payoffs (March, 1991).

We propose that an action-oriented approach is fruitful for increasing PIP for four reasons. First, research on improvement has found that employees stop identifying and trying to resolve issues themselves when senior managers fail to take action on known problems (MacDuffie, 1997). Conversely, employees are more willing to engage in the discretionary behaviors required for improvement if they believe that managers will act on, rather than simply study, their ideas (Gandhi, et al., 2005, Morrison and Phelps, 1999). Thus, there is a hidden psychological cost of the analysis-oriented approach of asking employees for ideas that are subsequently disregarded. Not solving ideas that the organization has solicited may lead to cynicism and a lack of participation in future efforts (Clarke, 1999, Tucker, 2007).

Second, the pressure of having identified more problems than can be resolved can create an unproductive firefighting culture where problems are patched rather than fully resolved, so their underlying causes are never addressed (Bohn, 2000). The queue of problems awaiting resolution can be shortened by addressing easy-to-solve problems first.

Third, research has found that major accidents typically result from an unpredictable combination of small magnitude problems rather than from a single large magnitude problem (Perrow, 1984, Reason, 2000). According to the “Swiss Cheese Theory,” accidents occur when errors remain unaddressed because of latent weaknesses—represented by holes in a piece of cheese—in multiple defensive layers (each represented by a slice of cheese) and reach the patient (Cook and Woods, 1994, Reason, 2000). It therefore can be beneficial to resolve what appear to be low priority problems because it is difficult to predict which problems will align to contribute to the next accident (Perrow, 1984).

Fourth, improvement requires resolving problems, but an analysis-oriented approach can expend significant managerial and financial resources on analyzing rather than resolving problems (Johnson, 2003). For example, the U.S. Aviation Safety Reporting System spent an average of \$100 per report on analysis alone (Johnson, 2003). Similarly, a hospital physician estimated that his institution spent \$1.6 million per year managing incident reports, with the bulk spent on analysis rather than resolution

(Wachter, 2009). With an action-oriented approach, this money would be spent on addressing problems rather than analyzing them. For these reasons, we hypothesize that an action-oriented approach—characterized by resolving many easier-to-solve problems—will be associated with improved PIP.

Hypothesis 2b (H2b): Work areas that resolve a higher percentage of easier-to-solve problems will have higher perceived improvement in performance than work areas that resolve a lower percentage of easier-to-solve problems.

One method for facilitating either the analysis or action-oriented approach is to have managers ensure that problems selected for resolution efforts actually get resolved. Regardless of whether an analysis or action-orientation is used, successful implementation of suggestion-based programs requires senior management commitment to problem resolution (Frankel, et al., 2005, Pronovost, et al., 2004). This recommendation would seem to contradict the process improvement literature, which emphasizes the importance of having frontline employees involved in identifying and resolving issues (Jimmerson et al., 2005). However, research has found that senior managers can be helpful to frontline workers' resolution efforts because they control financial resources needed to address issues that involve capital investment (Carroll et al., 2006). They also possess the authority necessary to solve problems that cross organizational boundaries (Carroll, et al., 2006). Furthermore, resolving problems requires time away from direct production responsibilities (Victor et al., 2000), which can be difficult for frontline employees. Senior managers can provide slack resources that provide the capacity for resolution (Edmondson, 2003). Thus, we hypothesize that assigning senior managers the responsibility of ensuring that a problem gets addressed will achieve better results.

Hypothesis 3 (H3): Work areas that have a higher percentage of problems assigned to a senior manager to ensure resolution will exhibit greater perceived improvement in performance than those with a lower percentage of problems assigned to a senior manager.

3. Methodology

We test our hypotheses in a field study of U.S. hospitals that were randomly selected to implement an MBWA-based improvement program. The program was launched in January 2005 and lasted for 18 months. Hospitals are an appropriate context to study MBWA-based improvement programs because of the widespread adoption of such programs as a tool for improving patient safety (Frankel, et al., 2003). For example, over 79 hospitals in the United Kingdom have implemented the program (National Patient Safety Agency, 2011). Furthermore, the importance of improving patient safety has been highlighted in several high profile reports (Institute of Medicine, 1999, Institute of Medicine, 2001).

3.1. The MBWA-based Program

We drew on prior research to design the MBWA-based program (Frankel, et al., 2008, Pronovost, et al., 2004, Thomas, et al., 2005). The MBWA-based program consisted of repeated cycles of senior manager-staff interaction, debriefing, and follow up. Senior managers, such as the Chief Executive, Operating, Medical, and Nursing Officers (CEO, COO, CMO, and CNO, respectively), interacted with frontline staff to generate, select, and solve improvement ideas. Their interactions took two forms: visits to the organizations' frontlines to observe work, which were called "work system visits;" and special meetings, called "safety forums," with larger groups of staff to discuss safety concerns. The two activities were conducted in the same work area, such as the emergency department. In work system visits, four senior managers would each spend 30 minutes to two hours visiting a particular work area to observe a person doing work. The senior managers would each observe a different process, such as medication administration, or person, such as a nurse or physician, to shed cross-disciplinary insight into the work done in the area. The purpose was to build senior managers' understanding of the frontline work context and gather real-time, grounded information about safety problems (Frankel, et al., 2008). In addition to the visits, managers also facilitated a safety forum in the work area. The safety forums were designed to enable a larger group of frontline workers from the work area to tell senior managers about their safety

concerns and points of pride (Sobo and Sadler, 2002). By supplementing work system visits with safety forums, the program addressed research suggesting that interaction with more frontline staff increased MBWA-based programs' positive impact on culture (Thomas, et al., 2005).

The MBWA-based program continued with a “debrief meeting,” which organized the information collected from the system visits and forum. The senior managers attended, as did the work area managers, selected frontline workers, and the hospital's patient safety officer. They compiled the improvement ideas identified through manager-staff interaction, discussed the ideas, in some cases prioritized them, and decided next steps, ranging from doing nothing to suggesting solutions and assigning responsibility. Managers were encouraged to communicate with staff about implementation efforts, describing what changes were made in response to identified ideas. The patient safety officers entered the ideas generated and actions taken into an electronic spreadsheet and sent it to our research team for analysis.

Each round of activities constituted one cycle. Each cycle focused on a clinical work area of the hospital and took approximately three months to complete, which is comparable to the time reportedly required for improvement teams to solve problems (Evans and Dean, 2003, Pronovost, et al., 2004). After completing a cycle, the management team would move to a different work area. The program focused on the following four work areas: operating room or post anesthesia care unit (OR/PACU), intensive care unit (ICU), emergency department (ED), and medical/surgical ward (Med/Surg). Senior management teams were able to customize the order in which they conducted the program. Cycles continued over the 18-month implementation. On average, hospitals conducted cycles in 4 work areas.

