Co-Evolution of Process and Content in Organizational Change: Explaining the Dynamics of Start and Fizzle

by

J. Bradley Morrison

Bachelor of Science, Chemistry (1978)
Bachelor of Science, Management Science (1979)
Massachusetts Institute of Technology

Master of Business Administration, Finance (1985)
University of Chicago Graduate School of Business

Submitted to the Alfred P. Sloan School of Management in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Management

at the

Massachusetts Institute of Technology

February 2003

© 2003 Massachusetts Institute of Technology
All Rights Reserved

Signature of Author: J. Bradley Morrison
Sloan School of Management
January 2003

Certified by: John D. Sterman
Professor of Management Science
Thesis Supervisor

Accepted by: Birger Wernerfelt
Chairman, Ph.D. Committee
Sloan School of Management
Co-Evolution of Process and Content in Organizational Change:  
Explaining the Dynamics of Start and Fizzle

by

J. Bradley Morrison

Submitted to the Alfred P. Sloan School of Management  
in January 2003 in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Management

ABSTRACT

Essay One: Explaining the Start and Fizzle of Organizational Change: Co-evolving  
Process and Content

The theme that sustaining efforts to change is difficult appears in many guises throughout  
the organizations. The tendency for change processes to run out of energy and  
momentum is widely recognized. In this essay, I offer an inductive study of the  
dynamics of organizational change that describes situated human interactions to explain  
the observed trajectory of organizational behavior. As a window through which to view  
change, I study a manufacturer adopting production practices based on concepts of lean  
manufacturing and the so-called Toyota Production System (TPS). I develop a grounded  
theory based on extensive fieldwork observing a change effort that began with temporary  
improvement but subsequently entered a phase of decline, the characteristic start and  
fizzle of change. I examine how the actions of various groups interact with each other  
and with characteristics of the workplace. I then induce a model that characterizes the  
dynamics of participatory change. The model draws on a close examination of how  
people do the work of process improvement to identify a set of mechanisms through  
which process and content interact to constitute the patterns of organizing that result. I  
find that front-line participation led workers to generate ideas, but the work of  
implementing those ideas placed demands on key support personnel such as  
manufacturing engineers. The support personnel quickly become overwhelmed, facing a  
growing backlog of ideas to implement. As the support personnel modified their work  
practices to address the mounting workload, the change process evolved and so did the  
content of the changes it produced.

Essay Two: Sustaining Organizational Change: Simulation Analysis of the Tipping  
Point in the Dynamics of Process Improvement

Organizations often adopt improvement programs driven by involvement from front-line  
workers who generate improvement ideas. These ideas create demands for support  
personnel (e.g., engineers, skilled tradespeople, and managers) with limited availability.  
This paper explores how implicit and explicit policies governing how implementation
tasks get done influence dynamic patterns of implementation over time. The study is grounded in participant observation at a manufacturer implementing lean manufacturing improvements. Workers generate improvement ideas more quickly than they get implemented, creating an increasing backlog of tasks for support personnel. Support personnel can accomplish tasks through working collaboratively with workers, or alternatively, they can perform the work on their own. This essay uses a system dynamics model to explore the dynamic consequences of the two means of implementing ideas. The collaborative approach fosters learning among workers that builds understanding of work processes and skill about collaborating effectively, which leads to benefits in subsequent collaborative efforts.

The essay is organized as follows. The first section describes the observed pattern of "start and fizzle" that will serve as a reference mode of the dynamic behavior to understand, presents a dynamic hypothesis in the form of a causal loop diagram, and explains its rationale. The second section develops a formal system dynamics model of the underlying feedback structure. The third section presents simulations and analysis of the model's dynamic behavior. The final section discusses some implications for theory and practice.

Essay Three: The Right Shock to Initiate Change: A Sensemaking Perspective

Changes in patterns of organizing often follow disruptions, so-called shocks, in the way that people understand their organization and organizing practices. People engage in a process of sensemaking, and new meanings emerge to guide future practices. Yet, organizations often deal with such shocks in ways that do not lead to any fundamental differences in the observed patterns of organizing practices. Managers and scholars often conclude that such shocks were simply not "big enough" to occasion change. The purpose of this paper is to raise awareness that shocks that are "too big" may also fail as occasions for change. The paper develops a framework that describes how organizational activity following a shock might unfold in a manner that does not occasion cognitive restructuring and organizational change. Four alternative means of disposing of shocks are discussed. Organizations might not notice these shocks, might not take action, might not take novel action, or might not undergo cognitive restructuring. Shocks, which may arise from both external and internal sources, may be either too small or too large relative to the organization's current cognitive schema. The paper juxtaposes the four means of disposing of shocks against a dimension representing the magnitude of the shocks to form a framework used for discussion of how small shocks and large shocks might occasion each of these four modes of disposition, drawing on relevant literature and providing examples. An important implication for the practice of organizational change is that agents who wish to induce organizational change should benefit from attending to cues from their organizations that indicate whether shocks to occasion sensemaking are too small or too large.

Thesis Supervisor: John D. Sterman
Title: Professor of Management Science
Acknowledgements

My doctoral studies have been an experience in learning, about theory, about research, and about myself. I owe a great debt of gratitude to the many people who have helped to make the experience rewarding and assisted me in countless ways. I pause here to acknowledge the special contributions of a select few people.

It is truly exhilarating to have four great minds attending to my thoughts and ideas with an eye towards shaping and strengthening them. I am quite fortunate to have for my thesis committee four people who are all distinguished scholars and who are also all people with whom I truly enjoyed spending time.

John Sterman helped in many ways. As chairman of my thesis committee, he was masterful. He is a gifted teacher from whom I have learned much, both in the classroom and out. He set an example with his careful attention to detail, analytical rigor, clarity of presentation, and zealous pursuit of truly interesting findings. I can only hope that to some small extent my own work reflects the high standards he showed me. He more than anyone kept me focused on the goal of completing the dissertation. Without his frequent admonishments to do "nothing but thesis" I would have fallen prey to even more of the many distractions that diverted my attentions.

Nelson Repenning has been my guide and mentor from the day I entered the doctoral program. He taught my first class in system dynamics and has continued teaching me much since then. Nelson was the "junior" faculty on my committee, but there was nothing junior about his contribution. If John Sterman was my Thesis Chairman, then Nelson was my Thesis Advisor. He was always available, quick to read and comment on the many iterations of the models and written drafts, especially helpful in preventing me from getting lost in my model or my abstract thoughts, and steadfast in his confidence and support for what I was trying to do. Nelson belongs in the hall of fame of thesis advisors for many reasons, but most notable for me is his skill as a coach. Nelson never told me what to do, but somehow he got me to do things that were useful. The many modeling sessions and discussions I have had with him are the highlight of my doctoral experience and directly responsible for most of what might be interesting in this dissertation. He was always the first one to whom I turned for advice and commentary, and he always delivered.

Wanda Orlikowski opened my mind to the merits of self-reflection and helped me to understand how important our own worldviews are in shaping our thoughts and experience. She has a magical ability to make a generative comment. She is tireless in her intellectual curiosity and seemed so often to find a fresh and interesting way of looking at something familiar in my work. My conversations with her frequently sent me away pondering a new intellectual challenge. She raised my awareness of many assumptions I had been making, allowing me to question them, articulate them, or appreciate how they influenced my thinking. Wanda was also a role model for scholarly character and integrity that I hope to emulate. She is genuine and sincere in her own
beliefs, yet always shows the utmost respect for the work of others working within different paradigms.

Michael Tushman accepted my request to "keep me honest" when I asked him to join my thesis committee. He believed in the merits of this work from the start and pushed me to make it better, more complete, more interesting, more relevant, and more central to organization theory. His instincts about where to search for the compelling insights and important contributions were invaluable guides. I also thank him for his perceptive advice regarding scholarly development and career management.

I gratefully acknowledge the participation of the respondents at my field site, whose willingness to allow me to observe them in their daily activities was essential for me to do my work. It was fun for me in large part because of their collegiality and sense of humor. I offer a special note of thanks to Don Kieffer, for inviting me to do this work and for all he did in support of the effort.

My fellow doctoral students did so much to help, each in their own way. Paulo Goncalves was a constant source of friendship and support from my start at Sloan, sharing many of the joys and frustrations of being a doctoral student. It has been my distinct pleasure to have him by my side through the many ups and downs over the past years. Laura Black also filled this role from the start, and then offered an additional motivation by showing me that it was indeed possible to finish. Scott Rockart was the only system dynamics doctoral student in residence when I arrived. His sagely advice, especially about navigating the terrain of Sloan, was enormously helpful. Mila Getmansky, Hazhir Rahmandad, Jeroen Struben, Gokhan Dogan, along with Paulo, Laura and me, were the SD Kids. They offered ongoing moral support and helpful commentary and were just plain fun to be with. Mike O’Leary and Sachi Hatakenaka, my classmates in the Organization Studies Group, were cheerleaders and inquisitive colleagues whose interest and critical thinking yielded many questions that helped me.

Jim Hines influenced me in many ways. His unbridled enthusiasm for system dynamics was contagious and most certainly infected me. He taught me to deeply appreciate the value of insight and how much we can learn from models and modeling. He was a constant source of encouragement and support and became a friend I will always cherish. Anjali Sastry came back to MIT late in my doctoral program and quickly took an interest in my work. She became a valued source of advice, helping to shape my research and my development as a scholar. Through his writing and speaking, Peter Senge helped me discover system dynamics and see its value in the study of organizational phenomena. I owe a debt also to Jay Forrester for his pioneering work some 45 years ago to give birth to the field of system dynamics. I am grateful that he generously shared his time with today's corps of doctoral students, giving me the rare opportunity to learn the ethos of our field directly from the founder.

Grandma Seagrave impressed upon me from a young age the importance of being able to communicate my thoughts. With intensity befitting only a grandmother, she has been perpetually interested in everything I do, giving me plenty of practice in the art of
explaining. My sisters Barbara and Betty have been with me in spirit in everything meaningful I have done in my life. Betty has always been a benchmark for selflessness and thoughtfulness to which I can only aspire. Barbara has been my soul mate for decades, occasionally perking me up when I have needed it but more often keeping me in check when I have needed it more. She has always made sure that I don't lose sight of what really matters. My parents were always there to celebrate small successes and console during the disappointments. My mother's generosity and compassion seems to know no limits. She showed me the power of believing in possibilities and allowed me to dream. Her unwavering confidence in me endowed me with the courage to try, to stumble, and to learn.

The most deeply personal note of gratitude goes to my Dad. The only way I can find to express in part what he has done for me is to say that through my eyes he has always been first and foremost a father. As I bring this dissertation to fruition, I am acutely aware that it is he who taught me the virtue of humility, without which no learning would be possible. I vividly recall the many dinnertime conversations of my childhood in which he nurtured my incessant desire to ask, "Why, why, why?" He has unobtrusively followed my progress through the doctoral program and, with the wisdom of one who has done it himself, offered understated words of encouragement and guidance that have kept me moving in the right direction.

My wife Sally is worthy of superhero status. Sally has been a true partner in the accomplishment of this thesis, willingly and lovingly walking with me every step. I have become so accustomed to the wonders of having her at my side that I am surely guilty of taking for granted the many things she has done to help. I thank her for all she has done, but especially for having the good sense to know when to leave me alone with my work and when instead what I really needed was for her to interrupt and say, "Honey, you need a hug."

With profound gratitude for all they have done, I dedicate this work to my Mom and Dad and my wife Sally.
To Mom, Dad, and Sally
Table of Contents

Essay One
Explaining the Start and Fizzle of Organizational Change: Co-evolving Process and Content

Abstract ..................................................................................................................... 11
1. Introduction ......................................................................................................... 13
2. Methods ............................................................................................................... 17
   Data Collection .................................................................................................... 18
   Data Analysis ....................................................................................................... 20
3. Overview of the Implementation Challenge ...................................................... 21
   The Research Setting .......................................................................................... 21
   Stalled Change: The Pattern of Start and Fizzle .................................................. 22
4. A Model of Organizational Change .................................................................... 23
   Learn a Little, Do a Little .................................................................................... 23
   The Process of Doing a Little ................................................................................ 28
   Do A Little and Do A Little More: Responses to a Backlog of Work ............... 33
   Do It Alone, Learn A Little Less: Consequences of the Shortcuts ................. 40
   Responses of the Work Group ............................................................................ 45
   Interaction with the Managers ............................................................................ 46
5. Discussion and Implications ............................................................................. 53
6. References .......................................................................................................... 59

Essay Two
Sustaining Organizational Change: Simulation Analysis of the Tipping Point in the Dynamics of Process Improvement

Abstract ..................................................................................................................... 67
1. Introduction ......................................................................................................... 69
2. Dynamic Hypothesis ........................................................................................... 71
3. The Model .......................................................................................................... 77
   Process Capability .................................................................................................. 77
   Task Accomplishment ........................................................................................... 78
   Support Personnel Time Spent Doing Tasks ....................................................... 81
   Productivity ............................................................................................................ 83
   Improvement Idea Quality ..................................................................................... 85
4. Behavior of the Model ....................................................................................... 89
   Good Ideas that Lead to Ruin ............................................................................... 89
   New Ideas, Extra Resources, and the Start and Fizzle Outcome ...................... 96
   Temporary Success: Benefits of Early Collaboration ....................................... 99
   Sustained Success: Engaging the Positive Loop ............................................... 106
5. Discussion .......................................................................................................... 111
6. References .......................................................................................................... 115
Essay Three
The Right Shock to Initiate Change: A Sensemaking Perspective

Abstract ................................................................................................................. 119
1. Introduction ....................................................................................................... 120
2. The Framework .................................................................................................. 123
   - Shocks that Get No Attention .................................................................. 130
   - Shocks that Get No Action ..................................................................... 132
   - Shock that Get No Novel Action ............................................................. 137
   - Shocks that Do Not Occasion Cognitive Restructuring ......................... 140
3. Implications ...................................................................................................... 143
   - Implications for Theory and Research .................................................... 144
   - Implications for Practice ......................................................................... 146
4. References .......................................................................................................... 147
Explaining the Start and Fizzle of Organizational Change:
Co-evolving Process and Content

J. Bradley Morrison

Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA
Morrison@mit.edu
(617) 354-9585

ABSTRACT

The theme that sustaining efforts to change is difficult appears in many guises throughout the organizations. The tendency for change processes to run out of energy and momentum is widely recognized. In this essay, I offer an inductive study of the dynamics of organizational change that describes situated human interactions to explain the observed trajectory of organizational behavior. As a window through which to view change, I study a manufacturer adopting production practices based on concepts of lean manufacturing and the so-called Toyota Production System (TPS). I develop a grounded theory based on extensive fieldwork observing a change effort that began with temporary improvement but subsequently entered a phase of decline, the characteristic start and fizzle of change. I examine how the actions of various groups interact with each other and with characteristics of the workplace. I then induce a model that characterizes the dynamics of participatory change. The model draws on a close examination of how people do the work of process improvement to identify a set of mechanisms through which process and content interact to constitute the patterns of organizing that result. I find that front-line participation led workers to generate ideas, but the work of implementing those ideas placed demands on key support personnel such as manufacturing engineers. The support personnel quickly become overwhelmed, facing a growing backlog of ideas to implement. As the support personnel modified their work practices to address the mounting workload, the change process evolved and so did the content of the changes it produced.
"It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of light, it was the season of darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us..." (Charles Dickens, A Tale of Two Cities, 1859).

When Charles Dickens penned what was to become one of the most famous of all opening lines in English literature, he was writing about conditions in Victorian England, not about the state of affairs in the modern organization. Yet, his intriguing opener captures a fundamental paradox relevant to life in organizations. Things may be going well and at the same time going poorly. And even more vexingly, it is often difficult to know how things are going. Gareth Morgan (1986, p. 339) proclaimed: "Organizations are many things at once!" Such is often the case when organizations attempt to implement change. Consider for example a common co-occurrence in the dynamics of organizational change: a change process starts and subsequently fizzes. The organization achieves some short-term objectives, yet simultaneously fails to establish an underlying capability to sustain improvement (Beer & Eisenstat, 1996). It is the best of times based on immediate results, yet it is also the worst of times as revealed by the subsequent course of events.

The theme that sustaining efforts to change is difficult appears in many guises throughout organization theory. The tendency for change processes to run out of energy and momentum is widely recognized (Beer, Eisenstat, & Spector, 1990; Pettigrew, 1998; Pettigrew, Woodman, & Cameron, 2001; Repenning, 2002). Quite simply, "it is hard to make changes stick" (Kanter, Stein, & Jick, 1992, p. 5). Temporary adoption followed by ultimate rejection of new behaviors necessary to achieve enduring change is a common dynamic pattern in change initiatives (Armenakis & Bedeian, 1999). In studies of innovation implementation, analysts increasingly identify unsuccessful implementation rather than ineffectiveness of the innovation itself as a cause of failure to achieve intended benefits (Hackman & Wegeman, 1995; Klein & Sorra, 1996; Repenning, 2002). Some authors refer to gaps between knowing and doing in organizations (Pfeffer & Sutton, 2000) or an improvement paradox as good ideas do not reliably help
organizations achieve improved performance (Sterman, Repenning, & Kofman, 1997). Others document the frequency of abandoning business tools (Rigby, 2001; Rigby, 1994) or describe the faddish nature of administrative innovations (Lawler & Mohrman, 1985). More broadly, the liabilities of undertaking organizational change have been implicated in organizational failure rates (Amburgey, Kelly, & Barnett, 1993; Hannan & Freeman, 1984).

Despite the attention of scholars from a wide range of disciplines and the pervasive concern among practicing managers, the difficulty organizations have putting administrative innovations into use remains a key issue facing organizational theory (Pfeffer, 1997). Scholars concerned about the limits of existing theory of organizational change have called for research to resolve two deficiencies in our current theories. First, there is a need for increased attention to both the process and content of change. Content refers to "what" changes, and process refers to "how" change occurs (Barnett & Carroll, 1995). "Enough research has been conducted on organizational change to make it clear that, in most contexts, both content and process factors ought to be evaluated. Yet theories and analyses of organizational change often tend to only one dimension." (Barnett et al., 1995, p. 210). "One particularly worthwhile domain in which process and content concerns converge is the translation of knowledge into action. There are many instances in which organizations know what to do (content) but have difficulty in actually implementing that knowledge (action). ... The challenge for organization studies in the future is to find ways of understanding the connection between content and process, between knowledge and action, and between theory and practice" (Pfeffer, 1997, p. 202). Pfeffer suggests that such a line of inquiry will lead to better understanding of "why well-informed individuals and organizations within them pursue ineffective activities and promote dysfunctional policies and practices" (Pfeffer, 1997, p. 202).

The second deficiency is a shortage of attention to the link between macro and micro perspectives. Organizational theorists have been intrigued for some time by the observation that apparently small inputs or events can lead to dramatic outcomes. Scholars in many other domains such as economics, system dynamics, and physics have
also pointed out similar connections between small initial changes and largely different consequences (Arthur, 1989; David, 1992; Forrester, 1961; Nelson & Winter, 1974; Prigogine & Stengers, 1984). The phenomenon is sometimes described as amplification or the enlargement of small cues (Weick, 2000). The scholarly call is to shift more attention to the ordinary activities in organizations in search of explanations for interesting outcomes. "Because of the magnitude of some changes in organizations, we are inclined to look for comparably dramatic explanations for change, but the search for drama may often be a mistake. (March, 1981, p.564). Echoing the need, Weick noted that "macro perspectives are hollow unless linked with micro dynamics" (Weick, 1993b, p.34).

To address some of these deficiencies, I offer an inductive study of the implementation of innovation that details how key feedback relationships that arise from situated human action explain the observed trajectory of organizational behavior during change. The central research question guiding this study is to understand how the actions of various groups in organizations interact with each other and with physical characteristics of the workplace to give rise to the observed pattern of behavior. What physical and social structures and processes influence how situated action at a micro level contributes to organizational outcomes? How do the behaviors of social actors in one organizational locale interact with those in others to yield observed patterns of organizational behavior?

Organizational researchers over the past decade have made some inroads into understanding the connection between situated action and observed macro outcomes. Scholars following the structuration perspective have been increasingly interested in models of organizations based on notions of recursive interaction between organizational actors and the resources available to them (Giddens, 1984; Orlikowski, 1992). Some work has been concerned with understanding the appropriation of various aspects of technology or the enactment of technology-in-use as users go about their day-to-day activities, thus conceptualizing emergent change in organizations (Barley, 1986; DeSanctis & Poole, 1994; Orlikowski, 2000; Orlikowski & Gash, 1994). Other work has used improvisation as a metaphor for organizational activity, emphasizing the unplanned
and ongoing nature of organizational activity (Crossan & Sorrenti, 1997; Hatch, 1997; Orlikowski, 1996; Weick, 1993a; Weick, 1998). These scholars view change as evolving through the course of ongoing organizational action. Such emergent change is not anticipated or planned, but rather is the realization in action of a new pattern of organizing in the absence of explicit, a priori intentions (Orlikowski, 1996; Weick, 1995; Weick, 2000). Studies in the emergent change school have been largely descriptive in nature, concerned mostly with establishing the existence of the emergent change phenomenon. They describe changes over time that lead to organizational situations or states that do not correspond to those that were anticipated or intended at the beginning of the change process. The authors offer these descriptions as examples of emergent change and generally distinguish these processes from alternative views embodying a designed or planned outcome, one that is anticipated or intended by leaders or technology designers.

An emerging stream of research explicitly applying feedback theory to understanding phenomena in organizational change serves as another foundation for the present work. Early work in this tradition applied a feedback lens to investigate phenomena such as the growth and demise of a corporation and the theory of punctuated equilibrium (Forrester, 1968; Hall, 1976; Sastry, 1997; Tushman & Romanelli, 1985). More recently, scholars have studied the challenges of sustaining process-focused improvement technologies, such as total quality management (TQM) and business process reengineering, in both manufacturing and product development (Keating & Oliva, 2000; Oliva, Rockart, & Sterman, 1998; Repenning, 2002; Repenning & Sterman, 2002; Sterman et al., 1997). Repenning and Sterman (2002, p. 292) describe how difficulties that organizations face in the implementation of improvement are "rooted in the ongoing interactions among the physical, economic, social, and psychological structures in which implementation takes place." The study calls attention to the critical importance of the interactions between physical characteristics of the workplace and the actions and beliefs of organizational actors. However, like the previous work, the authors treat the work of improvement activity at an aggregate level in order to focus on tradeoffs between activity in production

16
and activity in improvement. A close examination that includes the work of doing improvement activity at a micro level is still needed.

In this paper, I use the implementation of planned change through process improvement as a window through which to study connections between the process and content of change and between situated action and organizational trajectories of change. Process improvement efforts are a useful context in which to study how micro processes contribute to broader organizational change (Repenning et al., 2002). Dynamic behavior during implementation of such organizational improvement activities often exhibits a pattern characterized by an initial burst of successful activity followed by a decline in performance levels. "Start and fizzle" describes a problem in the sustainability of organizational change initiatives that is of interest to both scholars and practitioners. I develop a grounded theory that explains the start and fizzle pattern, drawing on data from field work I have conducted at a manufacturing firm as it attempted to adopt production practices based on concepts of lean manufacturing and the so-called Toyota Production System (TPS). I begin with a description of research methods followed by background on the research setting and an overview of the change initiative. In the following section, I develop an explanation for the observed pattern -- start and fizzle -- by examining how the actions of various groups (e.g., managers, production workers, and other shop floor workers) interact with each other and with the physical characteristics of the workplace. I then induce a model of these critical interactions and use it to explain some consequences of these interactions for organizational outcomes. I conclude with a discussion of implications for managers undertaking change efforts and for scholars concerned with organizational change.

METHODS

The organization I observed is an automotive vehicle manufacturer in the United States. The site of the research is a plant that manufactures parts and assembles engines used in the company's vehicles. In the spring of 2000, the general manager of the plant became interested in "taking [the company] to a new level" of manufacturing performance. He
began efforts in his plant to learn about and adopt new practices, drawing heavily on
approaches identified as “lean manufacturing” (Womack, Jones, & Roos, 1990) or the
Toyota Production System (Monden, 1983). I had the opportunity to begin data
collection with the general manager and the plant in the early stages while the general
manager was beginning to think about how to introduce new production practices to the
company.