3.2. Recruitment

Our study employed a quasi-experimental design, including a pretest and post-test of work areas. We drew a random sample of 24 US acute-care hospitals, stratified by size and geographic region. No financial incentives were provided; however, participation fulfilled a national accreditation requirement. Data on PIP were collected through surveys before implementation of program activities (2004) and again after the program was completed (2006). At each hospital, we surveyed a random 10% sample of frontline

workers, with additional oversampling in OR/PACUs, EDs, and ICUs in 2006 to improve sample size. The baseline (2004) response rate was 52%; the follow-up (2006) response rate was 39%. For the analyses in this paper, we used data from registered or licensed vocational nurses (n=417 in 2004 and n=433 in 2006). Twenty of the 24 treatment hospitals completed the improvement program in at least two work areas. The four that did not complete the treatment dropped out because one went out of business, one was purchased by another organization, and two experienced significant senior management turnover. As a result, they were unable to complete more than one cycle of activities and did not provide data on ideas generated, selection, actions taken, and feedback provided to frontline workers, or the posttest survey. We thus excluded these hospitals from our analysis. There was no difference in PIP in 2004 between the four types work areas targeted by the intervention in the four hospitals that dropped out of the treatment and in the 20 that did not ($t = -.904$, NS).

3.3. Data and Measures

Using a data collection spreadsheet that we provided, treatment work areas reported 1,245 patient-safety problems that were identified during the visits and forums. Each row of the spreadsheet represented a unique safety concern. The columns included hospital name, work area, who participated in the MBWA activity, the problem, recommended actions for resolving the issue, what actions were taken, and who was responsible for ensuring the problem was resolved. Each hospital provided us with a list of the senior managers, which we used to determine whether a senior manager attended the program activity (e.g. work system visit or safety forum) and whether a senior manager was assigned responsibility for the problem.

Work areas could use three columns to prioritize identified problems. Work areas in eight hospitals filled out these columns. The first column was a rating of the problem's safety risk on a scale from 1 to 10, with 1 = low to no risk, 3 = mild discomfort, 5 = would require intervention, and 10 = could cause harm or death. The second column was frequency of occurrence, with 1 = Very unlikely (hasn't occurred yet to our knowledge); 2 = Possible (has been known to occur); 3 = Very likely (occurs regularly). The third column rated difficulty of solution, with 1 = easy, can be done within 30 days, 2 = moderate,

multiple departments' approval required, 90 days; and 3 = difficult, multiple departments, process changes, and major budget, 6 months.

3.3.1. Independent Variables. Following Frankel et al.'s method (2003), we calculated a priority score for each problem by multiplying the work area's rating of the problem's severity by its frequency. To test H1a, for each work area we created a variable, "mean priority top quartile identified", that was the mean priority score of the top quartile of identified problems. To do this, we first ordered the individual problems in a work area by their priority scores in descending order. If there were ties, problems that were solved were placed ahead of non-solved problems to ensure we gave credit to work areas for solving problems that might lie on the boundary of the top quartile. We also counted the total number of problems in the work area. We used this information to determine each problem's priority percentile score. Next, we created a variable that was 1 if the problem was in the top quartile for priority percentile and missing otherwise. We multiplied the priority score by this top quartile dummy to get a new variable that was the priority score of the top quartile identified problems. Finally, we collapsed this variable to the work area level to calculate the mean priority score of top quartile of identified problems. As an alternate test of H1a, we also created a variable for the highest priority score of the identified problems in the work area.

To test H2a, we followed a similar process to create a variable, "% top quartile resolved," that measured the percentage of the top quartile priority problems that were resolved. To do this, in the dataset of individual problems, we generated a variable that was 1 if a problem was in the top quartile and resolved, 0 if it was in the top quartile and not resolved, and missing if the problem was not in the top quartile. We then collapsed this variable in the individual problem data to the work area level to calculate the percentage of top quartile problems that were resolved. As an alternate test of H2a, we also created a variable, "Was top ranked problem resolved" that indicated whether or not the top ranked problem in the work area was resolved (1 = yes, 0 = no). The alternate specifications for H1a and H2a allowed us to more precisely test the prediction from the innovation literature that performance is driven by identifying and solving the highest magnitude idea (Girotra, et al., 2010).

To test H1b, we created a variable, “% identified that were easy-to-solve” that was the percentage of identified problems in a work area that were rated as easy to solve (a “1” on the difficulty of solution scale). To test H2b, we generated another work area level variable, “% resolved problems that were easy-to-solve” by calculating the percentage of the set of resolved problems in a work area that were rated as easy-to-solve.

To test H3, we calculated the percentage of identified problems in the work area for which a senior manager was assigned responsibility to ensure that the problem was resolved.

3.3.2. Outcome Measure. Our outcome measure was “change in PIP from 2004 to 2006.” The measure was derived from four survey items: “The quality of services I help provide is currently the best it has ever been;” “We are getting fewer complaints about our work;” “Overall, the level of patient safety at this facility is improving;” and “The overall quality of service at this facility is improving.” Using a 5-point Likert response scale ranging from 1=strongly disagree to 5=strongly agree, we asked respondents the extent to which they agreed with the items. Agreement indicated that respondents thought quality and safety performance were improving. The scale exhibited high reliability, with a Cronbach’s alpha of .85 for the combined 2004 and 2006 individual-level, nurses-only data set (Nunnally, 1967).

The use of change scores is an appropriate method for testing change over time (Fitzmaurice, 2001). To create a change score for each work area, we first used the 2004 data and calculated for each nurse the mean for the four items. We then calculated the 2004 mean for each work area by averaging the mean scores of the nurses who worked in that work area. We repeated this process for the 2006 data. Then, we subtracted the each work area’s 2004 mean score from its 2006 mean score.

We calculated intraclass correlations (ICC) and mean interrater agreement score (r_{WG}) to test whether aggregation of PIP was appropriate. Significant intraclass correlations (ICC[1]=.06, $F=5.69$, $p\text{-value} < .000$, and ICC[2] = .82) supported aggregation (Bliese, 2000). The r_{WG} for nurses’ rating of PIP was 0.60, which also was sufficient for aggregation (Zellmer-Bruhn, 2003).

3.3.3. Control Variables. To account for the fact that work areas with a low PIP in 2004 had more opportunity to achieve a large improvement in PIP, we included a dichotomous variable to indicate

whether PIP in 2004 was in the lower quartile (“bottom quartile 2004 PIP”) to control for the starting point, which was coded as a “1” for the bottom quartile of work areas in PIP in 2004 and a “0” for all others (Fitzmaurice, 2001). This method enabled us to test for the change in PIP while also controlling for a low starting point.

For testing H1a-b and H2a-b, our sample size was the 24 work areas that formally prioritized their problems using the coding scheme we had provided. As a result, for these hypotheses, we were limited in our ability to use non-significant control variables. However, our random selection of hospitals helps alleviate concerns that our model may be missing a variable that explains our results (Antonakis, 2010). We did not include control variables for unit type (e.g. ED, ICU, OR/PACU) as none were significant and their inclusion did not change our results. We also tested for hospital-level control variables, such as teaching status, number of hospital beds, non-profit status and geographic region, but none of these were significant and their inclusion did not change our results.

In our test of H3, we did not use any data from work areas’ ratings of problem severity, so our sample size included the full set of 58 intervention work areas. For testing this hypothesis, we were therefore able to include more control variables. To better isolate the impact of a senior manager being assigned responsibility for problem resolution, we controlled for the percentage of problems within a work area that were resolved. We created a variable “% of problems resolved” by first coding a problem as having had solution effort if there was evidence in the data set that action had been taken to address the problem. We also controlled for the fidelity of implementation with the following variables: the number of work system visits that were conducted in the work area, whether a work system visit was conducted by a senior manager (1=yes, 0=no), and whether a safety forum was conducted in the area (1=yes, 0=no).

3.4. Testing of Hypotheses

We used Stata 11.1™ to test our hypotheses. We used linear regression with robust standard errors and clustered by hospital (Rabe-Hesketh and Everitt, 2004) with change in PIP as our outcome variable. The Shapiro-Wilk test showed that the residuals were normally distributed (V close to 1 and $p > .10$)

(Royston, 1992). Multicollinearity was also not an issue as all Variance Inflation Factors were well below the threshold of 10 (Chatterjee and Hadi, 1986).

3.5. Qualitative Data Collection and Analysis.

We visited each intervention hospital to observe prescribed activities, such as a work system visit or a safety forum. In addition, we discussed and observed specific examples of changes implemented in response to problems identified through the program to verify accuracy of the data submitted to us. There were no discrepancies. We also interviewed frontline staff, department managers, and the CEOs. Interview questions addressed the nature of performance improvement in the hospital in general and as it related to implementing the MBWA-based program. Interviews were recorded and transcribed. After each visit, investigators wrote a journal of the day's activities from notes taken during the day. The journal and interviews were combined into a transcript, which provided qualitative data on the hospital's improvement process.

4. Results

Intervention work areas implemented the MBWA-based program, as evidenced by the following statistics. On average, the 58 work areas identified 19 problems and took action on 11. Ninety-one percent had at least one work system visit; 79% had senior manager participation in a visit; and 48% conducted a safety forum. On average, senior managers were assigned responsibility for 10% of the identified problems. The evidence also suggests that the identified problems were legitimate concerns. Using the subset of work areas that prioritized their problems, the mean priority score for all identified problems was seven. As shown in Table 1, on average, the top quartile of identified problems had a mean priority score of 17 and the highest priority score, on average, was 19 (out of a maximum of 30).

We found variation in the program's effectiveness among intervention work areas. The bottom quartile of work areas experienced a decrease in PIP from a minimum drop of 0.375 points to a maximum drop of 2.25, or nearly half of the 5-point scale. Of these 15 work areas, four were already below median

in 2004, suggesting this is not merely a regression to the mean effect. The top quartile increased PIP by a minimum of .38 points to a maximum increase of 1.33. Of these 13 work areas, three had above median PIP in 2004. Given the large variation in results, we examined the analysis-orientation versus action-orientation to explain implementation success. Descriptive statistics and correlations for our main variables are shown in Table 1.

Insert Table 1 about here

4.1. Regression Results

Model 1 in Table 2 shows the result from testing H1a and b, and H2 and b. Identifying a top quartile of problems with higher average priority score was not associated with improved PIP (coefficient = .02, NS), nor was identifying a higher percentage of easy to solve problems (coefficient = -.60, NS). Thus, problem identification does not appear to impact PIP and Hypotheses 1a and 1b are not supported. Model 1, however, does provide evidence for Hypothesis 2b. A higher percentage of problems solved that were rated as “easy-to-solve” was associated with higher % change (coefficient = 1.00, $p < .05$). A one standard deviation (27%) increase in the percentage of solved problems that were easy-to-solve resulted in a 1.0 point increase in change in PIP, which was a 26% improvement ($1/\text{mean change in PIP of } 3.77 = .26$). The percentage of the top quartile priority problems that were solved was not significant (coefficient = -.22, NS). Thus, Hypothesis 2a is not supported. Likewise, when we used the alternative method for specifying an analysis-orientation by including the highest priority score of identified problems (coefficient .02, NS) rather than the mean priority of the top quartile identified and whether the top ranked priority problem was resolved (coefficient -.01, NS) rather than the percentage of the top quartile priority problems, these variable were also not significant (Table 2). Thus, theory from the innovation literature suggesting that performance is driven by identifying and solving the highest magnitude idea was not supported. However, the percentage of problems resolved that were rated as “easy-to-solve” remained

significant (coefficient = .82, $p < .01$), though the percentage identified that were easy-to-solve remained non-significant (coefficient = -.45, NS). Thus, it appears that an action-oriented approach that focuses on addressing problems that have been identified regardless of priority rating is more successful at increasing PIP than an analysis-oriented approach that seeks to identify and solve the high priority problems.

An alternate explanation for our finding that an action-oriented approach was successful (H2b) could be that work areas were more successful because they spent more money on problem solving rather than because they allocated more of the same amount of money on action rather than analysis. To control for this potential “spend more” explanation, the authors individually rated the rough cost of each solved problem on a scale of 1 to 3 with 1=low (cost \leq \$500, low cost of solution), 2=medium (cost $>$ \$500 $<$ \$150,000), and 3=high (cost \geq \$150,000) based on the description of what the work areas had done to solve the problem and independent research to check the cost of products or services whose costs could not be easily estimated. Examples of each class of solution cost are: for “1” apply a coating to one window to improve patient privacy; for “2” purchase new lighting in a catheterization laboratory to illuminate procedures; and for “3” hire several new personnel to address patient transportation needs. We compared scores and discussed our rationale until we reached consensus for all solved problems. We then summed the total estimated solution costs, estimating 1 = \$250; 2=\$5000; 3=\$150,000, for all of the solved problems in each work area.

Another possible explanation of H2b is that variation in quality of solution efforts (e.g. some work areas might have engaged in only superficial steps while others might have systematically resolved underlying causes) impacted the results. We also controlled for this “higher quality” explanation by hiring ten nurses not affiliated with the treatment hospitals to rate the solution effectiveness of each solved problem using a scale ranging from 1 to 10, with 1 being “no information given” and 10 being “systemic fix” that would prevent recurrence (scale is available from authors). Agreement on their ratings was fair ($\kappa = 0.23$) (Landis and Koch, 1977). The average rating for solution effectiveness was 5.9 for solved problems (“solution action in progress” on our scale) and 2.7 (“no solution implemented”) for unsolved problems, which validates their coding.

Given our small sample size, we omitted the analysis-oriented related variables in our testing of the cumulative cost and solution effectiveness variables. As Model 3 shows, the variable for the cumulative “cost of solving problems” was not significant (coefficient 0.00, NS). This may be because work areas could improve PIP without having to spend a lot of money on solutions. Solution effectiveness was also not significant (coefficient = -.11, NS). Percentage of solved problems that were easy-to-solve remained significant (coefficient = 1.22, $p < .05$), indicating that the results are similar after accounting for spending and solution effectiveness.