Data Collection
The field research began in June of 2000, as the general manager was beginning to
formulate his approach, and continued for 20 months. Data collection comprised
ethnographic material, documentary and archival data, and in-depth interviews
(Pettigrew, 1990). My early efforts were focused on understanding the context, setting,
and culture within which the focal change efforts were taking place. I spent several days
shadowing the general manager and the production managers (direct reports of the
general manager) of the facilities. I attended at least one of each of the various meetings
that take place regularly at daily, weekly, monthly, quarterly, and yearly intervals. I
attended several dozen daily production meetings, known as "hot meetings," in which
representatives of the various production units and support activities meet at the
beginning of the morning and evening shifts to report daily status information and to
coordinate activities. I attended many other meetings related both to ongoing plant
activities and to the implementation initiative. In some cases, I made audio recordings of
the sessions. For many others, I took detailed field notes. I also participated in a group
organized by the general manager comprising representatives from both salaried
management and elected union leaders at each of the company's six production facilities.
The group was chartered to lead efforts across the enterprise to adopt new and better
production practices and met monthly for two-day sessions. I attended all of these
monthly meetings. My later efforts brought me on to the shop floor observing and
engaging in conversations with workers. I followed the production of key component
parts from start to finish through a production cell and the assembly of an engine along
the entire assembly line, talking with workers at each step. I spent upwards of 100 days
on site, observing managers and workers in their daily activities. Over the course of this
field research, I have had many informal conversations or unstructured interviews with informants, recording these conversations in field notes. I gathered documents, electronic files, and emails generated during the course of the research, among other archival data. I also conducted a series of semi-structured interviews with management personnel, union leaders, and front-line workers. The data I gathered includes more than 1200 pages of field notes and over 200 hours of audiotapes of meetings and interviews.

I interviewed selected individuals who were directly involved with the change initiative. Respondents included the plant general manager, the plant production managers, other members of the plant management team (direct reports of the general manager), both work group advisors for the production cell selected as the pilot area, elected union officials, hourly production workers, members of an implementation team assigned to the change initiative, and other support personnel such as a plant engineer. The interviews lasted from 45 minutes to 2 hours, and some respondents were interviewed more than once. The interviews generally began by asking respondents to describe what they first remembered about the focal change initiative and then requesting they construct a timeline of the course of events. During these accounts of the change initiative, they were asked to describe their own roles, the factors that helped or hindered the efforts, and their impressions about progress and the success of the initiative. I often asked for their opinion about the persistence of the problems that they identified, which in hindsight seemed somewhat obvious. Finally, I asked them how the initiative had affected them personally, in particular what they may have learned through the experience. The interviews were audio taped, then transcribed.

Together, these methods draw on the strengths of the various methods and facilitate verification through an iterative process. In particular, the interviews provided opportunities for in-depth examination and the use of audio recording for accuracy, while the extensive direct observation allowed for comparisons between what people said in interviews and casual conversations and what they actually did (Pettigrew, 1990). These methods contribute to a rich database that I have developed comprising longitudinal data that includes both contemporary and retrospective reports from respondents. The
resulting data meet four of Pettigrew's (1990) five ideal characteristics for longitudinal field research on changes: the data are processual, pluralist, historical, and contextual. The data fall short on the fifth characteristic because they are not based on a range of studies for comparative purposes.

Data Analysis
As Barley (1990, p. 234) notes, "the analysis of field data actually begins during a study's observational phase." The ongoing analysis of field data offers the possibility of developing interim hypotheses and directs attention to relevant data (Glaser & Strauss, 1967). My data analysis followed traditional qualitative methods (Eisenhardt, 1989; Miles & Huberman, 1984; Yin, 1994). I also made frequent use of a graphical representation technique known as causal loop diagrams that is especially useful in the analysis of complex systems (Masuch, 1985; Sterman, 2000; Weick, 1979). The diagrams capture the feedback structure of a system of interacting elements and help to reveal the potential dynamic behavior of the posited relationships. During the course of my fieldwork, I made frequent sketches of such causal loop diagrams, often in the evenings during my trips to the research site. The diagrams were one explicit means by which I was "always trying to make sense of one's data and thinking about what more one can find out" (Feldman, 2000, p. 615).

Data analysis included listening to the recorded interviews and reading the transcriptions, coupled with a review of field notes. I identified patterns of interest and recurring themes in the data, bounding the analysis with a focus on efforts to implement change in the first production cell. As is typical in developing grounded theory, I organized the data into categories, which I represented with variables and causal relationships between them (Glaser et al., 1967). I combined variables and causal relationships to begin identifying causal loops as a description of the feedback processes gradually emerging from this analysis. During the data analysis, I occasionally translated portions of the emerging feedback structure into formal mathematical models and simulated their behavior in order to gain a richer understanding of the relationship between the feedback structure and the dynamic behavior. The iteration between the grounded data, causal loop diagrams, and
formal mathematical models led to additional insights and generated new questions that I could explore in the available data or pursue with my respondents. On occasion, I reviewed interim results of my analysis with members of the plant management team, who often identified examples that were useful to fill in some gaps. My data analysis approach follows methods used by other researchers applying a feedback lens to the study of organizational phenomena (e.g., Perlow, Okhuysen, & Repenning, 2002; Repenning et al., 2002).

OVERVIEW OF THE IMPLEMENTATION CHALLENGE

The Research Setting

The company I observed manufactures and markets a line of motor vehicles and related products targeted at automotive enthusiasts. The company has a rich heritage and a well-known brand. Market demand for its product continues to be strong. Despite this success, management recognizes the need to stay competitive. It is commonly acknowledged that the company's manufacturing costs are higher than those of their competitors.

The engine plant, like the company's other manufacturing facilities, employs a unionized workforce. Management and union leaders alike attest to a strong, cooperative relationship between the union and the management. For several years, work practices have been guided in part by the collective bargaining agreement, which details elements of the organization structure and responsibilities of the union leaders and workers and the salaried management. Elements of the agreement include work groups organized around production units, "work group advisors" rather than supervisors, joint representation in many decision processes, and a high degree of information sharing with employees, as well as other practices typically characterized as high-involvement work practices (Pil & MacDuffie, 1996).
Stalled Change: The Pattern of Start and Fizzle

Managers in the plant recognized a need to improve the manufacturing capabilities in their organization. They discussed the idea with union leaders at the plant and agreed to proceed. As one union leader said:

I came to understand that we are archaic in the way we do our manufacturing here. We all know that competition is always knocking on our door. We know that if we’re going to be competitive, especially here at this plant, because we are such a low-budget, low-profit margin, here, that we’re going to have to continuously improve to be competitive, and keep the [product] at a price where we can get people, ... first time buyers, into our family.

The workers throughout the plant were already organized in work groups corresponding to manufacturing cells that produced various engine parts. The plant managers and union officials jointly selected one work group to be "a pilot area that we're going to try this on" in the words of a production manager. The work group runs a production cell that comprises machining and some assembly operations to make several parts used in the company's engines. After several months of work in the pilot area, the management and shop personnel were proudly pointing to the initial success of the effort: A union official describing the early progress said, "It was going along pretty good there. The area was starting to really look uniform over there." Enthusiasm was high, as one hourly employee said:

The people that were there, they seemed very excited about it. When they saw the results, and what could happen, got a visual look, basically, of how we're doing things today and how it can actually be, they got kind of excited about it.

And the work group was beginning to show tangible evidence of business results, as noted by an engineer on the implementation team:

They got people looking at the machines, finally fixing them. We got the layout running. You got a pull system in place. And – start looking at some of the overall numbers – they're outstanding. Scrap has come down [thousands of dollars]. Performance went up from 70% to 94% [in-stock levels of finished goods inventory meeting or exceeding target minimums].
Yet several months later, some new work practices had been abandoned and performance had deteriorated. One informant described the situation as "the wheels are coming off." Another said, "If you go over there a couple of months later, after all this stuff, the [physical appearance] started to deteriorate." Several respondents reported similar characterizations of the improvement activity in this work group: an early phase of improved performance followed by a plateau and then decline, the start and fizzle pattern. As noted in the introduction, similar patterns are well documented in the literature on organizational change. The purpose of the following analysis is to develop a dynamic model that captures the ongoing interaction between the work of improvement activity and the organizational context. The model describes a processual theory that explains how the recursive relationship between process and content gives rise to the start and fizzle pattern observed in the improvement initiative of the pilot work cell.

In what follows, I draw on my field data to induce a model that describes how the interactions of various organizational members and the physical characteristics of improvement activity can generate the observed pattern of behavior. I represent the model using causal loop diagrams to describe the feedback structure enacted by the organization and discuss how this structure gives rise to the observed pattern of start and fizzle. I develop the model in stages, presenting the relevant data and then interpreting the data in diagrammatic form.

A MODEL OF ORGANIZATIONAL CHANGE

Learn a Little, Do A Little

A distinguishing feature of most participative improvement programs is that employees working in their own situated environments make significant contributions to improving the work processes that they perform on an ongoing basis. "The core of a high-performance work system (HPWS) in manufacturing … is that work is organized to permit front-line workers to participate in decisions that alter organizational routines. This may be achieved by using shop-floor production teams or through employee participation in problem solving or quality-improvement teams and statistical process control" (Appelbaum, Bailey, Berg, & Kalleberg, 2000). In TQM companies, employee
involvement mechanisms are invariably introduced. “According to the Conference Board (1991), 65 percent of TQM organizations create employee suggestion systems, and 70 percent have quality meetings between managers and employees and/or focus groups to solicit ideas about quality” (Hackman et al., 1995, p. 317).

Similarly, authorities on lean production and the Toyota Production System assert the importance of the front-line employee as the key contributor to the problem solving at the center of process improvement. After studying some forty production plants in the United States, Europe and Japan, Spear and Bowen (1999) attempted to discern the essence of the Toyota Production System:

“Indeed, in watching people doing their jobs and in helping to design production processes, we learned that the system actually stimulates workers and managers to engage in the kind of experimentation that is widely recognized as the cornerstone of a learning organization. That is what distinguishes Toyota from all other companies we studied” (Spear & Bowen, 1999, p. 98).

This ongoing experimentation leads both to process improvements and to increases in the skill, motivation, and involvement of the work force. “Indeed, the term ‘lean production’ … captures the minimization of buffers but neglects the expansion of work force skill and conceptual knowledge required for problem solving under this approach” (MacDonald, 1995, emphasis in original). Moreover, as the skill of the workforce increases, workers are able to generate and implement more improvement ideas. As they do so, they continue learning, acquiring even more skill, generating even more ideas, and thus more learning as the process continues.

In the organization I studied, the general manager and many other managers espoused their intent to follow an approach of learning and discovery rather than a top-down, directive approach to change. The general manager expressly articulated his belief that the workforce must be actively involved and that people learn by doing. One production manager described the strategy they chose to follow:

I think that the strategy … that we took here was to try to create demand for this [program]. Let’s educate people. Let’s give them some knowledge. Let’s train them and let’s encourage their intellectual
curiosity and then let’s channel their thoughts and their ideas into things that are consistent with the tools of lean manufacturing and the process of lean manufacturing.

Moreover, the managers also recognized that they themselves were not experts. Thus, they decided to begin with the idea that they would learn along the way. The production manager continues:

What are the first steps? What do we do? Do we just go down there and say – ok folks, we’re going to implement lean and we’re the experts here? We’re going to tell you what to do and you just trust us. We won’t mislead you. We may make a few mistakes, but just stick with us. Do what we tell you and we’ll be ok – all right? And we said – no, we can’t do that. We don’t have enough knowledge to go down there, ... succeed, and create a demand for this [program] that we hope to create. So how do we do this? Well, what if we all learn this stuff together, you know? And that’s where the bus analogy came from. Let’s get everybody on the same bus and let’s all agree where we are today and get some fuzzy idea about where we want to go. Then let’s see what we can learn along the way to bring clarity to where it is that we want to go. We may make a few left and a few right turns along the way. We may have to back up the bus a couple of times here and there. So that's where [someone] came up with the concept – learn a little, do a little. Let’s teach these folks a little bit. Let’s learn a little bit about this stuff ourselves. Let’s put our heads together. Let’s decide what our needs are and what our priorities should be within this workgroup and see if we can’t apply some of these tools and some of these skills we’ve learned to bear on some of the issues that we identify.

The plant managers and union officials jointly selected one work group as the pilot site for the initiative. The selected work group comprises approximately 15 hourly production workers who run a machining and manufacturing cell that produces several parts used in the engines assembled in this plant and at another engine assembly facility in the company’s production network. Due in part to quality and cost problems, the company had recently considered discontinuing the in-house manufacturing of these parts and had begun purchasing a portion of their needs from a supplier. Plant management and the union leaders had convinced the company to continue internal production of these parts, but there was a clear sense that performance improvement was needed or the company would indeed discontinue the operation. Moreover, because some production
volume had recently been shifted to outside suppliers, there was a surplus of worker time that could be dedicated to improvement activity. Thus, this cell was chosen because the need for improvement was clear and because resources for improvement were available.

The managers formed a full-time implementation team, dedicating people to work with the pilot work group. Two of the team members were engineers from an internal consulting group in the company that works with the company's manufactured parts suppliers to implement lean manufacturing practices and who also had experience with lean manufacturing at previous employers. The manufacturing engineer already responsible for that area was assigned to the team. The work group advisor – the plant has no "supervisors" – had significant experience with lean manufacturing from a previous employer and was quite enthusiastic. The team recruited two volunteers from the work group to work three-fourths to full time on the implementation. The union asked an hourly worker, whose job title is union analyst, to join the team. A union analyst examines work processes and contributes to setting quantitative standards for work output in the various production jobs. The team also had some support from an external consultant who conducted some training and provided expertise in lean manufacturing techniques.

The implementation team provided some training for the workers and then asked them to generate ideas for ways to improve the manufacturing processes. Lean manufacturing systems aim to reduce various forms of waste in the system, often by implementing processes such as just-in-time production, kanbans, production smoothing, improved process layouts, reduced set-up times, and standardization of operations (Monden, 1983; Womack & Jones, 1996; Womack et al., 1990). The early training was focused largely on improvements such as workplace organization and redesign of the production layouts. Specific tools, analytical techniques, or structured problem-solving methods were used to help generate these improvement ideas. An hourly production worker described the prevailing approach of high employee involvement:

I mean I’ve worked at a number of other companies. [This company] really goes out of their way for employee involvement. It’s not just lip
service. Some of the ideas we discussed as a group, but a lot of the stuff that we came up with came from the hourly employees. Roller racks was one of the things that we introduced. Where it's essential for some of the products that we use to be first in, first out. Actually, the roller racks were designed by hourly people that actually worked on the jobs. So that was something. ... The company says you guys use it, you tell us what you want. So they did give us that leeway and that was very helpful as far as myself doing my job. ... Let’s make these things easier and more efficient, of course. That was one of the big things.

The hope was that as ideas got generated and implemented, workers would learn and improvements would accrue, setting in motion a process of continued improvement and learning propelled by the involvement and knowledge of the workers.

Figure 1 shows the way that improvement activity contributes to process capability. Workers contribute by Generating Ideas, which accumulate in a stock of Tasks To Do. The rectangle icons in Figure 1 represents stocks, which are integrations (accumulations) of inflows that increase the stocks less outflows that decrease the stocks. The pipe-and-valve icons represent the flows into and out of the stocks. Thus, Generating Ideas increases the stock of Tasks To Do, and Task Completion Rate decreases the stock. The diagrams I develop here use stocks and flows along with feedback loops to represent the system structure. Stocks give a system memory and inertia, create delays, and are critical to the dynamics of the system. As accumulations of past actions, stocks influence future actions which in turn influence the rates of flow that change the stocks, thus closing the feedback loops (Repenning et al., 2002).

As the work proceeds, Task Completion generates Process Improvements that increase the Process Capability. Figure 1 represents Process Capability as a stock. Process Capability is the ability of the firm to generate output that is valued by the market. In this case, Process Capability refers to the throughput of manufactured parts that the work cell can accomplish for each unit of production labor employed. The stock of Process Capability is increased by Process Improvements and decreased by Process Degradation, a natural consequence of ongoing use and changes such as customer requirements and supplier factors. The diagram shows that ideas lead to improvement tasks, and the
successful completion of improvement tasks results in the intended increase in the organization's capability. *Process Capability* increases as long as *Process Improvements* exceed *Process Degradation*. The representation of *Process Capability* as the cumulative difference of improvement and degradation follows earlier models of process improvement (Repenning & Sterman, 2000; Repenning et al., 2002).

*Figure 1: Basic Stock and Flow Map of Lean Manufacturing Process Improvement Activity*

[Diagram showing the flow of tasks and their associated improvements and degradation]

Stocks are represented by rectangles, and flows are represented by "pipes with valves." A stock is the accumulation of the difference between its inflows and outflows (see Sterman, 2000).

**The Process of Doing a Little**

An important characteristic of many of these improvement suggestions was the varied nature of the tasks required to implement them. Although the vast majority of the ideas originated as suggestions from the work group and the dedicated implementation team, the tasks required to execute these ideas frequently required assistance from other people in the plant. One manager summed up the pattern regarding who gets assigned responsibility for doing the tasks: "Very rarely is it the person that came up with the idea."

Consider for example the efforts to organize the finished goods produced by the work group by establishing and properly configuring an area they called the supermarket. The parts the work group makes are used by two engine assembly lines, one located in the same facility and one at another manufacturing location. Better organization of the finished goods supermarket would enable better signals about what parts the engine assembly lines needed based on timely and accurate information about what is currently available, and support a simple and effective process to replace what was used as parts are drawn from the supermarket. The work group developed plans to organize the
finished parts in an area specified for the supermarket. The work plan comprised several tasks. One task was to mark lines on the floor, clearly delineating the space for the supermarket and sections within it. The size of each section would be based on the capacity deemed necessary for each of the various parts they produced. Another task was to make and hang overhead signs that would identify the parts in each section and other information such as the desired stock level and the capacity of the section. These are tasks that are normally done by personnel from the maintenance department or other skilled trades. Thus, the ideas generated by the work group created demands for work by other people in the plant. As one work group member described, this was a common occurrence: "Like just for little things like hanging signs for a supermarket, we need signs for a supermarket. We need somebody who can do that." In this example, the critical support personnel were people from the maintenance department. As one manager observed, this was a common scenario: "Maintenance. [The workers] have got a great idea for maintenance to do. And then when it doesn't get done, maintenance is bad, right?"

The work group I studied frequently needed support from the plant's manufacturing engineers. Manufacturing engineers are typically responsible for tasks such as choosing, purchasing and installing new equipment and specifying the manufacturing processes required to use the equipment. In the supermarket example, a set of such tasks was to design, purchase, and install appropriate storage containers, such as rolling carts or flow racks. Another planned improvement was to redesign the layout of the manufacturing equipment, modifying the flow of work in process through the cell. The redesigned layout included a new piece of equipment, which the manufacturing engineer would need to purchase. "Engineering" was one type of support personnel who were needed, and often not available, as noted by a member of the implementation team:

So we went up there and did the analysis, got some of the tools ordered, and did some of the basic things. Then we were starting to require more and more engineering. ... That's where it really fell down. I think probably the biggest thing we were missing on that team is we had no one to go to.
The process of matching tasks with various individual people to accomplish them was rarely given much explicit attention. Rather, the plant personnel drew upon their knowledge of the specialization and expertise of individuals, specific constraints in the union contracts, and socially understood role definitions to associate a responsibility for a particular task with an individual or group. Tasks were "engineering jobs" or "maintenance jobs." The nature of the tasks combined with a taken-for-granted understanding of who does what in such a manner that a large number of tasks fell upon a limited number of people. For example, a manufacturing engineer described his role: "So I would be responsible for taking that, [this improvement], whoever came up with it, and putting it into the actual process." The norms about who did what were powerful, yet unspoken, as evident in this description from another manufacturing engineer:

When you are sitting in a group, ... [with] some work group members and ... work group advisors ... and you are talking about how you are going to change this process and you are talking about moving machines around and talking about changing the manual work ... You knew what your part of that job was. You had to do the process documentation. You were responsible for moving the equipment and making the layouts and stuff.

Figure 2 incorporates a new link to depict how tasks get done. I use the term support personnel to refer to the various plant personnel, external to the production work group, who are responsible for executing specific tasks associated with the improvement ideas. Support personnel include the plant engineers, the maintenance department, the materials handling department, the tool room, the work group advisors, plant management, and even in some cases production workers in other work groups such as the assembly line that is the proximate customer of the subject work group. To capture the idea that support personnel must do the work required to complete tasks, Figure 2 adds a positive link between Support Personnel Time and the Task Completion Rate. The link depicts the idea that, all else equal, as more (less) Support Personnel Time is spent doing improvement tasks to support the work group, the Task Completion Rate will be higher (lower).
Figure 2: Support Personnel Are Responsible for Task Completion

The arrows connecting the variables are labeled with "+" signs to signify that an increase (decrease) in the first variable leads to a increase (decrease) in the second variable, all else equal (see Sterman, 2000).

The rectangle around the variable name *Tasks To Do* indicates that when ideas are generated, they accumulate, so we can think of the variable as a backlog of tasks that need to be done but have yet to be executed. *Tasks To Do* will grow as long as *Generating Ideas* exceeds the *Task Completion Rate*, which was the common circumstance. As one manager said, "The workers will always be able to generate ideas faster than we can implement them." Thus, the backlog of tasks continued to accumulate.

An implementation team member discussed how the group continued generating ideas:

So we just kept on plodding on with our stuff. Had the people creating things. I think we got the work group so far ahead of engineering that – it was bad, but there was no way that I was going to wait for a support organization to give me resources while I got the people engaged. You can’t. It’s like – they have to catch up.

Moreover, in addition to ideas that were consistent with the premises of lean manufacturing, other ideas such as ideas that arose simply to satisfy individual wants and desires were included and became the responsibility of these support personnel. One of the support personnel, a manufacturing engineer, noted the continuing inflow of ideas and said, "You know some of the stuff might not have even been a part of the lean manufacturing plant. It is just [that] everything is labeled under [this program]." So, the
support personnel fell further and further behind. The manufacturing engineer, for example, described the challenges he was facing as the work backlog grew:

I still saw us coming up with all these new ideas, new plans for improvements. You know we hadn’t finished the first new improvement. … You have all of these things but you didn’t get anything accomplished because you have so much you have to do. You are trying to do it all in one span of time. … You have to work with one thing at a time and that was the missing piece to the puzzle. … You kept pouring in more improvements. People kept coming in with more inputs and that was before you had one output.

The work group and the implementation team were keenly aware that the things they wanted done by support personnel were not getting done and complained about it often, as evidenced in the following statements from union representatives, implementation team members, and work group members:

If we had to have something from maintenance, they said you have to put in a [requisition]. It took a month to put lines on the floor. With this kind of thing that we’re going through, you can’t do that. (Union official).

But getting the people responsible to do the stuff we asked for, we still have stuff stored in our work areas that don’t belong to us. Our space is being cluttered. That’s one of the big things that’s always a stickler. We’d get promised, oh, by tomorrow that stuff will be out of there. A week later it will be sitting there. Who’s responsible? Whose stuff is this? Right now we don’t need the space. So that’s why they had it there. But our department never looked neat because we had just junk in there. (Production worker).

How do you [get something done by maintenance]? Send in a work order. Well, you send in a work order and it disappears. How do you get the priority? Because we are trying to show some speed and show some commitment, but we didn’t know who to plug into. (Implementation team member).

It wasn’t really [anything] overt that we’re not going to do this and not going to do that, but it was just like [pause] you’re pulling an ox cart through a mud pit is basically what you felt like. (Implementation team member).
The model thus far captures three basic relationships: 1) improvement activity leads to increases in process capability; 2) tasks accumulate in a backlog when the rate of idea generation exceeds the rate of task completion, 3) support personnel are required to complete improvement tasks. The field data also show that the rate at which workers generated ideas exceeded the rate at which support personnel completed tasks, so the backlog did indeed grow. Taken together, these relations imply that a key challenge in the implementation of an administrative technology for process improvement, such as lean manufacturing, is found in the manner in which the organization addresses the accumulated backlog of improvement tasks. Next, I turn to a discussion of the behavioral responses from support personnel, production workers, and managers.