The accumulation of evidence in these three models supports H2b, which had predicted that an action-oriented approach was associated with higher PIP than an analysis-orientation. In our sample of work areas, it was more beneficial to solve a higher percentage of easy problems than high priority problems.

Insert Table 2 about here

Table 3 shows the results from testing Hypothesis 3, which proposed that senior managers taking responsibility for ensuring that identified problems get resolved would be associated with higher PIP. We controlled for the fidelity of implementation to ensure that poor results were not driven by a failure to implement the program (Nembhard, et al., 2009). Hypothesis 3 was supported (coefficient = .79, $p < .05$). Increasing the percent of problems assigned to senior managers by one standard deviation (23%) was associated with a 0.79 increase in PIP. This equates to a 21% increase in PIP ($.79/3.76 = .21$). The variables testing fidelity to implementation were not significant in explaining a positive change in PIP.

Insert Table 3 about here

4.2. Robustness Checks

Other scholars have used a different approach for testing improvement over time, which is using the post measure as the outcome variable and the pre measure as a control variable (Fitzmaurice, 2001). To make sure our results held using this approach, we also tested our hypotheses using 2006 PIP as the dependent variable (instead of the difference score) and 2004 PIP as a control variable (instead of the bottom quartile 2004 PIP). By this method, the results were the same for all hypotheses (available from authors).

4.3. Qualitative Results

To provide insight into the nature of implementation of MBWA-based programs, Table 4 presents qualitative data from the five work areas that improved the most on change in PIP and the five work areas that decreased the most. On average, the difference in 2006 and 2004 PIP was .85 for the top five work areas and -1.4 for the bottom five.

Examining the types of issues that were identified and the responses to these issues shows that the top work areas identified meaningful problems. Managers in turn took these problems seriously and addressed as many as possible. For example, Hospital 88's Med/Surg unit was one of the most improved work areas. One of the identified issues was that the small medication room prevented two nurses from preparing medications simultaneously, which was an inconvenience and delayed patient care. Senior managers discussed the issue with staff and they collectively made a plan to move the medication room to a larger space. The COO commented, "It's a little thing, but when you actually see them doing the process, you say, 'What a minute, that is difficult for them.'" An interview with a staff nurse highlighted management's general willingness to address issues. She commented, "These people address safety issues. It may not always get addressed the way you want it to, but it still gets addressed."

Conversely, in work areas that deteriorated on PIP, an emphasis on prioritizing problems limited solution efforts. For example, Hospital 129's ED identified important issues such as long lead times to receive lab results, which slowed patient flow through the work area and contributed to long wait times in the ED. However, in the safety forum we observed the manager spent the entire time getting staff input on

prioritizing the items (e.g. severity, frequency of occurrence, and ease of solution), leaving no time to discuss how the issues might be resolved. This exemplifies the potential pitfalls of an analysis-oriented approach and suggests why an action-orientation may be preferable. This work area did not solve any problems, despite investing considerable time in identifying and prioritizing them. In other work areas that had the lowest % change scores, such as Hospital 34's OR/PACU and Hospital 65's ED, identified issues had to be validated by an external group, such as the hospital's patient safety committee, before resolution efforts would be authorized. This extra analysis step created an additional hurdle for problems to be selected for improvement, substantially slowing the pace of change.

Insert Table 4 about here

5. Discussion and Implications

We examined the tradeoff between service organizations' orientation toward analysis versus action. Solving a higher percentage of the highest priority problems was not associated with increased PIP, while solving easy-to-solve problems was, lending support to the action-oriented approach. This signals the value of going after the "low-hanging fruit" rather than concentrating on the "big hits." In further support of an action-oriented approach, having senior managers assume responsibility for ensuring that problems get resolved was associated with increased PIP. One explanation for this finding is that organizational change often requires senior managers to provide financial resources to pay for required equipment, materials, or labor; and organizational support to get an upstream department in the organization to change how they do their work as benefits might accrue downstream. In other words, senior managers can help ensure that action happens. Given the improvement literature's emphasis on empowering frontline employees to solve problems (Powell, 1995), our finding may be interpreted as highlighting the importance of empowering frontline employees to identify and solve problems while supporting those efforts by ensuring that organizational obstacles to improvement are removed. Increasing senior

managers' involvement with performance improvement programs can result in positive change if they facilitate action on identified issues.

Our research finds little evidence that spending resources on an analysis-oriented approach is productive for improving PIP in hospitals.

5.1. Implications for Theory on Approaches to Problem Solving

Prior research has found an association between manager commitment and implementation success for performance improvement programs that rely on frontline employee participation, such as 6-sigma (Coronado and Antony, 2002), Scanlon plans (Miles, 1965, White, 1979), lean (Worley and Doolen, 2006), and TQM (Antony et al., 2002). However, these studies were typically case studies of a few firms' successful implementations or a survey used to correlate implementation success measures with managers' self-reported commitment. Although these studies suggest a link between management commitment and implementation success, they provide limited insight into how to foster commitment and demonstrate it to frontline employees or why it is associated with success.

In our study, we manipulated manager commitment by requiring senior managers to become involved in improvement efforts by visiting their frontlines to identify and resolve problems using an MBWA-based program. We found that managerial involvement was productive for some, but not all work areas. A possible explanation for our mixed results comes from Miles (1965). He postulated that managers held one of two beliefs about the value of interacting with employees. One belief was that interactions of senior managers with frontline staff were valuable because they increased frontline staff members' morale, though the actual ideas generated were unhelpful. Believing the "symbolic-value" of the interactions was not associated with improvement. The second belief, which was associated with better performance, was that these interactions were valuable because the ideas raised by frontline employees were useful. This belief in the value of interacting with frontline staff underlies a core principle of the Toyota Production System of respect for people (Liker, 2004). Miles' study suggests that managers' respect for frontline employees' concerns may be an important moderator variable for manager

involvement in MBWA programs. An implication is that rather than designing interventions to increase manager involvement, it may be critical to first design interventions that surface and modify managers' beliefs such that they value and therefore leverage the actual ideas raised by frontline staff.

Another possible explanation of the positive impact of an action-orientation in our study is that the landscape of problem priority was relatively flat. In a flat landscape, local search is beneficial and expanded search efforts are unhelpful because the difference between a local high point and the global high point is too small to justify the increased search costs (Sommer and Loch, 2004). Spending resources to discover and solve only high impact problems through programs like MBWA or incident reports, may not yield sufficient improvement due to the lack of a disproportionately high priority problem that if solved would yield dramatic improvement. Instead, substantial improvement might arise from solving the “lower tail” of problems that exist in a flat landscape (Brynjolfsson, et al., 2011). Our findings call into question the assumption that there is *always* a small set of individual problems that have a disproportionately high priority score, as is assumed by performance improvement experts and forms the basis of their recommendation to limit solution efforts to the handful of highest priority problems (Juran et al., 1999). An implication of our study is that little is gained from hospitals' current focus on increasing reporting of safety concerns (Evans et al., 2007) because the expanded search is unlikely to yield significantly higher priority problems. Instead, hospitals would be better served by deploying their limited resources to solve existing problems (Wachter, 2009). Our findings may be applicable to other organizations that have flat problem landscapes. We suspect that service industries that offer complex, customizable services co-created with customers, such as software, consulting, and legal industries, may have flat landscapes, but this remains a question for future research. Our findings might also apply to other improvement programs that rely on employee suggestions, which could result in flat problem landscapes, such as incident reporting systems and lean.

5.2. Implications for Practice

Many organizations' strategies for improving quality begin by trying to increase employees' reports of near misses and errors (Evans, et al., 2007). The implied assumptions are that (1) increasing the number of reports enables organizations to conduct trend analysis that illuminates the most important problems which can then be solved; and (2) resources can be effectively matched with identified issues because many issues will be of sufficiently low priority that they can be ignored at low or no cost to the organization. In contrast, our study found that there may be little benefit to this approach. Our study instead provides support for the lean community's recommendation that organizations focus on developing capacity to address identified problems and taking action on problems (Imai, 1986). Rather than seeking to increase reporting, organizations should instead seek to increase their problem-solving capacity, which will enable them to take action on an increased number of reports. This recommendation is supported by prior research that found that underlying capabilities, such as problem-solving capacity, explained differences in organizational improvement (Adler et al., 2003).

Related to the importance of first creating problem-solving capacity, our results suggest how MBWA-based programs differ from the performance improvement technique of Kaizen. Kaizen is a structured activity where workers are relieved of normal production responsibilities and use a standardized approach to identify and fix problems under guidance of a manager (Imai, 1986, Laraia et al., 1999). This program is similar to MBWA; however, Kaizen events occur after managers and frontline staff are trained on standardized problem solving techniques so they have the ability to address the issues that arise (Imai, 1986, Laraia, et al., 1999). Kaizen also focuses on taking action to solve as many problems as possible in a given time frame.

Organizations can increase problem-solving capacity through three routes. First, organizations can hire additional staff who are already skilled at performance improvement techniques (Bohn, 2000). Second, existing employees can be trained on these techniques. This has the added benefit of developing a common approach and language for improvement (Repenning and Sterman, 2002). Third, a focus on resolving problems—even if they are small—will create additional capacity by spreading the techniques to those who interface with these efforts (Repenning and Sterman, 2002). To be successful, performance

improvement needs to be part of the regular work of organizations (Victor, et al., 2000). Solving problems increases problem-solving capacity by spreading skills and knowledge through the organization (Adler, et al., 2003). Rather than viewing problem-solving capacity as a resource that is depleted with use, a better analogy might be a muscle that becomes stronger with use. It is important to educate managers and employees to expect the “worse before better” dynamic that occurs when resources get diverted from production to improvement (Repenning and Sterman, 2002). However, over time performance should improve and capacity increase as employees become more skilled at performance improvement and contribute more meaningful ideas (Arthur and Aiman-Smith, 2001).

5.3. Limitations

Our findings must be considered in light of study limitations. First, our small sample size limited our analysis. We had a small sample size for several reasons. Given the cost and time intensive nature of conducting an experiment with hospitals over an 18-month period, it is difficult to conduct field-based, interventional experiments with samples larger than 20 to 30 organizations. Furthermore, some work areas were dropped from our analysis due to missing data in either 2004 or 2006, in part because concerns about survey overload caused hospital administrators to survey a random sample of only 10% of frontline workers. In addition, despite our provision of a method of prioritizing identified problems, many organizations choose not to assign prioritization values and therefore work-area coded data on problem priority was not available for all work areas in our study. Future research with larger sample sizes could test more nuanced theory. For example, an action-oriented approach may be most successful for work areas that start from a weak position and therefore can benefit the most from action, whereas an analysis-oriented approach may be most helpful for high performing work areas that can be more selective about finding the highest leverage opportunities.

A second limitation is the perceptual measure of improvement. However, hospitals were unwilling to share actual safety incident measures with us. In addition, publicly available measures, such as mortality and readmissions, are very blunt measures and, when we explored these relationships, they were not

significant in explaining variation in this study. Third, hospitals did not track how much they spent on solutions efforts and therefore estimation was the only way of testing the alternate explanation that spending more money on process improvement yielded better outcomes. Future research could contribute to improvement theory by examining the cost of improvement efforts compared to the benefits. A related limitation is the lack of data on quality of solution efforts.

A fourth limitation is that we did not randomize action-oriented approach versus an analysis-oriented approach among the work areas. Instead those differences emerged naturally. A randomized assignment of the approaches would provide a stronger test of the hypotheses.

5.4 Conclusions

Understanding the impact of suggestion-based improvement programs is helpful for organizations that may be considering implementing them. In our study of an MBWA-based improvement program, organizations whose managers ensured that problems were addressed achieved better results. This suggests that improvement programs are more likely to change employees' perceptions when they result in action being taken to resolve problems than when they are a symbolic show of manager interest. Based on study findings, we recommend that organizations focus on increasing their capacity to act on improvement suggestions rather than generating suggestions and prioritizing them.