*Do A Little and Do A Little More: Responses to a Backlog of Work:*

Support personnel have several options to regulate the *Task Completion Rate.* I represent each option as a balancing feedback loop that works to adjust the *Task Completion Rate* to achieve an implicit goal, such as the size of the backlog or the time lag for acting on worker suggestions. Figure 3 shows the first and simplest option. As the backlog of *Tasks To Do* grows, it exerts *Schedule Pressure on Support Personnel,* such as the engineers and maintenance personnel in the examples described above. In the manufacturing plant, schedule pressures arise from several sources. The members of the work group that have generated the ideas have expectations that the ideas will be addressed. Some members of the implementation team were themselves assigned for a limited duration of time and were striving to make meaningful progress before they had to move on to another assignment. Because this was a pilot project, it had a high degree of visibility, and, as one respondent said, "You know, management has a tendency to get impatient." In response to the *Schedule Pressure,* support personnel might increase their *Support Personnel Time* supporting the work group. *More Support Personnel Time* doing tasks leads to an increase in the *Task Completion Rate,* which in turn works off the backlog of *Tasks to Do,* thus easing the *Schedule Pressure.* The feedback loop is named the *Getting Things Done* loop and labeled Loop B1.
Figure 3: Balancing Feedback to Regulate Task Completion

The loop identifier, B1, indicates a balancing (negative) feedback loop (see Sterman, 2000).

The option in Loop B1 to increase the Task Completion Rate by increasing Support Personnel Time is limited by the total support personnel time available. In the situation of a fixed or limited resource such as an engineer's time, the additional workload created by the accumulating improvement suggestions stretches personnel to their limits. Support personnel have many demands on their time. As one union member describes, "A guy like [our engineer] and those guys, those guys are busy. They've got so many things on their platter that they can't devote the time that you need for that process."

Nevertheless, as one of the implementation team members said, "Any time there was a problem it fell back on the engineers. … Everything came down to the engineers." A similar challenge arose with support personnel from other areas. Support personnel, facing task demands from other areas as well, were frequently unable to respond quickly to the needs of the work group. Another respondent, noting the sluggish responses from several groups, said, "You couldn't get anything fixed. Do you hear what I am saying? I mean, simple things like putting lines on the floor and getting things moved, it just didn't happen." As described by a production manager:
So, [imagine] I'm in a work group. I need to get something done on my machine, and I ask maintenance to go deal with that. And I want it done right then. Right? I am empowered to do that. I am behaving in a way that is consistent with this principle. Maintenance then says, "OK, I have got this work group that has asked just me for some help, but I have a down machine over here. I've got to make a choice as to where I am going to work, because I am a scarce resource."

In the face of this accumulating backlog of *Tasks To Do* and in the absence of any slack time available, the options were limited. When asked what he could do about the many tasks that were piling up, one of the support personnel, a manufacturing engineer said:

Nothing. You have to wait. That is pretty much it. You could have a meeting. Well, ... we had this action item review every week ... and we said, OK, here is the list of stuff that we want to get done. ... Whatever resources involved will report how much time it is going to take. And every week, you know, things hadn't happened. So that was just a follow-up. What we'd do is try to push the pencil on some of the issues, and that's about all you can do.

Pressure to get work done under conditions of time, budgetary, and resource constraints is commonplace in organizational life. Yet, resourceful, dedicated individuals find ways to make do with what is available. As one team member described:

You're given a problem and you're told to solve it. Your training says solve it. So you analyze it. You put your heart and soul into it, and you're going to work your butt off all the while getting hammered for things left and right. By God, I'm going to get this thing and its going to work. So you invest a lot.

So, the support personnel find other ways of getting things done. Some scholars view this process as one of improvisation or *bricolage* - the "use of whatever resources and repertoire one has to perform whatever task one faces." (Weick, 1993a, p. 352) The accumulating work to do challenges the support person, the *bricoleur* in our example, to respond. One option is to accept a lower rate of task completion, but this only leads to mounting pressure as ideas continue to accumulate. Another option is to increase the amount of time spent doing improvement tasks for this work group (as in Loop B1), but we have already seen that the support person (e.g., manufacturing engineer) is fully loaded with work to do. Allocating more support personnel is an option that might be
available to managers but not to this person. So under the prevailing condition of fixed resources, the only remaining options available to the overworked person facing this challenge must take the form of somehow increasing the Productivity of Support Personnel, which implies finding a way to accomplish more tasks in the same amount of time.

It is useful here to unpack the “doing” of improvement tasks by separately recognizing what gets done from how it gets done, a distinction that echoes the pairing of the content and process of change (Pfeffer, 1997). One way for support personnel to do a task is to do most of the work collaboratively with production workers to gain the benefits of their input and to enlist their support in making the indicated changes, engaging the work group as a partner. Another way for support personnel to do the task is to do the work alone and hand over the work product in the form of a mostly completed task. Both of these approaches are in the repertoire of how work gets done in the plant. Doing the work collaboratively is the prescribed process, consistent with the high involvement principles of lean manufacturing. But doing it alone is quicker. As a manufacturing engineer describes:

You'd kind of lay it out as an engineer yourself and come down and talk to the operators and they'd sign it and it's a done deal. [With] more input [from the work group], of course the process is going to be a lot longer. The process is longer to implement than actual layout moves.

The passage highlights the relative effects of the two approaches on the Productivity of Support Personnel: collaborating takes longer. For another example, consider a proposal to install a new piece of equipment that will improve the flow of parts through the manufacturing cell. One important set of tasks relates to selecting, purchasing, installing, and setting up the machine and specifying the actual procedure for using it to conduct the necessary machining operations. An engineer might accomplish these tasks without any worker involvement, a tactic that would more quickly result in accomplishing the specific physical tasks than a more drawn out approach that relies on active worker involvement in many of the decisions that need to be made. Consider the apparently simple question
as to where on the floor to install a piece of equipment, as described by a manufacturing
engineer tasked with procuring and installing the new machine:

You’ve got, in some cases, three to five operators across two or three
shifts and every one of them has got their own idea of how they should do
this. The machine needs to be three inches to the left. No, the machine
needs to be 8 inches forward. No, this machine needs to be turned 90
degrees. No it doesn’t. Ok, and it just goes on and on and on and on and
on and it’s never ending.

By reducing the amount of collaboration and focusing on executing the operational
content of the improvement tasks, the support personnel can increase their productivity.
An apparently subtle change in how the support personnel does the work, not what work
he or she does, influences the rate at which the work gets done. A production manager
described the tendency to shortchange the amount of collaboration:

People love to take shortcuts. [For example, I've seen people say] screw
it. I’m just going to tell the operator what to do and I’m not going to go
find the steward so that we can do it together. ... I want this operator to
do something, right? Rather than talk about the idea with the steward and
then jointly go and tell the operator that this is what we need to do, I’ll just
go and tell them. ... Rather than me go to the steward and say ... you
know, we’ve got this problem, what if we were to do this and what if [we
were] to ask [the production worker] to go and do this, this, this and this?
What do you think? You know -- we just say no, just screw it.
[Production worker] just go do this. And [the production worker] goes,
"You know, I’m not sure I want to do this." "I want to go talk to the
steward." And the steward says, "Well shit, I don’t know anything about
it." Well that implies that, you know, don’t do it, right? "Wait until I go
and talk to the work group advisor about this." And then the steward goes
and sees the work group advisor and says, "What the hell are you doing
telling [the production worker] to go and do this for?" "You know, I mean
that doesn’t make any sense. I mean [jeez], if you had talked to me -- now
you haven’t. Now he’s all pissed off and gee, I can’t support this." Right?

Working in a less collaborative manner is one type of shortcut. Other researchers have
documented the use of shortcuts in response to time pressure. For example in a
commercial bank, lending officers facing work pressure from a backlog of orders cut
corners to reduce the time they spend on each customer order (Oliva & Sterman, 2001).
Another shortcut is to do the easiest work first, as an engineer describes:
You keep adding to the action register list and you keep completing the first item. ... We didn't put an actual priority that I can remember on any specific issue. The easiest thing always got completed first. The hardest thing got completed last. It didn't matter if the hardest thing was the first thing coming along. It was the easiest thing that got completed first.

The use of shortcuts to increase productivity creates another balancing feedback loop, as shown in Figure 4. An increase in Tasks to Do causes an increase in Schedule Pressure on Support Personnel. In response, the support personnel conduct their work in a less collaborative manner, spending more of their time working alone and consequently less time working collaboratively. The character of how the support personnel do the improvement work changes, captured in Figure 4 as a reduction in the Degree of Collaboration. Because the less collaborative way of working is quicker, the lower Degree of Collaboration yields a higher Productivity of Support Personnel. The resulting increase in the Productivity of Support Personnel raises the Task Completion Rate, which draws down some of the backlog of Tasks To Do and thus relieves some of the Schedule Pressure. The balancing loop thus formed is named the Go It Alone loop and designated Loop B2 in Figure 4. The accumulating backlog of Tasks To Do is a different set of specific tasks for the various support personnel in the plant, but the basic dynamics are similar. The accumulating backlog yields pressures that increase the likelihood of executing tasks through less collaborative or participative means to complete them in a timely (or less tardy) manner.
The essential characteristic of these shortcuts is they accomplish similar content in the completion of the instrumental task at hand, but they do so with a different process. They contribute to the Task Completion Rate but they are not based on the involvement of the work group members. Equipment gets purchased, but workers may not have been involved in their selection. Layouts get designed, but the work group does not participate in the active discussions they would have with a high involvement approach. The instrumental character of the task is completed, but the social elements are given lower priority. The responsible individual is able to claim the task is completed. As a production manager said, "[You can tell your boss that you] got the machine." However, the benefits of engaging the workforce in the process have been circumvented. The production manager continues, "You believe that you'll be able to take the shortcut, but in the long run the shortcut is going to set you back." Indeed, it is the long run consequences of these shortcuts that are critical to the pattern of start and fizzle.
Do It Alone, Learn A Little Less: Consequences of the Shortcuts

Reducing the Degree of Collaboration leads quickly to a reduction in Schedule Pressure, but circumventing the involvement of the workforce has delayed costs. Two such costs arise from losing experience in the practice of collaborating and experience with the updated production process.

Collaborative work between support personnel and production workers is a source of experience in the practice of collaborating. As they accumulate experience doing work collaboratively, they build a resource available to them to facilitate future collaborative work, a form of social capital (Coleman, 1988). Collaborative work enables the development of trust, enhances social relations that can facilitate the flow of information, and may foster the formation of norms for how collaborative work is done (Coleman, 1988). Collaborating also facilitates the emergence of transactive memory systems (Wegner, 1986). Previous research has shown that team members who train together rather than separately trust each other's expertise more, coordinate task activity better, and develop more accurate perceptions of who knows what, with consequent improvements in performance (Liang, Moreland, & Argote, 1995; Mooreland, Argote, & Krishnan, 1996; Mooreland, Argote, & Krishnan, 1998). Performance may also be improved as individuals develop more skill with collaborative work in general, for example through learning to adjust their behavior to complement other team members (Reagans & Argote, 2002).

Figure 5 adds a new construct to the model: Collaborating Experience. The link from Degree of Collaboration to Collaborating Experience captures the relationship between doing work collaboratively and building experience. As the support personnel do work collaboratively, they learn by doing collaboration, and they build skill and social ties. The rectangle around Collaborating Experience represents the accumulation of this experience as the support personnel work collaboratively with production workers. The link from Collaborating Experience to Productivity of Support Personnel captures the effect on performance: the more Collaborating Experience, the greater the Productivity of
Support Personnel. The new links close an important new feedback loop, named the Social Capital loop (Loop R3 in Figure 5). In contrast to the balancing loops described before, the social capital loop is a reinforcing (deviation-amplifying) loop that tends to reinforce and amplify an initial disturbance (Masuch, 1985; Merton, 1948; Weick, 1979). Consider a reduction in the Degree of Collaboration, which might result from the response of the support personnel to Schedule Pressure. As the support personnel decrease their Degree of Collaboration, the level of Collaborating Experience declines, which in turn decreases the Productivity of Support Personnel and lowers the Task Completion Rate. Tasks To Do decrease more slowly, maintaining higher levels of Schedule Pressure on Support Personnel and leading to an even further reduction in the Degree of Collaboration. The cycle continues as support personnel further reduce their collaborative time, undercutting collaborating experience and reducing productivity still further. However, a reinforcing loop can operate in the favorable direction as well, as a virtuous cycle. As support personnel increase their Degree of Collaboration, they build more Collaborating Experience, increasing their Productivity and boosting the Task Completion Rate, which drains the stock of Tasks To Do more quickly, easing the Schedule Pressure on Support Personnel and favoring an even further increase in the Degree of Collaboration.
Figure 5: Consequences of How Improvement Work is Done (Collaborative Time per Task)

Figure 5 shows a second delayed consequence of collaboration. Collaborative work between support personnel and production workers enables the production workers to develop experience with both the particular production process and the improvement activity. With more such experience, the production workers are able to generate better ideas for future improvements. In contrast, when workers are not involved in implementation activities, their understanding of the modified production processes may even decline. For example, consider what happens when a newly acquired production machine experiences problems. A machine operator who has participated in the specification and selection of the new equipment is more likely to suggest effective solutions than an operator who has not been involved. A production manager described how the operator and manufacturing engineer collaborate to solve problems:

If something’s wrong, all right, [the manufacturing engineer will] definitely tap into the operator’s knowledge to help try to understand, to
help him understand what could possibly be going wrong with this piece of equipment.

Yet the ideas from the production workers vary in how much they contribute to improvement. In the following quote, a production manager explains how production workers use a new system to identify improvement ideas. A suggestion with adequate detail is actionable and useful, whereas a simple note, such as "winder down," is not.

Identify what the problem is. I'm very specific. It's not just "winder down." Okay? It's brass tag [number so and so] and it's fault code F9, wire tension fault. Okay? So if I collect that data and I can turn that into usable information, in other words, at the end of the week you know what? My number one downtime issue was fault code F9, low wire tension fault, right, on brass tag [so and so], I know that that's what I need to fix, right? ... "Winder down" doesn't help us. "Winder down, fault code F9 low wire tension" is what we need.

An interconnection between the process of change and the content of change arises because more worker involvement through collaboration (process) builds worker understanding, which in turn influences the improvement ideas (content) they generate. Collaborating with the support personnel in improvement activities provides production workers with an opportunity for learning by doing, leading to a subsequent boost in the quality of ideas for further improvements. As shown in Figure 5, when the support personnel increase the Degree of Collaboration (i.e., support personnel spend more time with production workers), the production workers gain Process Understanding enabling them to generate ideas that contribute more Improvement per Task, leading to a higher rate of Process Improvement. The additional links through Process Understanding do not close any new feedback loops, but the accumulation by the workers of Process Knowledge introduces an important delay that contributes to the dynamics. The effect of a change in the Degree of Collaboration on Improvement per Task will not be felt for some time. Thus, if support personnel decrease their Degree of Collaboration, there will be a gradual reduction in the workers' Process Understanding, a form of organizational forgetting (Argote, 1996). However, the Process Knowledge accumulated based on earlier collaborative work will remain useful and continue fostering good ideas for some time. Eventually, lower Process Understanding will compromise the quality of the
workers' ideas, so *Improvement per Task* will decline, leading to a slower rate of *Process Improvements*. If the rate of *Process Improvement* falls below the rate of *Process Degradation*, then *Process Capability* will also decline.

The notion of learning by doing has a long tradition in research starting from Wright (1936) who documented learning curves that describe the link between cumulative experience and some measure of performance. Although there is wide variation in the rates at which organizations learn from experience (Dutton & Thomas, 1984), learning curves have been found in many industries (Argote, 1996; Pisano, Bohmer, & Edmondson, 2001) and in a range of functions including manufacturing (Argote & Epplle, 1990), services (Baum & Ingram, 1998), surgery (Pisano et al., 2001), product development, and process innovation (Hatch & Mowery, 1998).

For ease of presentation, Figure 5 represents one specific shortcut, reducing the *Degree of Collaboration*, but there are other examples of shortcuts that lead to a short-term increase in the *Task Completion Rate* at the expense of undermining the collaborative involvement of the workers. One example comes from an implementation team member who was getting impatient about a lengthy delay in getting a new piece of equipment (a lathe). Rather than waiting for the assigned manufacturing engineer (a support person) to work through the collaborative process of developing a new work process, designing a machine layout for it, and developing the associated work standards, the implementation team member chose to do these tasks on his own. Because he himself is an engineer with strong background in lean manufacturing, one alternative for executing the design tasks was to take matters into his own hands. He describes how he reacted to the long delay:

> Still no lathe. Still no standard for the lathe. This thing was supposed to be in there in December. We don’t know how long it takes to do this and to do that. I can estimate it, but I can’t estimate it in [our] system [laughter]. ...Then it was the first couple weeks in January. I just got so frustrated. I finally took their standards and I created a layout for them, and I created a work path for them as options.

Note in the following quote how the support personnel follow an approach that shortchanges collaboration, how the result compromises the workers' ability to get
involved, and how nevertheless the approach leads to a perceived outcome of task completion. A production manager describes how a group of support personnel, including himself, took an approach with a low degree of collaboration:

We didn’t do a good job of sharing what we were doing. The vision, the knowledge, the tools, so that they [the work group] could then apply their thinking and have input into this process.

Q. Who’s we?

You know, a few of us. [The work group advisor], the engineer, one operator, me, and [an outside consultant] actually did it. Okay? And, the intern. And so you know, we did it, right? We implemented this new layout and the work group seems to be working reasonably well with it.

Responses of the Work Group
Support personnel manage Tasks To Do through balancing loops that work by adjusting the rate at which tasks flow out of the backlog. Because they are the primary idea generators, workers have an additional option: to adjust the rate at which tasks flow into this accumulation. Workers may become disillusioned when the response to their ideas is slow or nonexistent. Workers may also recognize that support personnel are overworked and perceive they are unable to handle any additional workload. The feedback from the backlog to the rate of Generating Ideas is another balancing loop, the worker response loop B4 shown in Figure 5. When the backlog of Tasks To Do is high, the Effect of Backlog on Idea Generation is to slow down the rate of Generating Ideas. In the extreme, Generating Ideas will reach a standstill. As one production worker said, "It got to the point where people [were] thinking about dropping out because we can’t get things done." A union official noted how the production workers reacted to the slow rate of getting improvement tasks done:

The people get disinterested. ... They see the results, they get somewhat excited about it. But that dissipates over time when you're not making progress. That's just the way it is. They were running into all kinds of obstacles out there, in terms of trying to get things done.
The presence of the worker response loop has a stabilizing effect on the behavior of the system. Moderating the flow of new ideas into the backlog when the support personnel are already experiencing high levels of pressure to get their work done prevents the backlog from growing so large as to overwhelm the support personnel. Without the worker response loop, a sustained rate of Generating Ideas much greater than the Task Completion Rate would cause the continued growth of the backlog and a concomitant continued increase in Schedule Pressure. Degree of Collaboration would drop and stay low, leading to the eventual loss of Collaborating Experience as well as worker Process Understanding, forcing the support personnel into habitual reliance on working without collaboration to implement ideas that became less and less helpful. The organization would be working harder and harder to get less and less done. Despite the stabilizing effect of the worker response loop, it has other consequences. Workers suppress their otherwise good ideas, so there is an opportunity cost associated with the ideas they do not contribute. Worse yet, the unspoken idea is also unrecognized, so actors in the system are unaware of the opportunity cost they are incurring.

Interaction with the Managers
The description so far has emphasized the activities among the work group and the support personnel. As events unfolded, support personnel were in great demand, and their limited availability was an issue. In this section, I turn my attention to the managers (or others who have influence on the support personnel in question) and how they dealt with this issue.

The work group and implementation team were continually frustrated by the delays in getting assistance from support personnel and called this problem to the attention of management. As one production worker describes:

"We always had updates to the leadership group on how things were going. ... They asked what we needed, and our request was we need a list of people that we can go to, maintenance, facilities, that we can call on the phone or whatever and get things done."
Thus, the request from the implementation team to management was to somehow increase the Support Personnel Time, as shown in a new balancing loop, the Hear What They Say loop B5 in Figure 6. The request if for managers to notice an increase in Schedule Pressure on Support Personnel and therefore increase the Managers' Allocation of Resources. The extra Support Personnel Time would increase the Task Completion Rate, thus reducing the backlog of Tasks To Do and easing Schedule Pressure. The allocation of more personnel as Schedule Pressure on Support Personnel goes up and is perceived by the managers balances the original increase in Schedule Pressure.

The production workers and implementation team members were quite aware of the lack of adequate response from support personnel to execute tasks in a timely manner. They frequently raised the issue, but the managers did not alleviate the situation. A union official described how pervasive this issue was:

[The managers] are running 99 miles per hour trying to get this done, but they are not thinking about all the things that need to be in place to get it done. I said you've got to stop and smell the roses. They said we're going to get this done. I said well who is going to do it. Well, we're going to get that done. Well, by who? Right? The answer is kind of always to just blow over that question. I said you know what, this is amazing, why you won't listen to what I'm trying to say to you? But it all comes down, again, to the resources, in my mind, ... but they won't commit to them.

The support personnel needed to support the improvement efforts were not readily available in the plant, but the managers could have freed some of these internal personnel by setting priorities. Alternatively, the managers could have turned to other areas in the company to secure personnel. Although the managers did not do so, one production manager explained that it could have been done:

We've got some resources internally and if we don't have internal resources, we've got capable people at the [other facility] that if we really needed to, we could make a case for re-appointing some people or asking for some support from the [other facility].

Were the managers simply being unresponsive to the requests from the workforce? Or, is there another explanation? To explore this question, consider how the actions of the work group and implementation team may have influenced the managers.
An important consequence of the shortcuts described earlier is that in most cases, these efforts generated signals that improvements were being made and even led to positive business results (e.g., improvements in cost, quality, or delivery). The most salient performance problem is a stock-out of parts when needed by the assembly line. The situation in the work group before the improvement initiative started, as described by the production manager was that:

We were not doing very well in [this manufacturing cell]. It was almost a daily phone call from [the other engine plant about] the fact that we shut them down for some reason or other because of a quality issue or lack of product. So we weren’t feeling really good about what was going on down in [that manufacturing cell].

But, the efforts in the work group had begun to significantly improve the situation. As one union member described:

They were making the right part at the right time. ... You didn’t have all of these emergencies [such that] if we don’t run Saturdays and Sundays, the [assembly] line isn’t going run on Monday. The things in the supermarket were being filled, and they were holding up pretty decent.

The performance improvements were thus quite salient. So, from the perspective of the managers, the improvement initiative was beginning to show tangible signs of progress and some meaningful business results.

These signs of progress are another source of feedback to managers and give rise to additional interconnections shown in Figure 6. An increase in the Task Completion Rate, after some delay for the outcomes of the task to become visible and the managers to notice them, leads to an increase in the Managers Perceived Progress. All else equal, as Managers' Perceived Progress goes up, the managers perceive less of a need for resources, so Managers' Allocation of Resources goes down. These new links from Task Completion Rate through Managers' Perceived Progress to the Managers' Allocation of Resources close another important balancing feedback loop, labeled the See What They Do loop (Loop B6 in Figure 6). To continue tracing around this loop, note that as Managers' Allocation of Resources goes down, Support Personnel Time also falls, leading to a slower Task Completion Rate. The See What They Do loop is a balancing
loop that seeks to adjust the availability of personnel to achieve an implicit rate of task completion desired by the managers. It works to offset an initial increase (decrease) in the *Task Completion Rate* by tightening (loosening) resource constraints, thereby causing a balancing decrease (increase) in the *Task Completion Rate*.

Figure 6: Interactions with Managers and Resource Allocation

The See What They Do loop shows an important unintended consequence of the local efforts of the participants to do what they could under the constraints they faced. Consider an increase in *Schedule Pressure* as *Tasks To Do* increases. With increased *Schedule Pressure*, the support personnel reduce their *Degree of Collaboration*, which accomplishes the immediate objective of increasing the *Task Completion Rate*. But the very success the support personnel have in accomplishing these tasks with the shortcuts they improvise, such as doing things less collaboratively, sends a signal to managers that
the work is indeed getting done. Seeing the work get done, the managers feel less
pressure to allocate more personnel, so the support personnel get locked in to a pattern of
working with less collaboration. The more productivity gained from reducing
collaboration, the stronger the boost in getting tasks done will be. The more effective the
shortcuts are at increasing productivity, the more they undermine the fundamental
solution to the problem they are addressing - excess work to do.

The managers face a difficult challenge to marshal an appropriate level and mix of
personnel availability. Balancing the supply and demand for support personnel is
complicated by many factors, such as the costs of the personnel, the complex array of
skills that may be required, the irregular temporal pattern of demand for them.
Commenting that the cost of personnel should be obvious, a production manager said:

People lose sight of that [idea that personnel aren't free]. They think it's
other people's money which means it's somehow not money.

Indeed, another production manager explained how a perceived personnel shortage was
not only commonplace, but expected:

We're never going to have enough resources to do everything that
everybody wants when they want it done. It just ain't going to happen.
So, it's not a problem that we can solve with resources. Really, it's a
tension that exists in the business. ... We have to manage it by indicating
that people are going to have to make choices, and it's not that they're just
blowing you off. ... We're never going to have enough maintenance
people to be responsive to every idea that every operator has on the shop
floor. Right? It ain't going to happen. That's just part of the reality of
how it works.

The challenge is more than just making more support personnel time available. The
challenge is to make the right person with the right skills available at the right time.
When asked why the managers did not allocate more support personnel, one answer from
a production manager is simply that it is difficult to anticipate the needs:

We [management and front-line workers] are not as good at defining what
resources are needed. ... I don't think that we do a very good job of being
clear about what is the work we need to do ... or when does it need to be
done ... and who are the resources to get it done [and] how we support
that work to get it done when we said it was going to get done. Around
here it seems to be it'll get done when it gets done.
Moreover, in many situations the key person is a "shared resource" with duties that span many areas in the plant. As another production manager explains:

People are going to wrestle with the fact that others need to be responsible to others. ... The people need to be able to accept the fact that right now I can't be responsive to that work group because I have a more urgent need for the business, and I have to be willing to accept that.

But despite some complaints about slow responsiveness and requests for more support personnel, changes were getting implemented, which provided the managers with tangible evidence that allowed the managers to believe that the personnel shortages were not that critical. Over the first several months of the change efforts, managers frequently described positive outcomes and made favorable assessments of progress. One production manager described other improvements he saw:

The actual flow of material through the workgroup. The actual staging of material. I mean, we [used to] run out of stuff because we couldn't find it, not because we didn't have it. And that's been a big improvement. Certain things go in certain places now that used to pretty much just be anywhere before. It's not as good as it needs to be. There still is a lot of work that needs to be done there, but we are not – it looked that somebody had just dumped a dump truck of cardboard boxes in the inbound supermarket, or on the inbound inventory side of the business. Things are a little more orderly than that now.

Figure 6 shows that two balancing loops, B5 and B6, work to influence the Managers' Allocation of Resources. These two loops often will be working in opposite directions. For example, high Schedule Pressure on Support Personnel indicates a need for more personnel, yet a high Task Completion Rate and thus high Perceived Progress indicates less need for personnel. The managers' policies for allocating personnel depend on relative strength of these two influences. There are many reasons to explain why the managers favored the salient and optimistic interpretation described in the See What They Do loop. First, the optimism is plausible. Managers were exposed to signals that may give them a mistaken impression about the progress of the improvement initiative. In light of a support personnel shortage, the people focused on the improvements chose tactics that circumvented the high involvement approaches that were the intent of the
program. These efforts did indeed generate some positive and salient results, even if temporary. Managers were thus given a false sense of the progress of the program. Second, accomplishing tasks is more salient than doing work collaboratively. Research has repeatedly shown that people overweight salient and available information (Taylor & Fiske, 1975; Tversky & Kahneman, 1982), a bias that has been implicated in failures to learn and in the pathology of process improvement (Repenning et al., 2002; Sterman, 1994). Third, reporting practices exacerbated this optimistic interpretation of progress. The work group and implementation team themselves often reported with pride the accomplishments they were making. Indeed, in the very presentations during which they asked the managers to assist with the shortage of support personnel availability, the implementation team reported continued progress to the managers. Several months after the start of the initiative, the work group and implementation team presented a progress report to the plant management. Using 23 PowerPoint slides, seven production workers conducted a presentation complete with before and after pictures, graphs and charts of progress. They reported improvement activity and claimed savings from reductions in floor space, set-up time, overtime and scrap materials and benefits from increases in in-stock percentages of finished goods. The presentation emphasized progress, success, enthusiasm, and positive business results. The group did raise the issue of the need for support personnel - in one bullet point on one of the 23 PowerPoint slides that was discussed for less than 2 minutes out of the 30-minute session. The plant management was so impressed with the progress that they arranged for the work group to make the presentation to three other groups of managers from the company, including the senior executive team of the manufacturing division. By the time the work group presented to these senior executives, the presentation had grown to 42 slides. The slide that noted the need for support personnel had been dropped, and a new slide showed an email to the plant managers and the work group from their main customer in the other production facility:

As you can see from the attached chart significant improvement has occurred over the last month. Keep up the good work!

Moreover, the managers themselves frequently described the successes of the improvement program in presentations to peers, conversations with plant employees, and
discussions with each other. Both the individuals doing the work and the managers put a positive spin on the results.

DISCUSSION AND IMPLICATIONS

In the preceding analysis, I have drawn from a field study of process improvement at a manufacturing plant to elucidate key feedback relationships and induce a model of organizational change. Following Repenning and Sterman (2002), I have posited a feedback structure that calls attention to critical interactions between characteristics of the workplace and the behaviors of the agents acting in the system. I offer a more finely grained analysis that examines the interactions among the work of several groups of personnel on the shop floor and the work of managers. The model explains how people's well-intentioned actions interact to determine the observed pattern of behavior at the organizational level. I find that a high involvement improvement program, such as lean manufacturing, constitutes a feedback structure with the potential to be self-limiting. The generation of ideas through high involvement activities generates demands on key personnel and induces the people involved to find ways to work around the resource constraints they face, but the workarounds undermine the effectiveness of future implementation. I have examined four consequences of these workarounds, the ways people make do with what they have available to do the best they can under the circumstances. First, the efforts do generally yield results - that is, the intent of getting the work done is realized. Second, the approaches they take include ones that circumvent the intended high involvement of the work force, thus precluding the development of resources in the form of human and social capital that benefit future implementation. Third, circumventing the high involvement of the workforce also compromises worker understanding of the new production processes. Fourth, the successes they generate with their ways of working around the support personnel constraints send signals to management that more personnel are not needed and thus counteract their requests for more support personnel.
The analysis draws on data from one longitudinal case study and thus carries the usual limitations for studies of this sort. Questions about limits to the generalizability of the results point to directions for future research. While additional longitudinal studies of the implementation of lean manufacturing would enrich the empirical database from which to induce more general theory, studies of other administrative technologies that rely on high involvement are also called for. Another dimension of generalizability relates to social capital. Future work might investigate the relationship between social capital and positive feedback, paying particular attention both to processes by which social capital is built, analogous to the working collaboratively in the present study, and to other influences of the social capital, analogous to enhanced productivity of collaborating. Despite the limitations, the paper offers another longitudinal study of change to contribute to our empirical data set and responds to a recognized need in the scholarly literature. "The most compelling [challenges] are the requirement to link context with action and the concomitant need to expose processes and mechanisms of change through temporal analysis. … The task becomes to identify patterns in the process of changing" (Pettigrew et al., 2001, p. 700).

My analysis identifies a set of reciprocal interactions between the process and content of change that highlights the mutual causality between action and structure (Barley, 1986; Giddens, 1984; Orlikowski, 1992). My analysis suggests that actors in the production facility drew on knowledge of their prior action, embodied in socially understood norms and rules about who does what, to assign tasks to individuals. In doing so, they reinforced rules for assigning work, constituting and reconstituting a process of changing. They applied deeply embedded rules regarding who does what in a manner that overloaded some agents (support personnel, such as manufacturing engineers) who bore responsibility for implementing many ideas. The ongoing process of accomplishing these tasks evolved, as overburdened actors were increasingly likely to forgo collaboration with other actors who had suggested the tasks in the first place. The recurrent practice of support personnel executing these tasks with a low degree of collaboration reconstituted and reinforced a rule that support personnel conduct such tasks. Moreover, the less collaborative process shortchanged opportunities for both the support personnel and the
work group to learn and cultivate skills in collaborating and implementing ideas. As the change process evolved, so did the content of the modifications it produced, as the nature of the improvement ideas drew on the accumulated knowledge of the workers.

This study has several implications for understanding organizational change. First, the study suggests that consequences arising from personnel shortages during periods of change may be far more insidious than we realize. The study highlights the problems that managers face in allocating support personnel. Managers know that fewer support personnel will mean less work gets accomplished in a given timeframe. However, my analysis identified another consequence of personnel shortages, one that managers are unlikely to account for in their allocation decisions. Under conditions of personnel shortages, industrious employees find ways to make due with what is available and indeed somehow manage the ongoing challenges of organizational life. But the short-term solution to a personnel shortage may have undesirable long-term consequences, especially in the erosion of organizational capability. The ongoing accomplishment of work under these conditions enacts a structure that influences the way work continues to get done. Under conditions of constrained support personnel, organization members learn ways of working and develop social relationships that determine the future course of organizational life.

The notion that building organizational capability requires an investment of resources is not new, but the study here adds two points to suggest that managers may be unwittingly undermining their own futures. First, the self-reinforcing nature of alterations in how work gets done recasts the magnitude of the concern. Initial changes in work practices to meet the challenges of constrained personnel can easily become locked in. Second, because managers are likely to be slow to recognize personnel shortages, the shortage conditions are likely to persist. Work does get done, and changes do get implemented, so managers see tangible evidence that things are not that bad. Negative consequences of personnel shortages are often less salient, less certain, and result only after a significant delay compared to the vivid, certain, and immediate outcomes they notice. Researchers have documented the tendency of people to overweight salient and available information,
to fail to recognize important delays, and to exhibit an aversion to risk (Dawes, 1988; Kahneman, Slovic, & Tversky, 1982; Sterman, 1994). Repenning and Sterman (2002) discuss how these effects yield a preference for work for which the outcomes are salient and are achieved with short delays and greater certainty. Taken together, these additional insights suggest that much of organizational life may take place in the aftermath of personnel shortages.

A second important implication of recognizing the co-evolution of process and content is that imitating another organization's practice is far more difficult than organizational theorists or practitioners often appreciate. An examination of another organization's current practice reveals the nature of their practices in the present, but it reveals only partial information about the trajectory through which the organization has evolved. Popular techniques such as benchmarking and the documenting of best practices focus attention on a static view of current practice, at the risk of missing the importance of the co-evolution that has preceded current practice. Moreover, an implicit assumption is that the imitating organization copies the other organization's practice. But such copying is fundamentally different from copying the co-evolution through which the organization has gone. Such a difference has important implications for framing the challenge. The former suggests the approach is one of implementing or adopting a change. The latter suggests more attention to the importance of making a transition. The transition challenge is apparent in the present study. A production system with highly developed participatory work practices may indeed reduce the need for support personnel, in the short run developing the system requires more support personnel. This study suggests that a richer appreciation of this transition problem is needed, not only among managers of organization change but also among the scholarly community.

The model presented here has several implications for improving how we manage planned change. Under some conditions, the organization begins to implement process improvements, enjoys a period of superior performance, and then subsequently enters a phase in which improvement implementation stalls and performance eventually declines. However, the feedback theory of this model suggests that under different conditions the
organization would begin a similar pattern of improvement activity but would then be able to lock in a sustained level of superior performance. The key difference in the two outcomes is that for sustained change the organization makes a transition to a regime that continues accomplishing a higher rate of process improvement. Because the human and social capital built through the experience of collaborating makes working collaboratively more efficient, the pressures to take shortcuts are small, working collaboratively is favored, and the human and social capital are maintained or built further. The transition is a form of "tipping" in which the system passes a threshold beyond which the new behavior becomes entrenched. Tipping behavior is observed in many systems, such as disease epidemics, product development, and in collective behavior (Granovetter, 1978; Repenning, Goncalves, & Black, 2001; Sterman, 2000). The challenge for managers is to orchestrate the conditions that favor tipping toward the desired regime. In this model, tipping results from the interaction of the Going It Alone loop (the balancing loop B2) and the Social Capital loop (the reinforcing loop R3). There is some critical combination of support personnel, decision rules, and social capital such that the reinforcing loop is just able to dominate the balancing loop. The next essay in this dissertation develops a formal mathematical mode to examine the tipping phenomenon more closely.

The suggestion to shift attention away from explicit improvement tasks and toward building the capability for future improvement echoes themes in organizational learning. The sustainable advantage is found not so much in improving the primary work processes but in improving the more general ability of the organization to continually improve itself (Senge, 1990). For example, in an empirical study of performance outcomes associated with TQM, Powell (1995) found that it was not the TQM tools and techniques but the tacit resources - such as an open culture, employee empowerment, and executive commitment - that can produce competitive advantage. Accomplishing tasks that lead to improvements in process capability can be considered a form of first-order change, whereas building experience so as to collaborate more effectively in the future can be considered second-order change (Bartunek & Moch, 1987). The way in which support personnel interact with shop floor workers is governed by organizational routines (Cyert & March, 1963; Nelson & Winter, 1982), and "engaging in organizational routines can be

57
a process of learning" (Feldman, 2000, p. 625). Routines that are learning processes are examples of double-loop learning (Argyris & Schon, 1996). The present work builds on the organizational learning literature in two ways. First, it locates the organization's capability to improve as a component of a critical positive feedback process. Second, it describes one process, in the balancing shortcut loop, that can undermine efforts to build such learning capabilities.

This work has implications for how we study change in organizations, because it focuses our attention on the interactions between managers, shop floor workers, and other groups of employees. The groups, apparently acting in their own best interest, generate outcomes that are in conflict with their apparent objectives. Similarly, this work has implications for practice, since it calls attention to some of the potential pitfalls in a high involvement improvement program. In particular, this work highlights the critical importance of support personnel. Moreover, the feedback structure posited here describes some potential barriers to the organization's ability to sustain momentum for change. The interactions are subtle, masked by the apparent progress that effective shortcuts generates, and embedded in complex feedback processes. Powerful biases favor interpreting the course of events as the best of times, even when the critical resource of social capital is deteriorating, so also enacting the worst of times. Managers and researchers both stand to gain from a shift of attention away from the salient content of improvement activity and toward the way in which work gets done builds the capability for future success.
REFERENCES


Mooreland, R., Argote, L., & Krishnan, R. 1996. Socially Shared Cognition at Work: Transactive Memory and Group


Sustaining Organizational Change
Simulation Analysis of the Tipping Point in the Dynamics of Process Improvement

J. Bradley Morrison

Sloan School of Management
Massachusetts Institute of Technology
Room E53-358
50 Memorial Drive
Cambridge, MA 02142
Morrison@mit.edu
(617) 354-9585

ABSTRACT

Organizations often adopt improvement programs driven by involvement from front-line workers who generate improvement ideas. These ideas create demands for support personnel (e.g., engineers, skilled tradespeople and managers) with limited availability. This paper explores how implicit and explicit policies governing how implementation tasks get done influence dynamic patterns of implementation over time. The study is grounded in participant observation at a manufacturer implementing lean manufacturing improvements. Workers generate improvement ideas more quickly than they get implemented, creating an increasing backlog of tasks for support personnel. Support personnel can accomplish tasks through working collaboratively with workers, or alternatively, they can perform the work on their own. This essay uses a system dynamics model to explore the dynamic consequences of the two means of implementing ideas. The collaborative approach fosters learning among workers that builds understanding of work processes and skill about collaborating effectively, which leads to benefits in subsequent collaborative efforts.

The essay is organized as follows. The first section describes the observed pattern of "start and fizzle" that will serve as a reference mode of the dynamic behavior to understand, presents a dynamic hypothesis in the form of a causal loop diagram, and explains its rationale. The second section develops a formal system dynamics model of the underlying feedback structure. The third section presents simulations and analysis of the model’s dynamic behavior. The final section discusses some implications for theory and practice.
Organizations often adopt improvement programs to build their capabilities or achieve greater fit with their environments. Many methods for building organizational capability rely on improving a firm's core business processes through participatory improvement efforts driven by the active involvement of and contribution from front-line workers who generate improvement ideas, implement, and learn. Implementing these ideas often requires support personnel, people with specialized skills or in unique organizational roles, whose availability is limited. How, not just what, tasks get done influences the dynamic pattern of implementation over time. Examining how implicit and explicit policies guide improvement activity will likely add insight to our understanding of the challenges of implementation. In this paper, I employ a system dynamics model developed from field study data to explore the self-limiting nature of a participative improvement program and how this arises from the interaction between shop-floor workers and support personnel.

Many improvement programs and management approaches incorporate some notion of employee participation in the improvement activity. Kolodny (1998) lists examples of such programs including "business process reengineering, total quality management and all its derivatives ('kaizen' and continuous improvements, quality circles, just-in-time manufacturing and various forms of 'kanban', customer-supplier chains, quality function deployment, statistical process control and many other control charting methodologies, etc.), employee involvement, empowerment, teams, quality of working life, sociotechnical systems, high commitment work systems, high performance work organizations, organizational architecture, stratified systems, value or supply chains, etc." (p. 2). These organizational technologies differ along some dimensions such as the focal activities or business processes they address and the problem solving approaches and tools they employ; however, they have in common the characteristic that they rely on the active involvement of and contribution from front-line employees to generate improvements.

Despite the much-heralded anecdotal success of some programs, failure rates in such efforts are high. For example, a rich literature studying the success and failure of Total
Quality Management (TQM) programs describes their disappointing track record (Dean & Bowen, 1994; Hackman & Wegeman, 1995; Lawler & Mohrman, 1985; Oliva, Rockart, & Sterman, 1998; Repenning, 2002; Repenning & Sterman, 2001; Sterman, Repenning, & Kofman, 1997). A pattern characterized by an initial burst of apparent success followed by decline is common. While some organizations are able to incorporate new ideas and technologies, others fail with the same ideas and technologies, a phenomenon sometimes called “The Improvement Paradox.” (Keating, Oliva, Repenning, Rockart, & Sterman, 1999; Maier, 2000) Factors other than the characteristics of the ideas and technologies themselves must have an influence on outcomes. Researchers studying the paradox have described several feedback mechanisms that can account for start and fizzle, implicating loss of job security from fear of excess capacity (Repenning, 2000; Sterman et al., 1997), inadequate support (Repenning, 2002; Repenning et al., 2001), lengthy and differential half-lives for process improvement, and attribution errors (Repenning & Sterman, 2000; Repenning & Sterman, 2002). This stream of research generally treats improvement activity at an aggregate level, as an alternative to direct production activity, and develops important insight into the dynamics of process improvement. In the present work, by looking at the practice of improvement activity itself, I examine a micro level feedback mechanism that can generate start and fizzle even in the absence of the factors implicated in the prior work.

In this paper, I develop a causal feedback structure that can endogenously generate the observed start and fizzle pattern of behavior. The feedback structure is grounded in data from fieldwork at a manufacturing firm undertaking a change initiative (Morrison, 2002). I formulate a dynamic hypothesis in the form of a causal loop diagram that postulates that a small set of feedback loops is sufficient to generate the observed pattern of behavior over time. I then construct a mathematical model of the causal feedback relationships. I use this model to test the dynamic hypothesis, showing that this simple causal loop set can generate the observed pattern of behavior. I also explore other possible behavior modes this small loop set might generate.
This paper contributes to our understanding of how organizations change by calling attention to the dynamic consequences of interactions between the content and the process of change. Moreover, the paper provides an operational description of organizational processes that give rise to inertial behavior. The analysis examines how a virtuous cycle of positive feedback can overcome inertial tendencies to return to old ways of working and instead lead to lasting change in organizational capability. The findings have important implications for the practitioner as well. In particular, the results suggest the merits of increasing managerial attention to the "how" of change and decreasing attention on the short-term, instrumental outcomes regarding the "what" is changing.

The paper is organized as follows. The first section describes the observed pattern of "start and fizzle" that will serve as a reference mode of the dynamic behavior to understand, presents a dynamic hypothesis in the form of a causal loop diagram, and explains its rationale. The second section develops a formal system dynamics model of the underlying feedback structure. The third section presents simulations and analysis of the model's dynamic behavior. The final section discusses some implications for theory and practice.

**DYNAMIC HYPOTHESIS**

To develop a dynamic hypothesis, I draw on data from participant observation field work at a manufacturing firm during its implementation of a manufacturing technology known as lean manufacturing or the Toyota Production System (Monden, 1983; Womack, Jones, & Roos, 1990). The firm manufactures automotive products throughout the United States. At one of its several manufacturing facilities, the general manager became interested in taking the firm "to a new level" of manufacturing excellence and engaged the salaried and union management groups to undertake a major initiative aimed at adopting some of the practices of lean manufacturing.

Authorities on lean production and the Toyota Production System assert the importance of the front-line employee as the key contributor to problem solving. After studying
some forty production plants in the United States, Europe, and Japan, Spear and Bowen (1999) attempted to discern the essence of the Toyota Production System:

"Indeed, in watching people doing their jobs and in helping to design production processes, we learned that the system actually stimulates workers and managers to engage in the kind of experimentation that is widely recognized as the cornerstone of a learning organization. That is what distinguishes Toyota from all other companies we studied." (Spear & Bowen, 1999, p. 98)

This ongoing experimentation leads both to process improvements and to increasing the skill of the work force. "Indeed, the term 'lean production' ... captures the minimization of buffers but neglects the expansion of work force skill and conceptual knowledge required for problem solving under this approach" (MacDuffie, 1995, p. 198, Italics in original).

The union and salaried management selected one "natural work group," production workers who work together to manufacture a small set of component parts, to be the first area to work on lean manufacturing. After several months of work, the management and shop personnel were proudly pointing to the initial success of the effort. Yet several months later, some new work practices had been abandoned and performance had deteriorated. Specific performance data for the production unit are confidential but follow a similar pattern. Costs of scrap material, often due to quality problems in production, at first declined but then rose again later. Production output measures compared to standards for the work unit increased for a period of several months after the improvement effort started, peaked, and then began falling toward earlier levels. Respondents testified in the early months to the enthusiasm and participation of the workforce, but several months after the successful start, even intangible measures such as worker enthusiasm were declining. One informant described the situation as "the wheels are coming off." This phenomenon recalls a pattern frequently observed in improvement programs and the implementation of new technology, a pattern I describe as start and fizzle, and the one that will serve as the reference mode for this paper. I next turn to a brief description of the improvement activity as a basis for presenting a causal feedback structure that accounts for the start and fizzle. A more complete description of the field data for this analysis is found in the previous chapter and elsewhere (Morrison, 2002).
The firm recruited some volunteers from the work group and assigned some other personnel to form an implementation team to guide the improvement activity. Managers intended the effort to draw on active involvement of the front-line workers. The General Manager described a basic principle of the approach, saying, "People learn by doing." Management authorized overtime expenditure and conducted weekend sessions to train the work group in basic principles of lean production, such as the central idea to continuously identify and eliminate waste of all kinds (Monden, 1983). The training invited the workers to suggest improvement ideas consistent with the principals of lean manufacturing. Early efforts focused on cleaning up the work space, better organizing the areas for raw materials and finished goods, and reconfiguring the machine layout to streamline the work process and reduce work in process. My field observations show that workers generated improvement ideas more quickly than they got implemented, which created a backlog of work tasks. Most of the tasks were jobs not for the production workers themselves but for support personnel such as plant engineers, skilled tradespeople, and managers. For example, skilled tradespeople were responsible for painting lines on the floor and making signs for labeling sections in the newly organized finished goods area. Plant engineers were responsible for purchasing new storage racks, specifying new production equipment, and managing the acquisition and installation process. This paper considers two stylized means support personnel use to accomplish the various tasks. One means is through collaborative work engaging the front-line workers in work along with the support personnel. The second means is for the support personnel to perform the work on their own. For example, to purchase a new piece of equipment, a plant engineer might choose to solicit input and select the equipment jointly with the workers who will operate the equipment (the collaborative means) or he may choose to investigate options and make a final choice entirely on his own (the work alone means).

There are dynamic consequences of tradeoffs between these two means to implementing improvement ideas. The work alone means will generally be more efficient in the short-run if the only objective is to accomplish a task. The collaborative approach fosters learning among the workers that leads to more skill and understanding of the work.
processes they are trying to improve. Moreover, the collaborative approach fosters learning about how to collaborate more effectively, which leads to subsequent benefits of future collaborative efforts.