References

- Adler, P.S., P. Riley, S.-W. Kwon, J. Signer, B. Lee, R. Satrasala. 2003. Performance improvement capability: Keys to accelerating performance improvement in hospitals. *Calif. Management Rev.* **45**(2) 12-33.
- Anderson, C. 2004. The long tail: Why the future of business is selling less of more. *Wired* (October),
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On making causal claims: A review and recommendations. *Leadership Quarterly* **21**(6) 1086–1120.
- Antony, J., K. Leung, G. Knowles, S. Gosh. 2002. Critical success factors of TQM implementaton in Hong Kong industries. *Internat. J. of Quality and Reliability Management* **19**(5) 551-566.
- Arthur, J.B., L. Aiman-Smith. 2001. Gainsharing and organizational learning: An analysis of employee suggestions over time. *Acad. Management J.* **44**(4) 737-754.
- Bagian, J.P., C. Lee, J. Gosbee, J. DeRosier, E. Stalhandske, N. Eldridge, R. Williams, M. Burkhardt. 2001. Developing and deploying a patient safety program in a large health care delivery system: You can't fix what you don't know about. *Joint Commission J. Quality and Patient Safety* **27**(10) 522-532.
- Bagian, J.P., C.Z. Lee, J.F. Cole. 1999. *A method for prioritizing safety related actions: Or now that I have succeeded in increasing the reporting of patient safety related issues, how do I deal with the avalanche of reports?* National Patient Safety Foundation, Chicago.
- Beer, M. 2003. Why total quality management programs do not persist: The role of management quality and implications for leading a TQM transformation. *Decision Sci.* **34**(4) 623-642.
- Benning, A., M. Dixon-Woods, U. Nwulu, M. Ghaleb, J. Dawson, N. Barber, B.D. Franklin, A. Girling, K. Hemming, M. Carmalt, G. Rudge, T. Naicker, A. Kotecha, M.C. Derrington, R. Lilford. 2011. Multiple component patient safety intervention in English hospitals: Controlled evaluation of second phase. *BMJ* **342** d199.
- Benning, A., M. Ghaleb, A. Suokas, M. Dixon-Woods, J. Dawson, N. Barber, B.D. Franklin, A. Girling, K. Hemming, M. Carmalt, G. Rudge, T. Naicker, U. Nwulu, S. Choudhury, R. Lilford. 2011. Large scale organisational intervention to improve patient safety in four UK hospitals: Mixed method evaluation. *BMJ* **342** d195.
- Bliese, P.D. 2000. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. K.J. Klein, S.W.J. Kozlowski, ed. *Multilevel theory, research, and methods in organizations*. Jossey-Bass, San Francisco.
- Bohn, R. 2000. Stop fighting fires. *Harvard Bus. Rev.* **78**(4) 82-91.
- Brynjolfsson, E., Y.J. Hu, D. Simester. 2011. Goodbye pareto principle, hello long tail: The effect of internet commerce on the concentration of product sales. *Management Sci.* **57**(8) 1373-1386.
- Carroll, J.S., S. Hatakenaka, J.W. Rudolph. 2006. Naturalistic decision making and organizational learning in nuclear power plants: Negotiating meaning between managers and problem investigation teams. *Organization Studies* **27**(7) 1037.
- Chatterjee, S., A.S. Hadi. 1986. Influential observations, high leverage points, and outliers in linear regression. *Statistical Science* **1** 379-416.
- Choi, T.Y., O.C. Behling. 1997. Top managers and TQM success: One more look after all these years. *Acad. Management Executive* **11**(1) 37-47.
- Clarke, S. 1999. Perceptions of organizational safety: Implications for the development of safety culture. *J. Organ. Behavior* **20**(2) 185-198.
- Cook, R., D. Woods. 1994. Operating at the sharp end: The complexity of human error. M.S. Bogner, ed. *Human Error in Medicine*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 255-310.
- Coronado, R.B., J. Antony. 2002. Critical success factors for the successful implementation of six sigma projects in organizations. *The TQM Magazine* **14**(2) 92-99.
- Dixon-Woods, M., C.L. Bosk, E.L. Aveling, C.A. Goeschel, P.J. Prono. 2011. Explaining Michigan: Developing an ex post theory of a quality improvement program. *Milbank Quarterly* **89**(2) 167-205.

- Edmondson, A.C. 2003. Speaking up in the operating room: How team leaders promote learning in interdisciplinary action teams. *J. Management Studies* **40**(6) 1419-1452.
- Evans, J.R., J.W. Dean. 2003. *Total Quality: Management, Organization, and Strategy*. Thompson.
- Evans, S., B. Smith, A. Esterman, W. Runciman, G. Maddern, K. Stead, P. Selim, J. O'Shaughnessy, S. Muecke, S. Jones. 2007. Evaluation of an intervention aimed at improving voluntary incident reporting in hospitals *Quality and Safety in Health Care* **16**(3) 169-175.
- Farley, D.O., A. Haviland, S. Champagne, A.K. Jain, J.B. Battles, W.B. Munier, J.M. Loeb. 2008. Adverse-event reporting practices by US hospitals: Results of a national survey. *Quality and Safety in Health Care* **17**(6) 416-423.
- Fitzmaurice, G. 2001. A conundrum in the analysis of change. *Nutrition* **17**(4) 360-361.
- Frankel, A., T.K. Gandhi, D.W. Bates. 2003. Improving patient safety across a large integrated health care delivery system. *Internat. J. Quality in Health Care* **15**(1) i31-i40.
- Frankel, A., E. Graydon-Baker, C. Neppel, T. Simmonds, M. Gustafson, T. Gandhi. 2003. Patient safety leadership walkrounds. *Joint Commission J. on Quality and Safety* **29**(1) 16-26.
- Frankel, A., S.P. Grillo, E.G. Baker, C.N. Huber, S. Abookire, M. Grenham, P. Console, M. O'Quinn, G. Thibault, T.K. Gandhi. 2005. Patient safety leadership walkrounds at Partners Health Care: Learning from implementation. *Joint Commission J. Quality and Patient Safety* **31**(8) 423-437.
- Frankel, A., S. Pratt Grillo, M. Pittman, E.J. Thomas, L. Horowitz, M. Page, B. Sexton. 2008. Revealing and resolving patient safety defects: The impact of leadership walkrounds on frontline caregiver assessments of patient safety. *Health Services Res.* **43**(6) 2050-2066.
- Gandhi, T.K., E. Graydon-Baker, C.N. Huber, A.D. Whittlemore, M. Gustafson. 2005. Closing the loop: Follow-up and feedback in a patient safety program. *Joint Commission J. Quality and Patient Safety* **31**(11) 614-621.
- Girotra, K., C. Terwiesch, K.T. Ulrich. 2010. Idea generation and the quality of the best idea. *Management Sci.* **56**(4) 591-605.
- Hansen, L., M. Williams, S.J. Singer. 2010. Perceptions of hospital safety climate and incidence of readmission. *Health Services Res.* **42**(2) 596-616.
- Harlos, K.P. 2001. When organizational voice systems fail: More on the deaf-ear syndrome and frustration effects. *Journal of Applied Behavior Science* **37**(3) 324-342.
- Hofmann, D.A., B. Mark. 2006. An investigation of the relationship between safety climate and medication errors as well as other nurse and patient outcomes. *Personnel Psychology* **59**(4) 847-869.
- Huang, D.T., G. Clermont, L. Kong, L.A. Weissfeld, J.B. Sexton, K.M. Rowan, D.C. Angus. 2010. Intensive care unit safety culture and outcomes: a US multicenter study. *Internat. J. Quality in Health Care* **22**(3) 151-161.
- Imai, M. 1986. *Kaizen: The key to Japan's competitive success*. Random House Business Division, New York.
- Institute of Medicine. 1999. *To Err Is Human: Building a Safer Health System*. Committee on Quality of Health Care in America, Washington, DC National Academy Press.
- Institute of Medicine. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academy Press, Washington, DC.
- Jimmerson, C., D. Weber, D.K. Sobek. 2005. Reducing waste and errors: piloting lean principles at Intermountain Healthcare. *Joint Commission J. Quality and Patient Safety* **31**(5) 249-257.
- Johnson, C.W. 2003. How will we get the data and what will we do with it then? Issues in the reporting of adverse healthcare events. *Quality and Safety in Health Care* **12**(Suppl II) ii64-ii67.
- Juran, J.M., A.B. Godfrey, R.E. Hoogstoel, E.G. Schilling. 1999. *Juran's Quality Handbook*. McGraw Hill, New York.
- Juran, J.M., F.M. Gyra. 1980. *Quality Planning and Analysis*. McGraw-Hill, New York.
- Landis, J.R., G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* **33**(1) 159-174.
- Laraia, A.C., P.E. Moody, R.W. Hall. 1999. *The Kaizen Blitz: Accelerating Breakthroughs in Productivity and Performance*. John Wiley & Sons, Inc., New York.