Based on analysis of the field study data following traditional qualitative methods (Glaser & Strauss, 1967), coupled with a feedback lens (Repenning et al., 2002), I developed a causal loop diagram that posits a feedback structure arising from the nature of this firm's approach to participative improvement. Table 1 displays excerpts from the field data to highlight the key premises from which I developed the causal structure. I summarize here these key premises:

- Improvement ideas generated by workers result in task work assigned to other people (the support personnel), such as maintenance workers, engineers, and other salaried employees.
- Support personnel accomplish tasks at a more productive rate when they are working alone than when they are working collaboratively with workers.
- Faced with an increasing backlog of work, the support personnel will modify their work practices in the direction of less collaboration.
- Working collaboratively leads to experience collaborating which increases the productivity of the time spent collaborating.
- Workers build knowledge of the production process through involvement in improvement activity, which fosters better idea generation in the future.
### Table 1 Key Premises in the Causal Structure of Improvement Activity

<table>
<thead>
<tr>
<th>Model Premise</th>
<th>Representative Quote from Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task work is assigned to Support Personnel.</td>
<td>&quot;Very rarely is it the person who came up with the idea [who has to do it]&quot; (Production manager)</td>
</tr>
<tr>
<td>Support Personnel productivity is higher when working alone than when working collaboratively.</td>
<td>&quot;[With] more input [from the work group] of course the process is going to be a lot longer.&quot; (Plant engineer)</td>
</tr>
<tr>
<td>Support Personnel shortchange collaboration in response to mounting backlogs.</td>
<td>&quot;Workers will always be able to generate ideas faster than we can come up with them.&quot; (Production manager)</td>
</tr>
<tr>
<td></td>
<td>&quot;We [support resources] didn't do a good job of sharing ... so the work group could have input into the process.&quot; (Production Manager)</td>
</tr>
<tr>
<td>Working collaboratively builds experience that increases productivity.</td>
<td>&quot;I think my people skills have gotten tons better. ... Learning how to address [people's] concerns so they are reduced.&quot; (Plant engineer)</td>
</tr>
<tr>
<td>Worker knowledge fosters better ideas.</td>
<td>&quot;[The engineer] will definitely tap into the worker's knowledge to help him understand what [is] wrong with this piece of equipment.&quot; (Production manager)</td>
</tr>
</tbody>
</table>

Figure 1 shows a causal loop diagram that summarizes the key feedback relationships developed from these premises. The basic stock and flow structure shows the way that improvement activity contributes to process capability. Workers contribute by Generating Ideas, which accumulate in a stock of Tasks To Do for the support personnel. As the work proceeds at the pace of the Task Completion Rate, Process Improvements increase the Process Capability. The change in Process Capability is the net of Process Improvements and Process Degradation, a natural consequence of ongoing use and changes such as customer requirements, part design, and tooling. As Tasks To Do accumulate, the increasing backlog reduces Schedule Slack. The support personnel respond to reduced Schedule Slack by shifting to more time spent working alone and less time Working Collaboratively, thus increasing the Productivity of Support Personnel and the Task Completion Rate, reducing the backlog of Tasks To Do and regaining some
Schedule Slack. This cycle forms a balancing loop, B1, governing the outflow of Tasks To Do, labeled the Work Accomplishment Loop in Figure 1. The workers note the rate at which the support personnel respond to ideas. When the backlog is high, the Effect of Backlog on Idea Generation is that workers reduce their rate of Generating Ideas. When the backlog is low, workers will increase Generating Ideas. This worker behavior forms a balancing loop, B2, moderating the inflow of new ideas, labeled the Idea Generation Loop in Figure 1.

![Figure 1: Causal Loop Diagram of the Basic Feedback Structure](image)

As support personnel spend more time Working Collaboratively, they build Collaborating Experience, which increases the Productivity of Support Personnel. All else equal, higher productivity generates a higher Task Completion Rate, draining more Tasks To Do from the backlog, thus creating Schedule Slack and encouraging the support personnel to spend even more time Working Collaboratively. The cycle forms a reinforcing loop, R3, labeled as the Collaborating Experience Loop in Figure 1. The loop can work as a virtuous cycle as just described or as a vicious cycle: if time spent Working Collaboratively decreases for some reason, the loop reinforces the initial decrease to lead to even further decreases in time spent Working Collaboratively.
The diagram also reflects another consequence of Working Collaboratively. As workers are involved in the improvement activity through Working Collaboratively, they develop more Process Understanding upon which to base new ideas from improvement. As these ideas get better based on richer Process Understanding, the Improvement per Task increases, contributing to more Process Improvements and thus greater increases in Process Capability.

The dynamic hypothesis states that this feedback structure can generate the observed pattern of start and fizzle. The next sections specify a dynamic model based on the causal structure and conduct various tests to explore behavior the basic feedback structure can generate.

THE MODEL

The model represents a simplified generic manufacturing firm in which process improvement activities take place as a means of maintaining and increasing the overall capability of the organization to achieve productive output (widgets) valued by the market.

Process Capability

The overall productive output of the firm increases with the process capability it achieves through the design and execution of its fundamental business processes, such as manufacturing. It is assumed that higher levels of process capability are desirable.

\[
PC_t = \int_0^t (PI_t - PD_t) dt + PC_0
\]

\[
PI_t = TCR_t \times AIQ_t
\]

\[
PD_t = PC_t / TPD
\]

where
\[ PC = \text{process capability (widgets/week)}, \]
\[ PI = \text{process improvements (widgets/week/week)}, \]
\[ PD = \text{process degradation (widgets/week/week)}, \]
\[ TCR = \text{task completion rate (tasks/week)}, \]
\[ AJQ = \text{average idea quality (widgets/week/task)}, \]
\[ TPD = \text{time for process to degrade (weeks)}. \]

Process capability \( PC \) is the accumulation of process improvements \( PI \) less process degradation \( PD \). Process improvements \( PI \) are determined by the task completion rate \( TCR \) and the average idea quality \( AJQ \), representing the amount of improvement on average that is achieved in completing a task. The simplest formulation for process degradation \( PD \) is to assume that the process degrades at a constant fractional rate according to the time for process to degrade \( TPD \).

**Task Accomplishment**

\[ TCR_t = \min (MCR_t, PCR_t), \]
\[ MCR_t = \frac{TTD_t}{MTC}, \]
\[ TTD_t = \int (IG_t - TCR_t) dt + TTD_0 \]

where
\[ MCR = \text{maximum completion rate (tasks/week)}, \]
\[ PCR = \text{potential completion rate (tasks/week)}, \]
\[ TTD = \text{tasks to do (tasks)}, \]
\[ MTC = \text{minimum time to complete tasks} \]
\[ IG = \text{idea generation (tasks/week)}. \]

The ideas generated by improvement effort accumulate in a backlog of tasks to do \( TTD \) until the support personnel, such as engineers and maintenance personnel, select and accomplish the tasks - that is, until they implement the ideas. The task completion rate \( TCR \) is determined by the lesser of the potential completion rate \( PCR \) based on the productivity and the amount of time spent on improvements and the maximum completion rate \( MCR \) based on the number of tasks in the backlog and the minimum time to complete tasks \( MTC \). The minimum time to complete tasks \( MTC \) represents the
minimum time required to select and complete a task that implements an idea that has already been generated.

\[ IG_t = NIG \times EBLIG_t \]
\[ EBLIG_t = f(W_t) \quad f' \leq 0 \]
\[ W_t = ACR_t / PCR_t \]
\[ ACR_t = TTD_t / ACT \]

where
- \( NIG \) = normal idea generation (tasks/week),
- \( EBLIG \) = effect of backlog on idea generation
- \( W \) = perceived workload of engineering and maintenance (dimensionless),
- \( ACR \) = appropriate completion rate (tasks/week),
- \( ACT \) = appropriate completion time (weeks).

Idea generation \( IG \) is from workers generating ideas at a rate that depends on how quickly the ideas they generate are acted upon. When the average time for an idea to get implemented equals the appropriate completion time \( ACR \), the workers generate ideas at the normal idea generation rate \( NIG \). The backlog influences the idea generation rate through the effect of backlog on idea generation \( EBIG \), a downward sloping non-linear function of the perceived workload \( W \) for engineering and maintenance. The function is shown in Figure 2. The perceived workload \( W \) for engineering and maintenance is the ratio of the appropriate completion rate \( ACR \) to the potential completion rate \( PCR \).

When the potential task completion rate \( PCR \) is high relative to the appropriate completion rate \( ACR \), the rate of idea generation \( IG \) increases. Conversely, when the appropriate completion rate \( ACR \) is high, such as when the backlog is large, relative to the potential completion rate \( PCR \), the rate of idea generation \( IG \) decreases. The appropriate completion rate \( ACR \) represents the rate of task completion that is required based on the current backlog such that the average time to complete a task is equal to the appropriate completion time \( ACR \) and is thus determined as the quotient of tasks to do \( TTD \) and the appropriate completion time \( ACR \). The appropriate completion time \( ACT \) represents the workers' expectations about a reasonable delay between generating an idea and acting on it. The worker response to the backlog forms the Idea Generation loop, a
negative feedback loop that works to control the inflow of new ideas to keep the backlog at manageable levels.

Figure 2: Table for Effect of Backlog on Idea Generation

\[ ACR/PCR \]

\[
PCR_t = PTDA_t + PTDC_t
\]

\[
PTDA_t = TSDTA_t \times PWA
\]

\[
PTDC_t = TSDTC_t \times PWC_t
\]

where

- \( PDTA \) = potential tasks done alone (tasks/week),
- \( PTDC \) = potential tasks done collaboratively (tasks/week),
- \( TSDTA \) = time spent doing tasks alone (hours/week),
- \( PWA \) = productivity of working alone (tasks/hour),
- \( TSDTC \) = time spent doing tasks collaboratively (hours/week),
- \( PWC \) = productivity of working collaboratively (tasks/hour).

The potential completion rate \( PCR \) represents the rate at which tasks could be accomplished given the time available from the support personnel, the way they use their time to work on improvement activities (i.e., the split between working alone and working collaboratively), and the productivity of their solo and collaborative efforts. The actual completion rate \( TCR \) will equal the potential completion rate as long as there are sufficient ideas available in the backlog. The potential completion rate \( PCR \) is the sum of the potential tasks done alone \( PTDA \) and the potential tasks done collaboratively \( PTDC \). The potential tasks done alone \( PTDA \) represents the rate at which support personnel accomplish tasks for that portion of their time spent doing tasks alone \( TSDTA \), assuming
adequate work to do is available, and is determined by that amount of time and the productivity of working PWA. Similarly, the potential tasks done collaboratively PTDC represents the rate at which support personnel along with workers accomplish tasks for that portion of the support personnel time spent doing tasks collaboratively TSDTC, assuming adequate work to do is available, and is determined by that amount of time and the productivity of working collaboratively PWC.

**Support Personnel Time Spent Doing Tasks**

The model assumes the support personnel have a given amount of time available to work on improvement activities and choose to allocate (all of) that time between two ways of working: working alone and working collaboratively.

\[
TSDTA_t = EMTA_t - TSDTC_t \\
TSDTC_t = EMTA_t \times FTC_t \\
FTC_t = \int_0^t \left( \frac{(IFTC_i - FTC_i)}{TAF} \right) \, dt + FTC_0 \\
EMTA_t = EMTA_0 + RG
\]

where
- **EMTA** = engineering and maintenance time available (hours/week).
- **FTC** = fraction of time collaborating with workers (fraction),
- **IFTC** = indicated fraction of time collaborating with workers (fraction),
- **TAF** = time to adjust fraction (weeks),
- **RG** = Resource Gift, a test input used to add resources (hours/week).

Engineering and maintenance time available **EMTA** represents the total amount of time each week that support personnel have available to work on improvement activities. By assumption, these resources apply all of this time to work on improvement tasks either by working alone or by working collaboratively with workers. The time spent doing tasks alone **TSDTA** is the total time available **EMTA** less the time spent doing tasks collaboratively **TSDTC**. The time spent doing tasks collaboratively **TSDTC** is determined as a fraction **FTC** of total time available. The fraction of time collaborating **FTC** with workers adjusts with a delay to the indicated fraction of time collaborating with workers.
IFTA. The time to adjust fraction TAF is the average lag and represents the time it takes for support personnel to perceive a need to change their mix of time allocation and to adjust to the desired mix through actions such as starting to work on different tasks and changing their methods for accomplishing work already underway.

\[
IFTA_t = EB (PFTC_t) \\
EB (PFTC_t) = f (PFTC_t) \quad f' \geq 0 \\
PFTC_t = FTC_t \times ESS (SS_t) \\
ESS (SS_t) = f (SS_t) \quad f' \geq 0 \\
SS_t = PCR_t / ICR_t \\
ICR_t = TTD_t / DTC
\]

where

\( EB (PFTC) = \) effect of bounds on pressured fraction (fraction),
\( PFTC = \) pressured fraction of time collaborating (fraction),
\( ESS (SS) = \) effect of schedule slack on fraction of time spent collaborating (dimensionless),
\( SS = \) schedule slack (dimensionless),
\( ICR = \) indicated completion rate (tasks/week),
\( DTC = \) desired time to complete tasks (weeks).

The effect of bounds on pressured fraction \( EB \) is a function of the pressured fraction that ensures that the indicated fraction is bounded between 0 and 1. The pressured fraction of time collaborating \( PFTC \) is determined by the current fraction of time collaborating with workers \( FTC \) multiplied by an effect of schedule slack on time spent collaborating \( ESS \). This representation describes a process in which the support personnel adjust from their current mix of work in response to some pressure based on the schedule slack \( SS \) to either decrease or increase their time spent collaborating in order to achieve higher or lower rates of task accomplishment. They continue to adjust upward or downward until the indicated completion rate \( ICR \) and potential completion rate \( PCR \) come into balance and the pressure is relieved. The effect of schedule slack on time spent collaborating \( ESS \) is an upward sloping function that varies non-linearly with the schedule slack. Schedule slack \( SS \) is the ratio of the potential completion rate \( PCR \) to the indicated completion rate \( ICR \). The indicated completion rate \( ICR \) represents the rate at which support personnel must accomplish tasks in order to complete them on average in the desired time to complete tasks \( DTC \). The desired time to complete tasks \( DTC \) represents the goal of the
support personnel for the average time to complete a task. As shown in Figure 3, when schedule slack is equal to one, corresponding to the situation when the potential completion rate equals the indicated completion rate, the effect of pressure is also 1. The support personnel perceive no pressure to change and will continue to work with the current mix of time spent working alone and working collaboratively. When schedule slack is greater than one, corresponding to the situation when the potential completion rate exceeds the indicated completion rate, the effect of pressure is greater than 1. The support personnel perceive an opportunity to spend more time collaborating and still accomplish the work at the desired rate and will thus shift to more time working collaboratively. When schedule slack is less than one, corresponding to the situation when the potential completion rate is below the indicated completion rate, the effect of pressure is less than 1. The support personnel perceive a pressure to increase the rate at which they accomplish work and will thus shift to less time working collaboratively.

Figure 3: Table for Effect of Schedule Slack

![Graph Lookup - Effect of Schedule Slack](image)

Productivity

Productivity compares the number of tasks done to the time spent doing tasks. The productivity of support personnel working alone doing tasks $PWA$ is an exogenous constant in this model, but the productivity of working collaboratively $PWC$ is a function.
of how much experience the workers and support personnel have collaborating with each other.

\[ PW_C_t = PWA \times RP_t \]
\[ RP_t = \int(RE_t) \quad f' \geq 0 \]
\[ RE_t = EC_t/EC_0 \]
\[ EC_t = \int_0^t (BE_t - LE_t) dt + EC_0 \]
\[ BE_t = TSDTC_t \]
\[ LE_t = EC_t/TLCE \]
\[ EC_0 = BE_0 \times TLCE \]

where
- \[ RP \] = relative productivity (fraction),
- \[ RE \] = relative experience (dimensionless),
- \[ EC \] = collaborative experience (hours),
- \[ EC_0 \] = initial collaborative experience (hours),
- \[ BE \] = building experience (hours/week),
- \[ LE \] = losing experience (hours/week),
- \[ TLCE \] = time to lose collaborative experience.

The productivity of working collaboratively, \( PW_C \), is the product of the productivity of working alone, \( PWA \), and the relative productivity \( RP \), which defines the ratio of the productivity of working collaboratively to the productivity of working alone. The relative productivity \( RP \) is an upward sloping nonlinear function of the experience the workers and support personnel have in working collaboratively, expressed as a dimensionless ratio to the initial level of experience collaborating. The function is shown in Figure 4. Experience collaborating \( EC \) accumulates the building of experience \( BE \) through working collaboratively less the losing of experience \( LE \). Building experience \( BE \) is just the time spent doing tasks collaboratively \( TSDC \). The simplest formulation for losing experience \( LE \) is a simple first-order exponential decay, which describes a process by which the average useful life of the experience in collaborating is expressed by the time to lose collaborative experience \( TLCE \). For simplicity, the time to lose collaborative experience is assumed constant. Thus, experience decays at a constant fractional rate.
Improvement Idea Quality

The rate of process improvement depends on the task completion rate $TCR$ and the average idea quality $AIQ$, which expresses the amount of improvement on average that is achieved in completing a task. In this model, the quality of improvement ideas varies as the workers gain or lose experience in the implementation of the improvements that build process capability.

$$AIQ_t = TIQ_t / TTD_t$$

$$TIQ_t = \int_0^t (NIQ_t - UI_t) dt + TIQ_0$$

$$UI_t = TCR_t * AIQ_t$$

$$NIQ_t = IG_t * IPNT_t$$

$$TIQ_0 = NIPT * TTD_0$$

where

$TIQ =$ total idea quality (widgets/week),
$NIQ =$ new idea quality (widgets/week/week)
$UI =$ using ideas (widgets/week/week)
\( IPNT = \) improvement potential of new tasks (widgets/week/task),
\( NIPT = \) normal improvement per task (widgets/week/task).

The average idea quality \( AIQ \) is a characteristic of the stock of ideas that have accumulated in the backlog of tasks to do \( TTD \). The average idea quality \( AIQ \) is the quotient of total idea quality \( TIQ \) and the stock of tasks to do \( TTD \). The stock for total idea quality \( TIQ \) keeps track of the improvement potential of all the tasks in the backlog. Total idea quality is the amount of improvement in process capability that would accrue if the entire stock of tasks to do were implemented. Total idea quality \( TIQ \) accumulates the new idea quality \( NIQ \) less the outflow from this stock from using ideas \( UI \). The formulation for using ideas \( UI \) assumes that idea quality for each task done is the average idea quality for all tasks in the backlog. Thus, using ideas \( UI \) is a rate of outflow from the stock of total idea quality \( TIQ \) determined as the product of the task completion rate \( TCR \) and the average idea quality \( AIQ \). The new idea quality is the inflow to this stock and is determined as the product of idea generation \( IG \) and the improvement potential of new tasks \( IPNT \).

\[
\begin{align*}
IPNT_i & = NIPT \times ERK_i \\
ERK_i & = f (RK_i) \quad f' \geq 0 \\
RK_i & = WK_i / WK_0 \\
WK_i & = \int_0^i (GK_i - DK_i) dt + WK_0 \\
GK_i & = FCT_i \times TCR_i \\
DK_i & = WK_i / TKD \\
FCT_i & = PTDC_i / (PTDC_i + PTDA_i) \\
WK_0 & = GK_i \times TKD
\end{align*}
\]

Where
\( ERK = \) effect of relative knowledge on improvement per task (dimensionless),
\( RK = \) relative knowledge (dimensionless),
\( WK = \) worker knowledge from doing improvements (tasks),
\( WK_0 = \) initial worker knowledge from doing improvement (tasks),
\( GK_i = \) gaining knowledge (tasks/week)
\( DK_i = \) decaying knowledge (tasks/week)
\( FCT = \) fraction of collaborative tasks (fraction),
\( TKD = \) time for knowledge decay (weeks).
The improvement potential of new tasks $IPNT$ represents the amount of improvement in process capability that will result from the implementation of a newly generated idea. The improvement potential of new tasks $IPNT$ depends on the level of worker knowledge from doing improvements $WK$. Define the normal improvement per task $NIPT$ as the amount of improvement from an average task based on ideas generated when workers have the initial level of knowledge from doing improvements. The improvement potential of new tasks $IPNT$ is then the normal improvement per task $NIPT$ multiplied by an effect of relative knowledge on improvement per task $ERK$. The effect of relative knowledge on improvement per task $ERK$ is an upward sloping nonlinear function of relative knowledge $RK$, which is worker knowledge from doing improvements expressed as a ratio to the initial level of worker knowledge from doing improvements. The function is shown in Figure 5. Worker knowledge from doing improvements $WK$ parallels the accumulation of experience collaborating and represents the accumulation of gaining knowledge $GK$ less decaying knowledge $DK$. Gaining knowledge $GK$ represents worker learning from the recent improvement activity in which the workers have participated collaboratively with the support personnel. Thus, gaining knowledge $GK$ is the product of the fraction of collaborative tasks $FCT$ and the task completion rate $TCR$. The fraction of collaborative tasks $FCT$ is determined as the potential tasks done collaboratively $PTDC$ divided by the total potential tasks done, which is the sum of potential tasks done collaboratively $PTDC$ and potential tasks done alone $PTDA$. Decaying knowledge $DK$, like losing experience, is modeled as a first-order exponential decay, with the average time for knowledge decay $TKD$ assumed constant.
Finally, the model chooses an initial quantity of engineering and maintenance time available that will result in the simulation runs starting in an equilibrium. The initial values of stocks are formulated such that inflows will equal outflows under initial conditions. For the stock of Tasks To Do $TTD$ to be in equilibrium, the task completion rate $TCR$, based on the initial fraction of time collaborating $FTC$ and the engineering and maintenance time available $EMTA$, must be equal to the inflow determined by normal idea generation $NIG$. The following equation for the initial value of engineering and maintenance time $EMTA$ establishes this equilibrium condition for initialization of the model:

$$EMTA_0 = NIG / (PWA \times RP_0 \times FTC_0 + PWA \times (1 - FTC_0))$$

The parameter values assumed for exogenous constants are shown in Table 1. The parameters were chosen to represent plausible values for a manufacturing firm such as the one studied in the field work, based on judgements I made from my observations in the field work and in conversations with managers.
Table 1: Base Case Parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter Name</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPD</td>
<td>Time to Process to Degrade</td>
<td>16</td>
<td>weeks</td>
</tr>
<tr>
<td>MTC</td>
<td>Minimum Time to Complete Tasks</td>
<td>1</td>
<td>weeks</td>
</tr>
<tr>
<td>NIG</td>
<td>Normal Idea Generation</td>
<td>15</td>
<td>tasks/week</td>
</tr>
<tr>
<td>ACT</td>
<td>Appropriate Completion Time</td>
<td>2</td>
<td>weeks</td>
</tr>
<tr>
<td>PWA</td>
<td>Productivity of Working Alone</td>
<td>1</td>
<td>tasks/hour</td>
</tr>
<tr>
<td>TAF</td>
<td>Time to Adjust Fraction</td>
<td>2</td>
<td>Weeks</td>
</tr>
<tr>
<td>FTC₀</td>
<td>Initial Fraction of Time Collaborating</td>
<td>0.3</td>
<td>fraction</td>
</tr>
<tr>
<td>DTC</td>
<td>Desired Time to Complete Tasks</td>
<td>2</td>
<td>weeks</td>
</tr>
<tr>
<td>TLCE</td>
<td>Time to Lose Collaborative Experience</td>
<td>12</td>
<td>weeks</td>
</tr>
<tr>
<td>TGK</td>
<td>Time for Knowledge Decay</td>
<td>12</td>
<td>weeks</td>
</tr>
</tbody>
</table>

**BEHAVIOR OF THE MODEL**

The following sections describe how the model is used to understand the patterns of behavior this feedback structure can generate.

*Good Ideas that Lead to Ruin*

The simplest test to begin understanding the possible behavior modes of this model is to begin a simulation under equilibrium conditions, then introduce a small disturbance, and observe behavior. The equilibrium condition at the outset means that the organization is completing just enough improvement activity to offset the normal degradation in process capability that takes place. Support personnel are working with a mix of working alone and working collaboratively that achieves exactly the rate of task completion that is required to match the rate of idea generation from the workers, so the backlog of tasks to do is constant. A small number of new ideas, equivalent to approximately one-half week's worth of idea generation by the workers, is introduced into the stock of Tasks To Do in week 8. Neither the equilibrium condition of the initial state nor the disturbance
chosen for this test are intended to represent the conditions in the manufacturing setting of interest. Nevertheless, this test bears some similarity to a sudden increase in worker enthusiasm and consequent idea generation that might follow a new training program or a particular weekly work group meeting focused on problem solving.

The simulation results, displayed in Figure 6, show some surprising behavior. The introduction of new ideas for improvement actually causes a decline in performance. These ideas were indeed implemented, and each idea did indeed yield some improvement in process capability. There is a short period of improved performance (process capability) from the introduction of the new ideas at week 8 until approximately week 15, but then the process capability declines below its initial levels and stays lower over an extended period of recovery back to almost the initial level by week 225. During the entire period from week 15 to week 225, process capability is compromised as a result of simply introducing some new ideas for improvement.

**Figure 6: Dynamics of Process Capability following the Introduction of New Ideas**

![Graph showing process capability over time](image)

**Process Capability**

- Pulse ideas: 4 + + + + + + + + + + widgets/Week
- Process Capability: eq 2 2 2 2 2 2 2 2 2 widgets/Week

90
To understand how the feedback structure generates the observed behavior, consider the manner in which the support personnel respond to the increase in Tasks To Do. Figures 7 and 8 examine the early dynamics more closely. Support personnel recognize an increased backlog and thus a reduction in schedule slack. Faced with more work but the same amount of time available to complete the tasks, they respond by reducing the amount of time they spend collaborating in order to spend more time doing tasks alone. This shift in the mix of work practices has three consequences. First as shown in Figure 7, the most immediate consequence is an increase in the task completion rate as the increase in time spent doing tasks alone yields a higher overall productivity for the support personnel time. This increase in task accomplishment is the support personnel's intended rationale as they shift how they spend their time. The increased rate of task completion drains the backlog of Tasks To Do, and this stock falls back to, and even briefly below, its original level. From this view, the support personnel appear to have adjusted their work practices to respond to this temporary disturbance in order to get their job done in the expected time. This early increase in the task completion rate accounts for the observed increase in process capability for the first several weeks. Notice that the task completion rate is above the initial level for the entire period shown.

*Figure 7: Adjusting the Task Completion Rate in Response to a Pulse of New Ideas*

![Graph showing the adjusting the task completion rate](image)

- Schedule Slack: pulseideas
- Fraction of Time Spent Collaborating with Workers: pulseideas
- Task Completion Rate: pulseideas

91
A second consequence of this shift in work practices is considerably more important in contributing to the observed dynamics and is displayed in Figure 8. As the support personnel shift to less time spent working collaboratively and doing more work on their own, the workers are included in less of the implementation work that is getting done. Consequently, the workers' first-hand understanding of the production processes and their enthusiasm to improve it is reduced, resulting in a drop in the quality of the improvement ideas they subsequently generate. The average quality of the ideas, which measures the amount of improvement from each idea implemented, declines and continues to decline as the time spent collaborating declines. The fraction of time spent collaborating reaches its minimum in week 13 just as schedule slack returns to its neutral value of one when the potential completion rate for a short time is equal to the indicated rate of doing tasks. The average quality of ideas continues to decline still further, until it reaches its minimum value in week 28. This minimum is approximately 15 weeks after time spent collaborating reaches its minimum, due to the lag in the gaining and losing of worker knowledge (time constant = 12 weeks) and the time that new ideas wait in the backlog of tasks to do (average residence time during this period ranges between 2 and 3 weeks).

**Figure 8: Influence of the Quality of Ideas in Response to a Pulse of New Ideas**

![Graph showing the effect of quality of ideas on time (week)]
The organization continues to implement improvement ideas, indeed at a higher rate than before, but the quality of these ideas remains below the initial levels. The support personnel have acted to get their work accomplished in the desired time, and they have done so to an admirable degree. However, an unintended consequence of the way in which they have done the improvement activity is to foster a loss in the quality of future ideas implemented because of lower levels of worker familiarity with the work processes. To demonstrate that this effect of the quality of ideas generates the observed dynamics, I conducted a simulation using the same pulse input of new tasks but under conditions in which the average idea quality is held constant (which is done in the simulation by setting the Effect of Knowledge on Improvement per Task = 1). This change knocks out the loop for the Idea Quality, and the results as shown in Figure 9 show that process capability increases first, then decreases, and is always at or above the initial level.

**Figure 9: Influence of the Effect of Idea Quality following a Pulse of New Ideas**

![Graph showing process capability with and without an effect on idea quality.](image)

Further, as the support personnel shift to less time spent working collaboratively, they begin to lose some of their experience in collaboration, which in turn lowers the productivity of collaborating. The initial dynamics that generate this effect are similar to
those for idea quality, as shown in Figure 10. The fraction of time spent collaborating decreases in response to the loss of schedule slack, followed by a loss of experience collaborating. Here again, experience in collaborating remains lower than the initial level for an extended period of time after the support personnel have already begun to increase the amount of time they spend collaborating.

Figure 10: Influence of Experience in Collaborating after a Pulse of New Ideas

The influence of this protracted loss of experience in collaborating here is different from the decline in idea quality, because the experience in collaborating is part of a positive feedback loop that partially governs the task accomplishment rate. When experience collaborating drops, the productivity of collaborating also drops, which means that ceteris paribus the (potential) task accomplishment rate will be lower. With a lower task accomplishment rate, the support personnel will experience less schedule slack, leading to less time collaborating followed by a further decline in experience in collaborating, completing the cycle. In this simulation, the positive loop works in the unfavorable direction, and the influence on behavior is that time spent collaborating stays low for even longer than it otherwise would. To demonstrate this effect, I conducted a simulation.
using the same pulse input of new tasks but under conditions in which the relative productivity of collaborating is held constant. This change knocks out the loop for the Experience in Collaborating. The results shown in Figure 11 compare this simulation to the base case with the Collaborating Experience loop active. The upper panel of Figure 11 shows that the fraction of time spent collaborating recovers much more quickly when the Collaborating Experience loop is not active. The lower panel of Figure 11 shows that as a consequence Process Capability also recovers much more quickly.

Figure 11: Response to a Pulse of New Ideas with and without the Effect of Experience in Collaborating

![Graph showing Fraction of Time Spent Collaborating with Workers](chart)

Fractions of Time Spent Collaborating with Workers: pulseideas, exploopoff

![Graph showing Process Capability](chart)

Process Capability: pulseideas, exploopoff

95
The simulations presented in the previous section are based on an influx of new ideas and no other exogenous inputs. In some ways, it is not surprising that an increase in the amount of work to do with no concomitant increase in the resources to do the work results in undesirable outcomes. The simulations in this section are based on a test in which there are two exogenous inputs: an influx of new ideas to introduce the change initiative (similar in nature to the test in the previous section) and an addition of engineering and maintenance time available for a limited duration. These tests are conducted by introducing in week 8 an additional task inflow equivalent to approximately two week's worth of idea generation by the workers. In addition, the test introduces extra support personnel also at week 8. For ease of interpretation in this test, the quantity of extra resources is chosen to be exactly the number of additional hours of support personnel time that would be needed to accomplish the additional tasks, working at the productivity at the time the tasks are introduced. Under the initial conditions in which the support personnel are working 30% of their time collaboratively, a productivity of collaborating of 0.5 tasks per hour, and a productivity of working alone of 1 task per hour, the initial overall productivity is 0.85 tasks per hour. Thus, an influx of 60 new tasks would require 60 tasks / 0.85 tasks per hour, or 70.6 total hours of extra work time. In the first scenario, these additional resources are introduced along with the new ideas; that is, the total quantity is introduced over the same period (in this case 1 week) as the new ideas.

Figure 12 shows the results of a simulation with a test input of new ideas and the matching quantity of additional support personnel. The dynamic behavior in this scenario is surprisingly similar to the previous test in which no additional resources were introduced: improvement for a short period followed by a drop below the initial process capability that persists for a protracted period of time. Process capability increases sharply at first, because with additional resources the task completion rate is higher as they implement more ideas. For the duration of one week, there are both more new ideas generated and more support personnel time available. Although the quantity of
additional resources is exactly that required to do the extra tasks, the first thing that happens is that schedule slack increases considerably, as shown in Figure 13. Because the expected time to complete the new tasks is 2 weeks, the expected rate of completing tasks increases less than the rate at which the support personnel can get work done with all of the additional time available. Consequently, the support personnel begin by shifting to a more collaborative way of working while the extra resources are present. After one week the additional resources are removed. Process capability suffers a modest decrease but then increases again for a short period, reaching another peak in week 16. This additional improvement activity is due primarily to an increase in the productivity of the support personnel, which is achieved by shifting to spending more time working alone to get the additional tasks done. The additional improvement activity is also bolstered by a brief improvement in the average improvement per task that follows the increased time spent collaborating when the extra resources were in place. Notice that the temporary improved performance here comes mostly from task accomplishment above the normal accomplishment rate because the support personnel work less collaboratively to get more done. Once the extra resources are removed, the remaining support personnel experience a pressure from the loss of schedule slack, and they decrease the time they spend working collaboratively, so they can spend more time working alone and achieve the desired increase in the task accomplishment rate. Notice also that even after the additional resources have left, process improvements continue at a rate above the initial rate, until week 19.

**Figure 12: Response to New Ideas AND Matching Additional Resources**

![Graph showing process capability over time](image-url)
Following this period of improved performance, process capability declines for almost 40 weeks, reaching a minimum in week 55 before it begins recovering over an extended period back to its original level. Although this decline does not start until week 16, the seeds for this degradation have been sown previously, when the support personnel reduced their time spent working collaboratively. During the entire period from week 9 when the extra resources leave until week 16 when the remaining resources have been able to bring the backlog down to a desirable level, schedule slack remains below the neutral value of one, and the amount of time spent collaborating decreases. As a
consequence of the reduced time collaborating, the knowledge the workers gain from doing improvements decreases, which in turn reduces the quality of the new ideas they subsequently generate. The quality of ideas becomes low enough to offset any small gains from a faster task accomplishment rate, so process improvements and then process capability begin to decline. Process capability continues declining until week 55 and then begins a prolonged recovery. As in the simulations shown in the previous section, idea quality stays low long after the time spent collaborating reaches its minimum in week 16, and the recovery is slowed by the additional effect of the loss of experience in collaborating that has taken place through week 26.

Temporary Success: Benefits of Early Collaboration

In this section, I continue exploring test conditions in which I introduce both new ideas and additional resources to work on those ideas. In the manufacturing setting I observed, the management allocated some additional support personnel to assist in the work group where the improvement activity was taking place. After a period of time, these additional people moved on to work on other projects, so the net effect was an increase in support personnel for a fixed duration of time. As in the immediately preceding tests, I introduce at week 8 a burst of 60 new ideas over the period of one week. Also as in the previous section, I introduce at week 8 a matching quantity of additional support personnel; however, in this test I distribute the additional quantity of support personnel time over a period of 10 weeks. The quantity of additional hours required to accomplish the newly introduced tasks is 70.6 hours, which are introduced over a period of ten weeks, implying a rate of 7.06 hours/week. Figure 14 shows the pattern of the exogenous inputs in this test.
Figure 14: Test Inputs to Introduce Ideas and Resources

Figure 15 shows the results of a simulation with these test conditions and, for ease of comparison, also includes the results of the simulation in the previous section. These two simulations are based on identical test conditions, except that in one case (matched) the additional resources are introduced over one week and in the other case (matchedover10) the resources are introduced over 10 weeks. The total quantity of additional person-hours is identical; only the dynamic pattern of adding resources is different. For the same investment in additional person-hours for the improvement effort, these two scenarios show dramatically different results. In both cases, the process capability increases at first, then suffers a modest decrease, and begins to increase again. But, when the additional resources are added and removed quickly, the short period of improved performance is followed by an extended period of performance below the initial level. Conversely, when the same quantity of extra resources is introduced over a longer duration, the resulting pattern of process capability shows a prolonged period of superior performance relative to the initial levels. Both scenarios result in an eventual return to the original level of process capability.
The simulation output in Figures 16 to 21 helps unravel the reasons for the difference in these two test scenarios. Figures 16 and 17 show that when the additional ideas and resources are introduced in week 8, there is an immediate increase in process improvement as a result of the faster task improvement rate that occurs because more resources do more tasks. When the additional resources are distributed over 1 week, the rate of task completion and thus of process improvement increases considerably more than when the resources are introduced over 10 weeks because there are more resources available to do tasks in the first case. In the first case, there is an apparent surplus of resources, which leads to schedule slack, and the support personnel in turn increase the amount of time they spend collaborating, shown in Figures 18 and 19. The workers also increase the rate of idea generation, as they note the responsiveness of the additional resources to their ideas. The benefits of this early increased time collaborating accrue and are seen in an early increase in the average idea quality (Figure 20) and in the experience collaborating (Figure 21), which translates to higher productivity of time spent collaborating. Conversely, when the additional resources are distributed over a 10 week period, the fraction of time spent collaborating begins to fall almost immediately, because the influx of new ideas at the start places task demands on the smaller quantity of resources that will be available over the extended 10-week period. Workers increase the rate of idea generation, but less so than in the first case because there are fewer support
personnel. Average idea quality falls at first, because the support personnel are doing many tasks on their own and workers do not gain a first-hand understanding of these new understanding of these new improvements. Experience collaborating increases a small amount, because even though the mix of time is shifted toward support personnel working alone, with more total resource time the net effect is more time spent collaborating.

**Figure 16: Comparison of Response to New Ideas and Matching Additional Resources in 1 week and over 10 weeks**

![Graph showing process improvements over time.](image)

**Legend:**
- Process improvements: matched
- Process improvements: eq
- Process improvements: matched over 10

widgets/(Week*Week)
Figure 17: Comparison of Response to New Ideas and Matching Additional Resources in 1 week and over 10 weeks

Task Completion Rate

Time (Week)

Figure 18: Comparison of Response to New Ideas and Matching Additional Resources in 1 week and over 10 weeks

Fraction of Time Spent Collaborating

Time (Week)
Figure 19: Comparison of Response to New Ideas and Matching Additional Resources in 1 week and over 10 weeks

Time Spent Doing Tasks Collaboratively

Figure 20: Comparison of Response to New Ideas and Matching Additional Resources in 1 week and over 10 weeks

Average Idea Quality
When the additional resources are withdrawn in week 9 in the first case, there is an immediate drop in the task completion rate. The remaining support personnel continue to face pressure from the backlog of tasks and reduce the amount of time they spend collaborating even further. Time collaborating decreases until week 16 when they have reduced the backlog to a level at which they are able to begin spending more time collaborating. The reduced time collaborating leads to a drop below initial levels in both average idea quality and experience collaborating. In contrast, when the additional resources are available for 10 weeks, the early pressures result in increasing task accomplishment by reducing the fraction of time spent collaborating and in a smaller additional rate of idea generation by the workers. Thus, by week 11 the task accomplishment rate and the backlog have been balanced (schedule slack = 1), and the support personnel begin increasing the time spent collaborating. Time spent collaborating climbs back above its initial levels by week 13 and stays higher for an extended period of time. By the time the additional resources are withdrawn in week 18, the collaborative work has led to accumulation of worker knowledge in the production process and experience in collaborating that pays off in higher productivity. Both average idea quality and experience collaborating grow until they reach their peaks in weeks 36 and 28, respectively, following the reduction in time spent collaborating after
the additional resources are withdrawn. Workers continue to generate better ideas long after the additional resources have departed, and the increased effectiveness of process improvements contributes to higher levels of process capability. Process capability eventually peaks in week 55, as the rate of process degradation grows with increasing process capability and eventually overtakes the slowly declining rate of process improvements.

*Sustained Success: Engaging the Positive Loop*

The results of the previous section are encouraging. With a temporary introduction of additional resources deployed long enough to foster collaboration and the consequent increase in idea quality and the productivity of collaborating, the firm is able to achieve higher levels of process capability for an extended period of time. Yet, these results are still discouraging in that process capability eventually returns to its original levels. The firm enjoys an extended transient effect of its focused improvement activity but does not succeed in transforming to a sustained level of higher capability. In this section, I turn to a scenario that leads to such a transformation.

The previous tests introduce a quantity of new improvement ideas and an allocation of additional support personnel to implement these ideas. The quantity of additional resources was determined to be exactly the quantity needed based on the number of new tasks and the current productivity of the support personnel. In this section, I vary the total quantity of resources introduced over a fixed duration, 10 weeks as before. For ease of interpretation, I vary the quantity of extra resources as a multiple of the indicated quantity, the amount used in the previous tests. Figure 22 shows the behavior over time of process capability after the introduction of new ideas at week 8 and of various quantities of additional support personnel from week 8 to week 18. The runs are labeled to show the quantity of additional resources introduced as percentages of the total indicated quantity. Thus, ad100 introduces exactly the indicated quantity, ad250 introduces 2.5 times this quantity, and so on.
These results show three basic patterns of behavior over time. The first pattern, as in run ad50, is characterized by a short increase in process capability followed by a decline to below the initial levels that persists for an extended period until process capability eventually climbs back to its original level. This pattern results here when the amount of additional resources is far below that required by the newly introduced ideas. This pattern and its underlying causes are substantially similar to the very first test presented in this paper. The second pattern, as in runs ad100, ad150, and ad180, is characterized by an increase in process capability, followed by a brief decline and then another increase to a level of process capability well above the initial level that persists for an extended period until it eventually falls back to its original level. This pattern results here not only when the quantity of resources is apparently adequate, but also when additional resources are 1.5 and 1.8 times the indicated quantity. This pattern and its underlying causes are discussed in the previous section. The third pattern, as in runs ad200 and ad250, begins much like the second pattern with an increase and a temporary interruption in the rate of increase. But, in this pattern, process capability continues to climb to a new level well
above the initial process capability that appears to be sustained in a final equilibrium. In other words, this pattern is one in which the firm has achieved a lasting increase in capability, a sustained success.

To understand the reasons for this sustainable improvement, Figure 23 compares more closely the run ad180 which returns to its initial equilibrium and the run ad200 which displays a permanent change in process capability. Figure 23 shows that in the first case, the fraction of time the support personnel spend collaborating eventually returns to the initial fraction but in the second case, there is a permanent shift to almost all time spent doing work collaboratively. Whereas in the first case, experience collaborating rises and then falls as the time spent collaborating falls, in the second case, experience collaborating rises to a permanently higher level. With a sustained higher level of experience collaborating, the productivity of the time support personnel spend collaborating is also permanently higher, and they are able to sustain a higher overall task accomplishment rate. In addition, the average idea quality, not shown, is also sustained at a permanently higher level in the second case, since worker knowledge from doing improvements is permanently higher because support personnel collaborate with them on such a high proportion of tasks.
Figure 23: Comparison of Results from Introducing Additional Support personnel: 1.8 and 2 times indicated quantity

The key to understanding why these two outcomes differ so dramatically is to recognize the importance of the reinforcing loop through experience in collaborating. As the support personnel spend more time collaborating, they accumulate experience in collaborating, which increases the productivity of the time they spend working collaboratively. At higher levels of productivity, they accomplish more tasks, which increases the amount of schedule slack and allows for a further increase in the time spent collaborating. When this loop is working in the favorable direction, an increase in the time spent collaborating leads to further increases in time spent collaborating. But, the immediate pressure to get work done often outweighs the expected future benefits of collaborating. In those circumstances, cutting back on collaboration leads to further cutbacks as experience collaborating is lost.
To see how this applies to the two cases of interest here, consider the pattern of schedule slack shown in the last panel of Figure 23. In the first case, after the influx of new ideas and subsequent departure of the extra resources, schedule slack never recovers to its neutral value of one, indicating that the task accomplishment rate is equal to the desired rate. While schedule slack is below 1, the support personnel continue to experience pressure to reduce the amount of time they collaborate in order to boost productivity. Consequently, the support personnel slowly adjust downward their time spent collaborating in response to this pressure and eventually bring the task accomplishment rate back to desired levels. They achieve this desired accomplishment rate by squeezing out the less productive collaborative work in favor of doing tasks more productively on their own. The reinforcing loop is working in a downward spiral, as less collaboration leads to lower productivity forcing further reductions in collaboration.

In contrast, in the second case the small amount of additional resources compared to the first case means that while the additional resources are available, the time spent collaborating is a bit higher. This allows for building a bit more experience collaborating and achieving a slightly higher productivity of collaborating. Consequently, when the additional resources are withdrawn, schedule slack does indeed recover and continue to increase above the neutral value of one, indicating that the task accomplishment rate is above the desired rate. While schedule slack is above 1, the support personnel recognize the opportunity to increase the amount of time they collaborate and still achieve the desired rate of task accomplishment. The support personnel collaborate more, leading to increasing productivity of collaborating, a higher rate of task accomplishment and thus even more schedule slack, which enables even further increases in time spent collaborating. More time collaborating also leads to increases in the quality of the new ideas workers generate, so process improvements are higher both because of a higher task accomplishment rate and a higher average idea quality. The reinforcing loop is working in a virtuous cycle, and time spent collaborating continues to grow until it reaches a practical limit close to virtually all of the available time. In practice, other limits not considered here are likely to have an influence before the shift of time gets skewed so far towards collaborative work. Nevertheless, the scenario described here allows the firm to
transition to conditions in which this reinforcing loop is, at least for some time, working in the favorable direction.

DISCUSSION

This paper develops a simple theory that can explain an observed pattern of start and fizzle in the behavior of process improvement. The theory draws from simple premises based on observations from field research in which shop floor workers generate ideas faster than support personnel are able to implement them. The backlog of tasks influences the way in which the support personnel do their work either alone or in collaboration with the shop-floor workers. By shifting to spend more time working alone in the interest of greater productivity, the support personnel work off the backlog, and the implementation of these ideas improves process capability. The increased productivity in the short-term comes at the expense of collaborating with workers, which has negative consequences later. When assigned resources move on to other projects, these consequences become apparent as less effective improvement ideas and lower productivity of future collaboration, which combine to cause a reduction in improvements and thus the observed fizzle. Conversely, if the support personnel build experience in collaborating, they become more productive in their collaborative activity, which can lead them to shift even more of their time to the collaborative approach. This sets in motion a positive feedback loop and leads to a sustainable increase in process capability even after the assigned resources are removed.

The paper presents a causal theory to describe the feedback structure, and formulates a dynamic hypothesis that the structure can generate the observed pattern of start and fizzle. Simulation of a system dynamics model based on the structure supports this hypothesis, showing the start and fizzle behavior, not only when resources are adequate to accomplish the ideas that workers generate but also when even more than adequate resources are allocated. However, when resource availability is adequate to allow the support personnel to cross a threshold and set the positive loop working in the favorable
direction, the firm achieves a lasting transformation to a higher level of process capability.

This work has important implications for managers embarking on a process of organizational change based on participative improvement. First, the explanation as to why improvements are not sustained that this paper offers is simple, but maybe easily overlooked. The reason for the fizzle ultimately traces back to the reduction in the support personnel that are implementing improvement. On the surface, it is not surprising that after resources are reduced, improvement activity wanes. However, the situation is more complicated. Because of important delays in the system, improvement activity continues and process capability increases long after the reduction in resources. The observed decline comes much later than the departure of the additional resources. It is well established that people fail to appreciate the time delays between action and response, and learning is compromised, especially when dealing with systems characterized by non-linear feedback and dynamic complexity (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1994). Managers experiencing a start and fizzle for the reasons described here are unlikely to attribute the decline to a simple explanation based on the departure of resources much earlier.

A second important implication concerns resource allocation policies. A resource allocation policy based simply on the work content of the implementation tasks and monitored by the rate of task accomplishment does not take into account the importance of the consequences of how work gets done. A traditional tool such as an implementation schedule to monitor progress may be misleading or perhaps even exacerbate the schedule pressure that leads to compromising ways of working that have future benefits but short-term costs. Task accomplishment is far more salient than intangibles such as the degree of collaboration in current work practices, the skill of the organization in doing work collaboratively, and the worker's understanding of the business processes they are executing. Yet, sustaining improvement in process capability depends heavily on these less salient factors.
This work also has implications for organizational theory in the area of organizational change. Several scholars have called for a greater attention to the interconnections between the content of change and the process of change. "Enough research has been conducted on organizational change to make it clear that, in most contexts, both content and process factors ought to be evaluated. Yet theories and analyses of organizational change often tend to only one dimension." (Barnett & Carroll, 1995, p.219). "One particularly worthwhile domain in which process and content concerns converge is the translation of knowledge into action. There are many instances in which organizations know what to do (content) but have difficulty in actually implementing that knowledge (action). ... The challenge for organization studies in the future is to find ways of understanding the connection between content and process, between knowledge and action, and between theory and practice" (Pfeffer, 1997, p. 202). Pfeffer suggests that such a line of inquiry will lead to better understanding of "why well-informed individuals and organizations within them pursue ineffective activities and promote dysfunctional policies and practices" (Pfeffer, 1997, p. 202). This paper identifies a critical interconnection between the content of participative improvement change and the process through which such change gets done. Moreover, the use of system dynamics to further explore such interconnections holds great promise.