- Leape, L.L. 2002. Reporting of adverse events. *New England J. Medicine* **347**(20) 1633-1638.
- Levinson, D.R. 2010. *Adverse events in hospitals: National incidence among medicare beneficiaries*. Department of Health and Human Services Office of Inspector General, Washington, DC.
- Liker, J. 2004. *The Toyota Way*. McGraw-Hill, New York, NY.
- Loch, C.H., L. Van der Heyden, L.N. Van Wassenhove, A. Huchzermeier, C. Escalle. 2003. *Industrial Excellence: Management Quality in Manufacturing*. Springer, Berlin, Germany.
- Luria, G., I. Morag. 2012. Safety management by walking around (SMBWA): A safety intervention program based on both peer and manager participation. *Accident Analysis and Prevention* **45**(0) 248-257.
- MacDuffie, J.P. 1997. The road to root cause: Shop-floor problem-solving at three auto assembly plants. *Management Sci.* **43**(4) 479-502.
- March, J.G. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* **2**(1) 71-87.
- Milch, C.E., D.N. Salem, S.G. Pauker, T.G. Lundquist, S. Kumar, J. Chen. 2006. Voluntary electronic reporting of medical errors and adverse events. *J. General Internal Medicine* **21**(2) 165-170.
- Miles, R.E. 1965. Human relations or human resources. *Harvard Bus. Rev.* **43**(4) 148-157.
- Morrison, E.W., C.C. Phelps. 1999. Taking charge at work: Extrarole efforts to initiate workplace change. *Acad. Management J.* **42**(4) 403-419.
- National Patient Safety Agency. 2011. *Patient Safety First: 2008-2010, The campaign review*. National Patient Safety Agency, London, UK.
- Nembhard, I.M., J.A. Alexander, T. Hoff, R. Ramanujam. 2009. Understanding implementation failure in health care delivery: A role for organizational research and theory. *Academy of Management Perspectives* **23**(1) 1-27.
- Nuckels, T., D. Bell, H. Liu, S. Paddock, L. Hilborne. 2007. Rates and types of events reported to established incident reporting systems in two U.S. hospitals. *Quality and Safety in Health Care* **16**(3) 164-168.
- Nunnally, J.C. 1967. *Psychometric Theory*. McGraw-Hill, New York.
- Parker, S.K., T.D. Wall, P.R. Jackson. 1997. "That's not my job": Developing flexible employee work orientations. *The Academy of Management Journal* **40**(4) 899-929.
- Perrow, C. 1984. *Normal accidents: Living with high-risk technologies*. Princeton University Press, Princeton, NJ.
- Peters, T.J., R.H. Waterman, Jr. 1982. *In Search of Excellence: Lessons from America's Best-Run Companies*. Harper & Row, Publishers, Inc., New York, NY.
- Powell, T.C. 1995. Total quality management as competitive advantage: A review and empirical study. *Strategic Management J.* **16**(1) 15-37.
- Pronovost, P.J., B. Weast, K. Bishop, L. Paine, R. Griffith, B.L. Rosenstein, R.P. Kidwell, K.B. Haller, R. Davis. 2004. Senior executive: Adopt-a-work unit: A model for safety improvement. *Joint Commission J. Quality Improvement* **30**(2) 59-68.
- Rabe-Hesketh, S., B. Everitt. 2004. *A Handbook of Statistical Analysis using Stata*. Chapman & Hall/CRC Press, Boca Raton, FL.
- Ramanujam, R., D. Keyser, C. Sirio, D. Thompson. 2008. Inside the black box: The use of voluntary reported incident data. Working paper, Vanderbilt University, Nashville, TN.
- Reason, J. 1990. *Human Error*. Cambridge University Press, New York, NY.
- Reason, J. 2000. Human error: Models and management. *British Medical J.* **320** 768-770.
- Repenning, N.P., J.D. Sterman. 2002. Capability traps and self-confirming attribution errors in the dynamics of process improvement. *Administrative Sci. Quart.* **47**(2) 265-295.
- Rogers, Pronovost, Isaac, Schoenfeld, Leape, Iezzoni, Blumenthal. 2010. Systematically seeking clinicians' insights to identify new safety measures for intensive care units and general surgery services. *Amer. J. Medical Quality* **25**(5) 359 - 364.
- Royston, J.P. 1992. Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing* **2** 117-119.

- Sagor, R., B.G. Barnett. 1994. *The TQE Principal: A Transformed Leader*. Corwin Press, Thousand Oaks, CA.
- Singer, S.J., S. Lin, A. Falwell, D. Gaba, L. Baker. 2009. Relationship of safety climate and safety performance in hospitals. *Health Services Res.* **44**(2 (Part 1)) 399-421.
- Singer, S.J., P. Rivard, J. Hayes, P. Shokeen, D. Gaba, A. Rosen. forthcoming. Improving patient care through leadership engagement with frontline staff: A VA case study. *Joint Commission J. of Quality and Patient Safety*.
- Sobo, E.J., B.L. Sadler. 2002. Improving organizational communication and cohesion in a health care setting through employee-leadership exchange. *Human Organization* **61**(3) 277-287.
- Sommer, S.C., C.H. Loch. 2004. Selectionism and learning in projects with complexity and unforeseeable uncertainty. *Management Sci.* **50**(10) 1334-1347.
- Terwiesch, C., K.R. Ulrich. 2009. *Innovation Tournaments: Creating and Selecting Exceptional Opportunities*. Harvard Business School Press, Boston, MA.
- Thomas, E.J., J.B. Sexton, T.B. Neilands, A. Frankel, R.L. Helmreich. 2005. The effect of executive walk rounds on nurse safety climate attitudes: A randomized trial of clinical units. *BMC Health Services Research* **5**(1) 28-36.
- Thompson, D.N., G.A. Wolf, S.J. Spear. 2003. Driving improvement in patient care: Lessons from Toyota. *Journal of Nursing Administration* **33**(11) 585-595.
- Tucker, A.L. 2007. An empirical study of system improvement by frontline employees in hospital units. *Manufacturing Service Oper. Management* **9**(4) 492-505.
- Victor, B., A. Boynton, T. Stephens-Jahng. 2000. The effective design of work under total quality management. *Organ. Sci.* **11**(1) 102-117.
- Wachter, R.M. 2009. Hospital incident reporting systems: Time to slay the beast http://community.the-hospitalist.org/blogs/watchers_world/archive/2009/09/20/hospital-incident-reporting-systems-time-to-slay-the-monster.aspx. Retrieved 04/08/11, 2011. Wachter's World.
- White, J.K. 1979. The Scanlon plan: Causes and correlates of success. *Acad. Management J.* **22**(2) 292-312.
- Worley, J.M., T.L. Doolen. 2006. The role of communication and management support in a lean manufacturing implementation. *Management Decision* **44**(2) 228-245.
- Zellmer-Bruhn, M.E. 2003. Interruptive events and team knowledge acquisition. *Management Sci.* **49**(4) 514-528.
- Zohar, D., Y. Livne, O. Tenne-Gazit, H. Admi, Y. Donchin. 2007. Healthcare climate: A framework for measuring and improving patient safety. *Critical Care Medicine* **35**(5) 1312-1317.