This paper highlights the importance of "how" work gets done in addition to "what" gets done. The key to achieving a sustained transformation to higher levels of organizational performance is to engage a positive loop that can work to overcome the strong balancing loops that work to favor the status quo. Organizational researchers have frequently noted this strong tendency of organizations to favor the status quo, often referred to as organizational inertia (Leonard-Barton, 1992; Sastry, 1997; Tushman & Romanelli, 1985). The present work provides a specific example of the means by which a reinforcing loop can work to overcome such inertial forces. Indeed, this work suggests a reconceptualization of participative improvement to incorporate a richer appreciation for the importance of building organizational capabilities to improve. The real challenge is not one of improving our organizations' core business processes but of improving the way in which we improve. This notion echoes the thoughts of some researchers in the field of
organizational learning (Levitt & March, 1988; March, 1991; Senge, 1990). Attention to
the importance of the reinforcing nature of improving the ability to improve holds the
potential to improve both management theory and practice.
REFERENCES


The Right Shock to Initiate Change:
A Sensemaking Perspective

J. Bradley Morrison

Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA
Morrison@mit.edu
(617) 354-9585

ABSTRACT

Changes in patterns of organizing often follow disruptions, so-called shocks, in the way that people understand their organization and organizing practices. People engage in a process of sensemaking, and new meanings emerge to guide future practices. Yet, organizations often deal with such shocks in ways that do not lead to any fundamental differences in the observed patterns of organizing practices. Managers and scholars often conclude that such shocks were simply not "big enough" to occasion change. The purpose of this paper is to raise awareness that shocks that are "too big" may also fail as occasions for change. The paper develops a framework that describes how organizational activity following a shock might unfold in a manner that does not occasion cognitive restructuring and organizational change. Four alternative means of disposing of shocks are discussed. Organizations might not notice these shocks, might not take action, might not take novel action, or might not undergo cognitive restructuring. Shocks, which may arise from both external and internal sources, may be either too small or too large relative to the organization’s current cognitive schema. The paper juxtaposes the four means of disposing of shocks against a dimension representing the magnitude of the shocks to form a framework used for discussion of how small shocks and large shocks might occasion each of these four modes of disposition, drawing on relevant literature and providing examples. An important implication for the practice of organizational change is that agents who wish to induce organizational change should benefit from attending to cues from their organizations that indicate whether shocks to occasion sensemaking are too small or too large.
Shocks that interrupt the ongoing flow of activity in organizations often lead to organizational change, yet not always. Managers and scholars often conclude that shocks that do not occasion organizational change were simply not "big enough." The purpose of this paper is to raise awareness that shocks that are "too big" may also fail as occasions for change. The paper employs a sensemaking perspective to develop a framework that describes how organizational activity following a shock might unfold in a manner that does not occasion cognitive restructuring and organizational change.

Scholars and managers have long been concerned with the processes and dynamics associated with changes in the beliefs and practices in organizations. Often a change in patterns of organizing follows some disruption in the way that organizational members understand their organization and organizing practices, demanding that they make new sense of it (Poole, Gioia, & Gray, 1989). At its most basic level, any substantive organizational change is associated with a modification of existing value and meaning systems (Gioia, 1986). The organization engages in a process of sensemaking, and new meanings emerge to guide new practices (Weick, 1995). Yet, organizations often deal with disruptions in ways that do not lead to any fundamental differences in the observed patterns of organizing practices.

Sensemaking provides a useful perspective for understanding organizational change. Sensemaking involves "placing stimuli into frameworks (or schemata) that make sense of the stimuli" (Starbuck & Milliken, 1988, p. 51). The stimuli might also be called shocks (Van de Ven, 1986). As organizations make sense of past actions, people sometimes modify their schemata or commonly held beliefs about how the world works. The updated beliefs guide and give meaning to behavior (Bartunek, 1984; Bartunek & Moch, 1987). Shocks that interrupt an ongoing flow might be occasions for sensemaking, but not all shocks lead to modifications in schema and organizational change. A conventional explanation is that such a shock simply was not sufficiently compelling, or "big enough," to occasion change. In this paper, I show that shocks that are "too big" may also fail as occasions for change. The key challenge will be to better understand the
nature of cues that occasion sensemaking and consequent organizational change and those that do not. In addition to the scholarly merits of an enriched understanding of these shocks, there are useful applications in the field of managerial practice. When an organizational participant, such as a manager, wishes to effect change, the initiation of sensemaking becomes a fundamental challenge.

Bartunek and Moch (1987) use the term second-order change for situations in which there are fundamental alterations of the cognitive schemas, as distinguished from first order-change, which describes the expansion and elaboration of existing schema as people incorporate new information (Bartunek et al., 1987). The discussion that follows focuses on occasions for sensemaking that lead to updating of cognitive schemas, or second-order change, so it will call upon two concepts that deserve a brief introduction here: sensemaking and cognitive schemas. Briefly, sensemaking is an interpretive process in which people assign meanings to ongoing occurrences (Gioia & Chittipeddi, 1991) and involves interpretation in conjunction with action (Thomas, Clark, & Gioia, 1993). Weick (1979) summarizes the theme of sensemaking with a basic recipe for sensemaking: “How can I know what I think until I see what I say?” (p. 133). The sensemaking process both draws upon cognitive schemas as a guide for action and updates these cognitive schemas in making sense of experience.

“A schema is an abridged, generalized, corrigeble organization of experience that serves as an initial frame of reference for action and perception” (Weick, 1979, p. 154). Schemas are dynamic, cognitive knowledge structures used by individuals to encode and represent incoming information efficiently (Markus, 1977). “Schemas” is a commonly used term to refer to cognitive structures, as described by Markus and Zajonc (1985) in their review:

“Cognitive structures are organizations of conceptually related representations of objects, situations, and of sequences of events and actions. What is stored in a cognitive structure can be the specific elements and features defining the object, event, or situation or it can be the rules defining the interrelationships among the elements, or both. Cognitive structures derive from past experiences with many instances of the complex concepts
they represent. Cognitive structures simplify when there is too much, and thus they allow the perceiver to reduce an enormously complex environment to a manageable number of meaningful categories. They fill in where there is too little and allow the perceiver to go beyond the information given. These structures help the perceiver achieve some coherence in the environment and in the most general sense provide for the construction of social reality. They are built up in the course of information processing and they function as interpretive frameworks.”

(Markus & Zajonc, 1985, p. 143)
THE FRAMEWORK

The framework developed here examines some reasons that shocks are ineffective in occasioning organizational change. Shocks may emanate from sources external to the organization or they may be introduced to the organization by people intending to initiate change. Thus, a shock might be a manager's articulation of a business case for change that describes impending demise if people do not undertake an improvement program. One component of the framework suggests four possible failure modes, each of which is a means by which an organization disposes of the shocks in a manner that does not occasion cognitive restructuring. I use the term failure mode to emphasize that the shock "fails" as a precursor to organizational change, not to convey any sense of failure of the organization. The manager's business case (the shock) may not lead to adoption of the improvement program (a failure in occasioning change), although foregoing the improvement program may be an optimal response for the organization. Table 1 lists the four failure modes. Shocks might fail to occasion change because organizational participants do not pay attention to them, because they do not act on them, because the actions they take are not novel actions, or because the actions they take do not occasion any cognitive restructuring. The second component characterizes the degree or magnitude of the shock along an aggregate dimension representing proximity to existing cognitive structures. In combination, the components reflect the notion that each of these four failure modes might occur when shocks are too small or when they are too large. I turn now first to a consideration of the two components of the framework. In the following section, I will examine the eight cells formed by considering the two components together.

The first component concerns the manner in which organizational actors respond to a possible shock. Shocks are interruptions to an ongoing flow of organizational activity. Yet not all interruptions, not all stimuli, not all information, not all events, and not all messages from managers in organizations act as shocks. Several authors have discussed the importance of shocks or interruptions that occasion change (Barley, 1986; Weick,
1995). The analysis here focuses on the several means by which organizational actors are able to dispose of stimuli that do not occasion change. The first component of the framework identifies four such means of disposing of shocks, labeled failure modes.

The first three of these failure modes are evident in an observation by Schroeder, Van de Ven, Scudder and Polley (1989). They observed that successful shocks “stimulated people’s action thresholds to pay attention and initiate novel action.” (Schroeder, Van de Ven, Scudder, & Polley, 1989, p 123). Thus, the shock must be such that 1) the actor pays attention, 2) the actor initiates action, and 3) the action is novel. The first failure mode arises when the organizational actors do not pay attention to the stimuli. It is useful here to recall the distinction between noticing and sensemaking made by several authors (Daft & Weick, 1984; Kiesler & Sproull, 1982; Starbuck et al., 1988). Noticing, which is analogous to Daft and Weick’s scanning and data collection, is an act that classifies stimuli as signals or noise. What people notice becomes an input to their sensemaking (Starbuck et al., 1988). What they do not notice becomes an example of this first failure mode.

The second failure mode arises when the organizational actors notice but take no action. There are at least three somewhat divergent perspectives on action evident in the literature on organization theory (Pfeffer, 1985). One widely held perspective represents organizations as intendedly rational but cognitively limited systems that act in pursuit of goals (March & Simon, 1958). A second view considers organizations as largely under the control of external forces. Population ecology (Hannan & Freeman, 1977) and resource dependence (Pfeffer & Salancik, 1978) theories are prominent in this view. Both the goal-directed and external control perspectives consider action the consequence of some stimulus, the source of which might be in the first case a goal or in the second case an external constraint or dependency. However, in either case, dissatisfaction with existing conditions prompts action directed at improving those conditions. Thus, this failure mode arises when the stimulus from the potential shock attracts the attention of organizational actors, but the stimulus does not stimulate the action threshold (Schroeder et al., 1989).
The third perspective on action characterizes the social constructionist assumptions, which consider action as part of an emergent process and encompasses the sensemaking perspective. Action unfolds over time, and meaning is constructed around these events in an ongoing process. The basic assumption is that organizational members actively create, or enact, their own reality. Enactment is conceptualized as the bracketing and constructing of portions of the flows of experience. "Experience is the consequence of activity," but "there is no such thing as experience until the manager does something" (Weick, 1979, p. 148). The distinction between noticing and acting is less precise here, but there can be noticing without any action (other than the "action" of ignoring, filtering, or not acting). For example,

"Each person watches someone else avoid certain procedures, goals, activities, sentences, and pastimes and concludes that this avoidance is motivated by ‘real’ noxians in the environment. The observer profits from that ‘lesson’ by himself then avoiding those acts and their presumed consequences. As this sequence of events continues to be repeated, managers conclude that they know more and more about something that none of them has actually experienced firsthand." (Weick, 1979, p.151-152).

The third failure mode arises when the organizational actors take action, but the action they take is not novel. Behavior in organizations is based on routines (Cyert & March, 1963; Nelson & Winter, 1982). Routines are based on interpretations of history and emphasize relatively constant responding (Levitt & March, 1988). This failure mode describes the common situation in which the organization experiences a shock that might occasion organizational change, but the tactic for disposing of this shock is the reliance on a habitual routine. Existing norms and routines are adequate to deal with the situation. Actors simply call upon the catalog of causal maps at their disposal, perhaps distorting information slightly to make it fit into existing, and familiar, maps. The habitual response occurs, and the consequence is interpreted as sufficiently close to what might be expected. Action is taken, but the interruption, if any, in the flow of activity and meaning is inconsequential.
A considerable body of literature suggests that the vast preponderance of organizational action is governed by routines (Levitt et al., 1988; Starbuck, 1983). Thus, routine action indeed may be the most common organizational response to a shock. Note again that the use of the term failure mode here is not meant to imply that the routines are causing failure. Although in some situations the over-reliance on routines leads to non-adaptive behavior with dysfunctional consequences (Leonard-Barton, 1992; Starbuck, 1983; Tushman & Romanelli, 1985), routines nevertheless play an important role in simplifying. The term failure mode refers to the “failure” of the shock to induce a restructuring of the cognitive schema, not to a failure of the organization. The response, based on habitual routines, may indeed be an appropriate response.

The fourth failure mode arises when there is novel action, but it does not occasion any cognitive restructuring. The shock occurs, and action occurs. Since this is novel action, it is not interpreted as just a routine response to the stimulus, and thus triggers the search for meaning. But the search for meaning to make sense of this action identifies an explanation found in the current cognitive schema. The difference between this mode and the previous one is subtle but important. Consider these modes from the perspective of an outside observer (perhaps the interested leader of change). The previous failure mode appears as a routine response. There is no discernible difference in action or behavior compared to historical observations. In contrast, in this fourth failure mode, the organization appears to address an interruption with a novel response. However, by making sense of the novel response with an existing explanation, the organization avoids any modification to the cognitive schema. Although invoking an existing explanation may not be observable, the consequence of continued use of the same schema might be. Rather than adopting a new schema, the result here is that the realm of application of the old schema has been expanded, so that it will now have a historical relevance to a wider range of future stimuli. Thus, the observer might notice a short-lived interruption followed by a reversion to old behaviors, even in the face of a repeat occurrence of the novel stimulus.

126
Weick (1995) suggests that two common occasions for sensemaking in organizations are uncertainty and ambiguity. This failure mode might result when neither of these conditions is present. Uncertainty and ambiguity are avoided because an available explanation can be adapted to make sense of the action just experienced. In essence, there is no experience of shock in the cognitive realm. No additional information is necessary, because the explanation is available, so there is no uncertainty. If in addition the available explanation, appropriately modified, is the only one available that fits with the action in question, there will be no confusion among competing explanations and thus no ambiguity.

Notice that it is possible for novel action to be reconciled with existing cognitive schemas. There are two characteristics of this possibility that are central to the sensemaking perspective. First, the existing schemas are invoked in order to associate a belief with an action that has already become part of experience. The sensemaker is an observer of action, taking the action as a given. The process of sensemaking is not so much interested in finding an explanation for why the action occurred as in making sense of the action in order to continue managing a stream of activity. Thus, the sensemaker is not concerned that the novelty per se of the action is not explained. Because sensemaking is retrospective, there is no need to explain how an old schema might generate a new response. Second, novel action is conceptually possible. Cognitive schemas are simplifications, and thus do not anticipate the infinite variety of possible incoming stimuli that make up the stream of activity. Schemas do not deterministically specify outcomes, but rather guide behavior. The emergent course of events may indeed include novel responses, and these may be enacted prior to the pairing of them with any new meanings.

In this failure mode, the possible occasion for sensemaking begins with action. It differs from the previous failure mode, because in the previous failure mode, the observed action, one that was governed by a routine, already has associated with it a belief that makes sense of the action. In the present mode, there is no such belief already attached.
The following excerpt from Weick (1995) explains the nature of sensemaking as a process of relating beliefs and actions to form meanings:

"There seem to be at least four ways in which people impose frames on ongoing flows and link frames with cues in the interest of meaning. Sensemaking can begin with beliefs and take the form of arguing and expecting. Or sensemaking can begin with actions and take the form of committing or manipulating. In all four cases, people make do with whatever beliefs or actions they start with. Sensemaking is an effort to tie beliefs and actions more closely together as when arguments lead to consensus on action, clarified expectations pave the way for confirming actions, committed actions uncover acceptable justification for their occurrence, or bold actions simplify the world and make it clearer what is going on and what it means. In each of these cases, sensemaking involves taking whatever is clearer, whether it be a belief or an action, and linking it with that which is less clear. These are fundamental operations of sensemaking. Two elements, a belief and an action, are related. The activities of relating are the sensemaking process. The outcome of such a process is a unit of meaning, two connected elements." (Weick, 1995, p.135).

Thus, this fourth failure mode might transpire as a novel, committed action uncovers an acceptable justification among the existing cognitive schemas, connecting the old belief with the new response.

The second component of the framework is shown as the horizontal axis in Table 1. The purpose of introducing this dimension is to allow us to unpack some of the reasons that shocks do not occasion change. Whereas the first component of the framework summarizes the nature of the organization’s response, this second component characterizes the shock. This dimension is an analytical simplification that characterizes the shock according to its proximity to existing cognitive structures. When we say in everyday language that a shock is inadequate or insufficient to occasion change, this evokes images of a shock that is too small. Introducing a dimension to describe the magnitude of the shock allows us to also consider that shocks may be too large. This concept then allows us to examine ways that small shocks and large shocks can lead to each of the four failure modes identified above.
Although shocks will vary in both character and magnitude, this simplification aggregates the multiple dimensions of the character of a shock into one unspecified dimension. For the present purposes, our need is simply to distinguish two extremes that describe shocks in relation to the current cognitive schemas. To conceptualize this single dimension, consider a multi-dimensional space with as many dimensions (possibly an infinite number) as necessary to describe the ongoing stream of organizational activity. Imagine that expectations based on the current cognitive schemas and the history of experience are all mapped as points in this complex, multi-dimensional space. Next, describe the shock along these multiple dimensions, and map it in the same space. A straight line from the closest expectation point to the shock point thus represents the size of the gap between the expectations based on current schema and the stimulus from the shock. For our purposes here, it is the size of the gap that matters.

A hypothetical example in a stylized one-dimensional world will help to elucidate the notion of the magnitude of a shock. Consider a manufacturing facility charged with meeting a daily throughput target. After many consecutive periods of meeting the production target, the facility will be operating in an ongoing equilibrium. An increase in the target production level can be considered a shock. The magnitude of the shock can be characterized by the amount of increase in the production target, which is the gap between the new target and the historical output levels. When the shock is quite small, the organization is able to achieve the new target with existing procedures, perhaps by using the available slack to simply work a bit harder. When the shock is moderate, the organization might incorporate some process improvement technique aimed at increasing throughput. When the shock is quite large, the organization might be overwhelmed by the large and continuing production shortfall relative to the new target and thus be incapable of allocating resources to incorporate the process improvements. What constitutes too large a shock depends in part on the amount of organizational slack. In computer simulations of process improvement scenarios similar to these, the improvement program is seen to fail when the shock is either too large or too small (Morrison & Repenning, 2000).
By combining the first component that describes four possible failure modes characterizing organizational responses with the second component capturing two extremes in the magnitude of a shock, I arrive at the framework of Table 1. I turn now to a brief examination of the eight cells formed by this combination.

_shocks that get no attention_

The two cells in the first row capture situations when the shock gets no attention. Consider first the situation in which a small shock fails to get attention. At this end of the shock magnitude dimension are the many possible stimuli that are filtered out without even any noticing. Psychologists use the term “just-noticeable differences” to refer to the smallest differences that people can classify (Starbuck et al., 1988, p 46). Thus, some stimuli are simply too tiny for the human perceptive apparatus to identify. Both individuals and organizations have well-developed processes for separating signals from noise in order to focus attention on those signals that are in some way deemed important. For example, one model of the manner in which executives arrive at their own renditions of a situation is based on a three-stage filtering process that includes limited field of vision, selective perception, and interpretation (Finkelstein & Hambrick, 1996). Both the limits in the field of vision and the selectiveness of the perceptual filter accomplish the screening out of noise so that limited attention resources may be focused on the stimuli that warrant attention.

The stimuli classified into this category are too small for noticing because they are too similar to the existing flow of stimuli to which the organization has become conditioned. What is too similar depends on the sensitivity of the organization’s data collection processes. The perturbations from the continuous flow of activity to which the organization does attend are too minor to be detected by the organization’s scanning apparatus. From the standpoint of the blunt instrument used by the organization, there is no disturbance. For example, when a supplier to an organization changes the price of a part by an increment of a few pennies compared to a several thousand dollar total price, the organization is not likely to notice such a difference. Moreover, the fewer units of this part that are used, the less likely is the organization to notice. Indeed, many of the
existing management reports that generate cost tracking information will not capture such a small difference. Many such small changes are lost as rounding errors. Other such small changes are somehow offset by yet other small changes occurring almost simultaneously so that the resulting net effect falls below the just-noticeable difference threshold.

Shocks can also go unnoticed because they are too large. These shocks are potential stimuli that are so far removed from the current stream of organizational experience that they do not fall within the range of the organization’s scanning systems. They are outside of the organization’s field of vision. “Effective perceptual filtering amplifies relevant information and attenuates irrelevant information, so that the relevant information comes into the perceptual foreground and the irrelevant information recedes into the background” (Starbuck et al., 1988, p. 41). This situation is described well by the colloquial saying that something is “not on my radar screen.” Whereas shocks that are too small to be noticed are filtered out as noise, these shocks that are too large to be noticed are filtered out as not related to any relevant concern.

Without intervention, “structures and systems focus the attention of organizational members to routine, not innovative activities” (Van de Ven, 1986, p 596). The organization’s scanning systems and the individuals’ foci of attention are somehow centered on what is happening in the here and now. Consequently, noticing is a complex function of what is familiar and what is not familiar to the decision-maker (Starbuck et al., 1988). To get noticed, a stimulus must be unfamiliar enough to get detected, but it must also be familiar enough to fall within the range of relevant concerns. Otherwise, the shock is just one of the enormous number of events that are not captured by the organization’s scanning systems.

An example of a large, unnoticed shock is found in the decline of the Swiss watch industry as chronicled by Glasmeier (1991). The Swiss watch industry’s share of the world export market fell from 40% to 10% in the decade beginning in 1974, largely
because the Swiss did not adopt quartz technology for watch movements. Glasmeier offers an explanation for why the Swiss did not notice the emerging quartz technology:

"The watch cartel insulated Swiss manufacturers from the effects of inter-firm competition. Enjoying (volume) control of the world market (based on mechanical devices), it was easy for firms to become myopic about external events and new technology introduced by distant competitors. Because [they] looked only to members of the Swiss Watch Industry Federation for market information, new developments outside Switzerland did not filter into existing information channels." (Glasmeier, 1997, p.33).

These large but unnoticed shocks might be more common than we realize. Van de Ven (1986) identified managing attention as one of the central problems in innovation. Thomas Kuhn (1996) describes the emergence and maintenance of scientific paradigms and argues that the activities of normal science are predominantly conducted within the frame of existing paradigms. These paradigms influence the collection of data, suppressing the perception of data inconsistent with current paradigms. Sterman and Wittenberg (1999) have developed formal dynamic models of Kuhn’s theory in which the self-reinforcing feedback between expectations and perceptions reduces the recognition of anomalies hence constraining the emergence of new paradigms [See also Sterman, 1985 #199].

*Shocks that Get No Action*

Organization participants act on events, attending to some and ignoring most, often talking to people to see what they are doing (Braybrooke, 1964). In the second row of Table 1, I distinguish those shocks that get noticed but do not occasion action by the organizational participants, the “most” that get ignored. First consider shocks that are too small. The notion underlying this failure mode is that there is a threshold below which a stimulus does not engage the organization in a mode of action. Several researchers suggest such a threshold. Van de Ven (1993) highlights the importance of shocks, from either external or internal sources, as triggering mechanisms for concrete actions to undertake specific innovations:

"Concentrated actions to allocate resources and initiate development are triggered by ‘shocks’ … These shocks are
sufficiently large to trigger the attention and action of organizational participants. When people reach a threshold of dissatisfaction with existing conditions, they initiate action to resolve their dissatisfaction.” (Van de Ven, 1993, p. 275).

This view follows the description of problem sensing in crisis situations. “A problem is perceived when a discrepancy or gap is perceived between the existing state (perceived reality, initial state) and a desired state (goal, standard of how things should be, terminal state)” (Billings, Milburn, & Schaalman, 1980, p. 302-303). Indeed, a classic model of organizations set forth by March and Simon (1958) is based on the premise that dissatisfaction with current conditions stimulates the search for improved conditions. A more interpretivist view is found in Dutton and Dukerich (1991). Their view is “that some organizational actions are tied to sets of concerns that we call issues. Issues are events, developments, and trends that an organization’s members collectively recognize as having some consequence to the organization” (Dutton et al., 1991, p 518). The situation reflected in the present failure mode is that the gap between the current and desired conditions is insufficiently large to trigger action. The stimulus is not recognized as having some consequence to the organization.