Table 1. Mean, Standard Deviation (SD), and Correlations for Work Areas that had the MBWA intervention and Prioritized Problems (n=24)

	Variable	Mean	SD	Min	Max	1	2	3	4	5	6
1	Change in PIP	.02	.53	-1.17	1.1						
2	Avg priority of top quartile identified	17.23	6.67	6	30	.298	1				
3	Highest priority score	18.75	7.43	6	30	.325	.952*	1			
4	% of identified problems that were easy to solve	36%	26%	0%	100%	-.016	-.305	-.289	1		
5	% of top quartile problems solved	88%	29%	0%	100%	.186	.091	.109	-.045	1	
6	Biggest priority problem was solved	88%	34%	0	1	.209	.11	.039	-.086	.799*	
7	% solved that were easy	33%	27%	0%	83%	.327	.097	.099	.551*	.432*	.35

* p<0.05

Table 2. Regression comparing change in PIP in treatment work areas (OR/PACU, ICU, ED, and Medical/Surgical) that rated the severity, frequency, and ease of solution of the problems, clustered by hospital with robust standard errors in parenthesis (n=24 work areas, 8 hospitals)

Outcome variable	Model 1	Model 2	Model 3
	Change in PIP	Change in PIP	Change in PIP
Bottom quartile 2004 PIP (1=yes)	.39* (.16)	.36^ (.19)	.38* (.13)
H1a. Mean priority top quartile identified	.02 (.02)	Not in model	Not in model
H1a. Highest priority score of identified problems	Not in model	.02 (.02)	Not in model
H1b. % identified that were easy-to-solve	-.60 (.49)	-.45 (.45)	-.90^ (.42)
H2a. % top quartile resolved	-.22 (.23)	Not in model	Not in model
H2a. Was top-ranked problem resolved (1=yes)	Not in model	-.01 (.26)	Not in model
H2b. % resolved problems that were “easy to solve”	1.00* (.30)	.82** (.21)	1.22* (.46)
Cum cost of solving problems	Not in model	Not in model	.00 (.00)
Ave effectiveness of solution effort	Not in model	Not in model	-.11 (.10)
Constant	-.25 (.48)	-.47 (.46)	.61 (.62)
Observations	24	24	24
Degrees of freedom	F(5,7)	F(5,7)	F(5, 7)
F statistic	10.99**	5.28*	7.08*
Adjusted R-squared	.06	.07	.08

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.10

Table 3. Impact of the Percentage of Problems Assigned to Senior Managers on Change in PIP in treatment work areas (OR/PACU, ICU, ED, and Medical/Surgical) (n=58)

Outcome variable	Model 1 Change in PIP
Bottom Quartile 2004 PIP (1=yes)	.56** (.15)
Percentage problems solved	.12 (.33)
Number of work system visits in the area	-.04^ (.02)
Senior manager participated in work system visit (1=yes)	-.12 (.23)
Safety forum in the area (1=yes)	-.12 (.14)
H3. % of problems assigned to senior managers for resolving	.79* (.32)
Constant	-.08 (.31)
Observations	58
Degrees of freedom	F(6, 19)
F statistic	2.96*
Adjusted R ²	.10

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.10

Table 4. Illustrative Problems, Solutions, and Quotes from Top and Bottom Five Work Areas

Hosp ID	Work Area	2004 Score	2006 Score	Change in PIP	Examples	Solution Efforts	Illustrative CEO Quote about Prioritization
116	OR/ PACU	3.3	4.6	1.3	Need more clinic space	Made new clinic rooms	The associates will prioritize with the managers, who have a good idea of what the staff want to do.
88	Med/ Surg	3.6	4.7	1.1	Medication room is very small for 2 people	After discussing with staff, changed medication preparation to a larger room.	These people address safety issues. It may not always get addressed the way you want it to, but it still gets addressed.
39	ED	4.0	5.0	1.0	Feel like "dumping ground" when the clinic closes; Roof leaks, Need more blood pressure machines	Relocated clinic in to expand ED patients; hired additional ED staff; fixed roof; provided blood pressure equipment	Nurse almost gave wrong medication because two similar drugs next to each other in Pyxis. Told CNO. Pharmacist came up right away and changed drawer.
100	Med/ Surg	2.6	3.6	1.0	Newly diagnosed diabetic patients can't get glucometers from insurance; buy different kinds, hard for nurses to teach.	Vendor donated glucometers, in-serviced nurses, made kits for newly diagnosed diabetic patients	Manager ordered new isolation carts to keep supplies for each patient outside the door to prevent spread of MRSA.
47	ED	3.0	3.8	0.8	Need prompt response from pharmacy for selected meds; need lift equipment for obese patients; Pyxis [#] IT display disposed to medication errors	Installed phone system with priority access to pharmacy; identified or added lift equipment; reprogrammed Pyxis IT display	We understand what needs to be done - trying to get rid of verbal orders, trying to set up our Pyxis machine differently.
122	OR/PA CU	5.0	3.9	-1.1	Keeping surgical equipment in order and performing, physician pressure to rush through certain protocols	They have equipment checks and big push on sterilization, still working on this and discussing it in meetings.	"A lot of people would say, 'I don't even know he is involved.' But all of my administrative people are involved."
119	OR/ PACU	3.8	2.6	-1.2	Need exhaust air, some equipment (chairs), back up of patients in ED, beds not ready	Changes made to improve air, equipment ordered and others repaired, working on flow in ED	It is hard to find the time and energy [to sustain this program] because there are other demands that pour in.
34	OR/ PACU	5.0	3.8	-1.3	OR table not safe for bariatric patients; insufficient checking of patient labs prior to surgery	No solutions listed	Anyone can submit safety idea to their vice-president. It gets sent out for review to applicable departments.
129	ED	4.4	3.0	-1.4	Long lead times for radiology and lab, ties up rooms, long waits in ED, units not taking patients	No solutions listed.	Spent 30 minutes of safety forum deciding on priority scores with no discussion about what would correct the problems.
65	ED	4.3	2.0	-2.3	Police bringing in dangerous patients with only two people on at night	Talk to police department about patients, have security cameras and panic buttons	You can't fix them all, but you have to prioritize. Our patient safety committee will end up doing that.

Pyxis™ is an automated medication dispensing device used by nurses to administer medications to their patient