One reason that gaps may be insufficiently large is that the assessments of both the goal and the current situation change over time in response to evolving conditions. Models of adaptive goals or aspiration levels are found in the literature on psychology (Lewin, Dembo, Festinger, & Sears, 1944) and organizations (Cyert et al., 1963; Levinthal & March, 1981), along with empirical investigations of aspiration level adaptation (Lant, 1992). Judgements about current conditions are also influenced by a tendency to overweight recent and available information, so people may anchor their judgements on recent experience (Tversky & Kahneman, 1974). Consequently, small or gradual changes over time do not reach people’s recognition thresholds for action (Helson, 1964). Suggesting that organizations are a bit like frogs, Senge (1990) recounts the parable of the boiled frog:

“If you place a frog in a pot of boiling water, it will immediately try to scramble out. But if you place the frog in room temperature water, and don’t scare him, he’ll stay put. Now, if the pot sits on a heat source, and if you gradually turn up the
temperature, something very interesting happens. As the temperature rises from 70 to 80 degrees F., the frog will do nothing. As the temperature gradually increases, ... the frog will sit there and boil. ... The frog’s internal apparatus for sensing threats to survival is geared to sudden changes in his environment, not to slow, gradual changes.” (Senge, 1990, p. 22).

Perhaps the most common manifestation of these shocks that are too small to generate action is found when organizations let talk substitute for action (Pfeffer & Sutton, 2000). Faced with a potential shock, organization members often react by finding, perhaps creating, reasons why they should not act. Examples of these situations highlight two main topic areas that often dominate discussions and shift attention away from action. First, organization members articulate reasons that a particular shock does not apply to their circumstances. For example, when managers are presented with information from benchmarking studies of competitors, a common reaction is to focus on differences that might explain why data from the comparison firm should not be compared with the managers’ own firm. Attribution theory predicts that managers will attribute the favorable outcomes achieved by other companies to situational causes in a manifestation of actor-observer effects, and these results have been documented in experimental settings (Wagner & Gooding, 1997). The second topic that diverts attention away from action is barriers to action. Managers engage in discussions that focus their attention on the various reasons they might be unsuccessful or that convince them that they are unable to act. Such responses are perhaps more likely in organizations that Daft and Weick (1984) characterize as passive organizations, which “do not engage in trial and error” (p. 288).

At the other end of the shock dimension, we find shocks that are too large to occasion action. The organization finds itself not responding to these shocks because the collective does not know what to do. This is the organizational equivalent of the proverbial “deer in the headlights” that is frozen at a standstill because the shock of the bright lights is so great as to overwhelm it. When the organization is exposed to an overwhelmingly large shock, the existing cognitive schemas are baffled, and the ensuing situation is marked by considerable uncertainty about what to do. “It is precisely in the face of massive uncertainty that beliefs of some sort are necessary to evoke some action,
which can then begin to consolidate the situation so that explicit inferences about cause/effect linkages can then be attempted” (Weick, 1983, p. 229-230). But, with a shock that is so extreme that it raises uncertainty well above the threshold for action, the organization will employ other mechanisms to dispense with the shock.

Isenberg (1986) observes that managers employ a process he calls “plausible reasoning” to “increase the certainty conditions beyond some critical threshold that changes as certainty conditions, riskiness, stakes, and other factors change” (p 247). The process begins as the manager “develops a different understanding of a phenomenon, often due to an experience of surprise” (Isenberg, 1986, p247), which we might call a large shock. The manager begins a selective search “to achieve a degree of certainty that will allow the manager to proceed … at minimal cost and minimal risk. The manager engages in action in the face of an incomplete but tentative understanding of the situation and uses the feedback of his or her actions to complete the understanding” (Isenberg, 1986, p247-248).

Unfortunately, a common approach to figuring out what to do involves consideration and analysis of many possible options, often at the expense of action. Moreover, once some decision as to a course of action is made, the organization embarks on a phase of predicting and planning that once again delays the initiation of action that will occasion sensemaking. Societal ideologies have reinforced the value of such planning, as Weick (1995) suggests in the following passage. (The words in parentheses identify seven properties of sensemaking.)

“Once people begin to act (enactment), they generate tangible outcomes (cues) in some context (social), and this helps them discover (retrospect) what is occurring (ongoing), what needs to be explained (plausibility), and what should be done next (identity enhancement). Managers keep forgetting that it is what they do, not what they plan, that explains success. They keep giving credit to the wrong thing – namely, the plan – and having made this error, they then spend more time planning and less time acting.” (Weick, 1995) p. 55

Another instantiation of a large shock that does not occasion change is found when the organizational participants see no relationship between the shock and the current realm of
action in which they are practicing. For example, evidence that suggests that a firm’s products will be obsolete in 10 years due to new technologies is not likely to have much effect on managers who believe their job is the manufacturing rather than the designing of products. Moreover, if they are accustomed to and occupied with acting and thinking about matters on a day-to-day or month-to-month time horizon, 10 years into the future may have no bearing on their current actions. In this scenario, uncertainty, confusion, or paralysis is not the reason that organization members do not know what to do. Rather, they do not know what to do because they are unable to translate the shock into an action that they believe is within their ability to control. Organizations decompose domains to prevent multiple simultaneous adjustments and treat the resulting subdomains as autonomous (Simon, 1997). If managers believe the limits on their domain of action prevent intrusion into other areas, they will turn to other matters relevant to their current domain and timeframe. Levinthal and March (1993) have described these tendencies to ignore the larger picture and to ignore the long run as two forms of organizational learning myopia.

A final example of large shocks that do not occasion action is a surprise layoff of employees at a paternalistic organization that has never before laid off employees. Brockner & Weisenfeld (1993) investigate the reactions of the people who remain, the survivors. Layoffs elicit survivor concerns about fairness and job security and arouse uncertainty about the situation and its meaning. The resulting state of confusion might restrain action.

“Situations that induce hesitation, alienation, or despair in anyone should be experienced as confusing because they make it harder for people to take actions around which meanings could crystallize. This hypothetical scenario sounds very much like what is reported by those (e.g., Brockner & Weisenfeld, 1993) who remain in their jobs after their co-workers are removed by downsizing. The confusion felt by those who remain stems not so much from their ‘survivor guilt’ as from their inability to act. Interventions that make it easier to bind people to action should reduce the confusion more quickly than would interventions designed to deal with feelings of guilt.” (Weick, 1995, p. 174).
Shock that Get No Novel Action

In this section, corresponding to the third row of Table 1, I discuss shocks that are noticed and acted upon, but the action taken is not novel. Organizational members can easily classify issues that are routine and expected and thus fit existing categories (Dutton et al., 1991). These issues elicit well-learned responses based on organizational “recipes” that are easily available and accepted by the organization as legitimate (Weick, 1979). Much organizational activity is the continuation of an ongoing flow. “Organizations frequently create action generators – automatic behavior programs that require no information-bearing stimuli” (Starbuck, 1983, p.93). Some activity follows these well-learned responses to easily classified stimuli in routine patterns of behavior (Cyert et al., 1963; Nelson et al., 1982).

Organizations use these routines to generate responses that deal with potential shocks. Such routine responses are found at the individual, group, and organizational levels. People are programmed into cognitive routines or habits that desensitize them to novel events (Starbuck, 1983). Indeed, “what people do most is often what they think about least” (Van de Ven, 1986, p 595). Groups make use of habitual routines, with both functional and dysfunctional consequences. Habitual routines in groups can reduce the likelihood of innovative performance processes (Gersick & Hackman, 1990). Organizations employ routines to generate the same response to a different stimulus. Weick (1996) refers to this as a non-learning sequence, and suggests that it “is probably the most common sequence found in organizations” (p. 166).

The point here is that when the shock is such that it is sufficiently similar to the ongoing flow of activity and the organization calls upon routine responses, the action that occurs already has meaning associated with it. The response is the routine response. The act is the application of existing cognitive schema to the situation, and the meaning attached is that it was based on the beliefs embodied in the existing schema. Small shocks are easily classified, and the perceived outcome is that the routine response has addressed the shock. The earlier example of a small increase in a production target at a manufacturing firm is just such a small shock. Since the new target is quite similar to the previous
production target, the organization classifies the stimulus in a category that elicits a routine response. The routine response to work a bit harder satisfies the production target, and the organization does not need to consider alternative responses, such as productivity or quality improvement. This is an example of first-order change, because it is "consistent with already-present schemata" (Bartunek et al., 1987, p.486).

Large shocks can also lead to responses that lack novelty. Consider the response of an organization to a shock that is perceived as a severe threat to the organization’s survival. Staw, Sandelands and Dutton (1981) draw upon findings from individual, group, and organizational levels of analysis to explain threat-rigidity effects. They argue that threat-rigidity effects lead to changes in both information processing and control processes that make the organization’s behavior less varied or flexible. Perceived threats restrict the number of alternatives for action that managers consider. In the face of threats, managers rely on fewer sources of information and emphasize information consistent with current schemas, simplifying and stereotyping in order to assimilate information into the current schemas. There is also a shift to more centralization of authority, more extensive formalization, and more standardization of procedures (Staw et al., 1981). Moreover, perceived threats seem to increase the importance of efficiency concerns, as organizations employ cost cutting, budget tightening, restriction of marginal activities, and intensified accountability (Starbuck & Hedberg, 1977). “Organizational actions associated with conservation of resources and tight control mechanisms are likely to be manifested in maintenance of the status quo, which in turn favors existing interpretation-action patterns” (Thomas et al., 1993, p. 244).

The notion that a large shock might lead to persistence of current beliefs and routines is evident in Ocasio’s (1995) analysis aimed at reconciling theories of failure-induced change with threat-rigidity theories. He concludes that:

“Economic adversity will increase the adoption of those types of organizational change that (1) have been well-learned, whether through prior organizational experience or through institutional mimetic processes; (2) are congruent with core assumptions and beliefs, as interpreted by decision makers; and (3) favor the
interests and identities of participants in the decision-making process” (Ocasio, 1995, p. 321).

Ocasio (1995) reminds us that the effect of economic adversity “is contingent on the social construction of mental models by participants in organizational subunits, as regulated by the institutional logic of the cultural system” (p. 320-321). Managers’ interpretations of the shock as either a threat or an opportunity influence information processing and decision-making (Dutton & Jackson, 1987). Thus, if the shock is interpreted as a severe threat, the organization’s response may lack novelty.

Ocasio (1995) points to a useful example in the U.S. automobile industry’s response to the increased threat from foreign competition in the 1970s and 1980s. Japanese manufacturers, most notably Toyota, developed and implemented various elements of lean production systems, which relied on low automation, high employee involvement, and management philosophies based on concepts such as just-in-time and one-piece flow (Womack, Jones, & Roos, 1990). General Motors’ responses included acquiring technology-based companies and increasing its level of automation, responses that were consistent with its core beliefs. However, they did not adopt the changes in manufacturing principles or human resource practices that challenged their core values and assumptions (Ocasio, 1995). The continued reliance on learned responses in the face of a large shock is also evident in the demise of the Saturday Evening Post, as chronicled by Hall (1976). Managers in this case continued to increase prices, a strategy that had been successful in the past, even as circulation was dropping (Hall, 1976). Greve and Taylor (2000) studied format changes in radio broadcasting to understand the effects of innovations by one firm on non-mimetic changes in other firms in the industry. They found that innovations by firms that were large and nearby were associated with less change and attributed this finding to threat-rigidity effects. The shock from an innovation by a large firm in the focal firm’s home or nearby markets is perceived as threatening, and novel responses are suppressed.
Shocks that Do Not Occasion Cognitive Restructuring

The fourth and last row of the framework of Table 1 deals with shocks that generate novel actions but do not occasion cognitive restructuring. On one extreme, we find small shocks that generate novel actions but no cognitive restructuring. The key premise underlying this category is that the beliefs that get attached to actions in the process of sensemaking are themselves versatile. Sensemaking is driven more by plausibility than by accuracy (Weick, 1995). There is a fuzzy vagueness about beliefs that endows them with the potential to be enacted in a variety of ways. When the sensemaker is searching to make sense of an action in the recent experience, an available explanation based on current beliefs, if plausible although not necessarily a perfect fit, will be satisfactory. Sensemakers, like decision-makers, engage in satisficing (Simon, 1955; Steinbruner, 1974). Consequently, when an action is novel but not too novel, the parsimonious sensemaker may dispense with this interruption, now the novel action, by plausibly mapping it into an existing belief.

In some sense, the situations included in this cell (no cognitive restructuring, shocks too small) of the framework are a subset of those in the cell immediately above it (no novelty, shocks too small). In the first case, the shock occasions action that is not novel; in the second case, the shock occasions action that is not novel enough. But, from the point of view of the sensemaker, these are indeed different sequences. In the first case, the sensemaker considers the stimulus, \( x \), to have been similar to previous ones, so the well-learned response, \( y \), is sensible because the routine recipe is “if \( x \) then do \( y \).” In the second case, the sensemaker acts in a manner that is not exactly \( y \) but is close and then considers the recipe to be adequate to explain “did almost \( y \) given \( x \).” Students of sensemaking are interested in “what happens when novel moments in organizations capture sustained attention and lead people to persist in trying to make sense of what they notice” (Weick, 1995, p. 86). What is missing in this failure mode is the element of sustained attention and persistence.
The tendency to associate a pre-existing belief with an observed action is related to two biases reported by psychologists studying judgment and decision-making. The hindsight bias denotes the tendency to perceive already observed outcomes as having been relatively inevitable and has been demonstrated in several studies by Fischhoff, who has also called this effect “creeping determinism” (1975, p. 288). People are generally not able to reconstruct the way an uncertain situation appeared to them before knowing the results. One way for the sensemaker to make sense of the observed action is to believe that it was inevitable under the circumstances. As long as the action is not especially novel, the sensemaker can resolve any uncertainty about why the action y should follow the stimulus x by believing is was clear all along. Now that action y did follow stimulus x, it must be what was expected. Confirmation bias may contribute to this effect as well. This bias describes people’s tendency to exclude the search for disconfirming evidence from their decision processes (Einhorn & Hogarth, 1978; Wason, 1960). An analogous tendency in sensemaking situations would mean that people are less likely to consider novel aspects of the shock and the action just observed, since the novel aspects could be disconfirming of the current schemas. People would give greater consideration to those situational aspects that are similar to experience. These similarities would provide adequate basis for the current schemas to make retrospective sense of an action that is only a little bit novel.

The final cell of the framework is for large shocks that occasion novel action but do not occasion cognitive restructuring. This cell describes the situation when observed actions change but the underlying belief systems do not. Having acted in a manner that is inconsistent with a person’s beliefs, an individual will experience an uncomfortable state of cognitive dissonance (Festinger, 1957). Dissonance theory suggests that individuals will attempt to reduce such states of dissonance. One way to reduce this dissonance is to update the beliefs to bring them into consonance with the recently observed action, which would be considered second-order change (Bartunek et al., 1987). However, in lieu of updating these beliefs, another approach to resolving the dissonance is increase the emphasis on information that makes the cognitions less dissonant, in short, to find an
alternative explanation. The context of the action taken may well make plausible alternative explanations available to the sensemaker.

The most apparent instantiation of this is a situation in which organizational members take novel action specifically as directed by an authority figure. These are situations in which the members adopt behavioral compliance with the proscriptions of their supervisors. The members recognize they are being urged to respond in a specific manner, and this urging by a legitimate authority figure provides a basis for compliance (Cialdini & Trost, 1998). Aronson (1995) refers to this as the “external justification” for the action (p. 198). A classic experiment by Festinger and Carlsmith (1959) studied this effect. Students performed some boring, repetitive tasks for a full hour as part of the experiment. Afterwards, the students were asked to lie by telling a confederate who was waiting to participate next that the task she was about to perform was interesting and enjoyable. One group of students was offered twenty dollars to lie, whereas the other group was offered only one dollar to lie. The key finding was in how these two groups differed in their subsequent ratings of the task. Those paid one dollar rated the task as an enjoyable one, and those paid twenty dollars rated it as dull. Those in the one-dollar condition in the absence of an external justification moved their beliefs (toward more enjoyable) in the direction of justifying their actions (of saying it was enjoyable). They reduced dissonance by updating their belief about the task. In the twenty-dollar condition, the subjects were able to justify their “lying” with the explanation that they did it for the money. They reduced dissonance without updating their beliefs about the task by relying on an external justification.

In an organizational context, sensemakers might preserve their cognitive schemas by justifying their behaviors with an external justification in the form of the explanation “my boss told me to do it.” Under the circumstances of a large shock that induces the organization to force managers to “walk the talk” (e.g., Kotter, 1995), this might result in avoiding the chance to explore alternative schemas:

“If they are forced to walk the talk, this may heighten accountability, but it also is likely to heighten caution and inertia and reduce the risk taking and innovation. This outcome occurs
not just because people are scared. It occurs because people who are forced to walk the talk prematurely often forgo exploration and walk on behalf of words that they barely understand. Because things that are poorly understood are things that tend to be seen as uncontrollable, they seem like threats rather than opportunities. Innovation shuts down. . . . What people forgo is the chance for walking to uncover something for which the current words are inadequate and for which new words are needed.” (Weick, 1995, p. 183).

Another characteristic of large shocks that might contribute to the likelihood of this failure mode is that responses to them take time to unfold. There might be a period of time during which organizational members are acting in a novel manner but have not yet come to closure about why they are doing so. For example, Isabella (1990) suggests that managers’ interpretations of key events evolve through a series of stages. They first “assemble rumors and other tidbits of information into an in-progress frame of reference,” then draw “on conventional explanations and comparisons to past events,” and later “compare conditions before and after an event” and “review the consequences of the event” (Isabella, 1990, p.14-16). After the introduction of a new alternative schema, time is needed for evaluation and comparison before the new schema becomes established (Labianca, Gray, & Brass, 2000). However, during this evaluation period, selective attention to information confirming existing beliefs might eventually lead the sensemaker to move the in-progress frame of reference towards the original cognitive schemas. The organization that introduces flavor-of-the-month change programs is prone to this effect, as organizational members adopt a wait-and-see approach, compliantly biding time while waiting for the program to die an early death. Moreover, a more radical new idea (a bigger the shock) might lead members to more readily seek information to confirm their belief that it will go away, such as information that the management does not really believe their own rhetoric. Thus, the novel action will recede and the old belief system will be maintained.

IMPLICATIONS
The purpose of this paper is to present a framework of how organizational activity in the aftermath of a shock might unfold in a manner that does not occasion cognitive restructuring and organizational change. Shocks, which may arise from both external and
internal sources, may be either too small or too large on a composite dimension representing proximity to the organization’s current cognitive schema. Four means by which organizations might dispose of these shocks are described. Organizations might not notice these shocks, might not take action, might not take novel action, or might not undergo cognitive restructuring. The paper uses this framework to discuss how small shocks and large shocks might occasion each of the four modes. The framework is presented as an initial organization of ideas, and it has implications for theory and for the practice of management.

*Implications for Theory and Research*

The discussion has implications for organizational theory and research regarding the role of shocks in the processes of sensemaking and organizational change and regarding the role of leaders. Several limitations of the foregoing treatment of shocks point to interesting questions for theory and research. First, the preceding discussion considered shocks essentially as one-time, isolated events. However, organizing is an ongoing activity (Weick, 1979), and shocks themselves may often have an ongoing nature. This gives rise to a dynamic process in which the cumulative nature of shocks may have important implications for the organizations that experience them (Rudolph & Repenning, 2002). Second, the representation of shocks along a one-dimensional continuum may be overly parsimonious. While this one-dimensional description is useful here as a simplification, it might be possible to identify a small set of dimensions that maintain the benefits of simplicity yet provide a richer framework for analysis and understanding. For example, a representation that included some temporal dimension and some other physical dimension might be useful to more clearly identify the boundaries between too small, moderate, and too large. Third, the foregoing discussion has ignored the influence of context. A similar shock in different contexts might occasion different responses. What are some of the contextual factors that might influence the likelihood of experiencing the various failure modes in the aftermath of a shock? How do these contextual factors effect the thresholds for too little or too much?
The discussion also points to some implications for studying the role of leaders during organizational change. Managers play a role in shaping shared understanding and conceptual schemes (Daft et al., 1984). The actions of management may play a symbolic role, and indeed one of the critical roles for managers is to construct belief systems (Pfeffer, 1981). The executive is exposed to different flows of information and often employs different perceptual filters than others in the organization (Finkelstein et al., 1996; Starbuck et al., 1988). Thus, managers are likely to influence their organization in both the selection and the interpretation of cues. In other words, managers influence the flow of shocks that organizational members might experience. Managers might create, translate, shape, or govern the timing of the flow of shocks throughout a firm. The discussion suggests that studying how managers influence these flows of shocks and the relationships between these flows of shocks and the various failure modes in the framework would be a fruitful stream of research. For example, consider managers’ use of stories, pictures, and symbols. “Collections of illustrations or stories, held together by a theory of action, provide a frame within which cues are noticed and interpreted” (Weick, 1995, p 120-121). Symbols are not just expressive media, associated with the symbolic role of management but are also one of the main means by which management accomplishes substantive action (Gioia, Thomas, Clark, & Chittipeddi, 1994). So, if stories and symbols are stimuli that either are themselves shocks or guide the interpretation of shocks, how do characteristics of these stories or symbols map into the framework presented here? How do managers employ stories and symbols to avoid the failure modes represented in the framework?

One value of the framework presented here, then, is an organizing concept for the study of the role of leaders as they offer cues and interpretations in the ongoing flow of organizing. This line of study may find some common ground between a top-down and a bottom-up view of organizational change. Organizational change has been characterized as a process of improvisation (Orlikowski, 1996), and the comparison with a performance of jazz music is often used (Weick, 1998). The interruptions from shocks might occasion improvisation (Weick, 1998). The leader has a role, not as the planner and director of change but as one who provides or guides the interpretation of shocks while allowing for improvisation and emergent change. The leader’s role in shaping the shocks to the
organization becomes one of forming belief systems for organization, generating the boundaries, and initiating improvisation.

Implications for Practice

The framework presented here has two implications for managers. Managers in some situations may wish to orchestrate shocks that are between the two extremes of too small and too large. The identification in this framework of failure modes associated with shocks that are too large should sensitize managers to the possibility that more change may result from a lesser shock. Second, managers should not expect to be gifted enough to orchestrate exactly the right level of shock on an ongoing basis. Rather, the informed manager will benefit from attending to cues that indicate whether the organization is responding to the shocks. Such a manager would monitor signals from the organization’s response to the shocks and adjust subsequent activity accordingly. When the organization seems to be responding with novel action and sensemaking activities, maintaining or perhaps increasing the shocks may be appropriate. Note that increasing the shock level here does not necessarily mean that the manager should overdo it. When the organization is not responding in the desired manner, the manager can use this framework to diagnose how the shocks need to be adjusted. When the manager observes responses that indicate a shock is too small, such as apathy, inaction, continued focus on excuses and barriers to action, reliance on old routines, or evidence of continued adherence to old ways of thinking, the manager should consider generating shocks that are greater in magnitude. Conversely, symptoms such as blatant lack of engagement, paralysis or outright fear of the future, excessive search for solutions without action or other strategies for marking time, or overt compliance accompanied by public attributions to policy as rationale indicate that the shocks are too large. The manager observing these symptoms should consider scaling back the magnitude of shocks. The key implication for many practicing managers will be to construct mechanisms by which they will have access to the appropriate cues to gauge these organizational responses. In other words, they must embed themselves in the appropriate feedback loops.
REFERENCES


147


### Table 1

Organizational Responses to Shocks that Do Not Occasion Cognitive Restructuring

<table>
<thead>
<tr>
<th>SHOCKS THAT ARE:</th>
<th>Too Small</th>
<th>Too Large</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FAILURE MODE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No Attention</strong></td>
<td><strong>Don’t notice</strong>&lt;br&gt;&lt;<em>Shock</em>: Below the perception threshold&lt;br&gt;&lt;<em>Response</em>: Filtered as noise</td>
<td><strong>Can't notice/don’t care</strong>&lt;br&gt;&lt;<em>Shock</em>: Not on the “radar screen”&lt;br&gt;&lt;<em>Response</em>: Filtered as unrelated to relevant concerns</td>
</tr>
<tr>
<td><strong>No Action</strong></td>
<td><strong>Don’t bother</strong>&lt;br&gt;&lt;<em>Shock</em>: Below the action threshold&lt;br&gt;&lt;<em>Response</em>: Ignored as aberrations</td>
<td><strong>Don’t know what to do</strong>&lt;br&gt;&lt;<em>Shock</em>: Outside the realm of action or above the uncertainty threshold&lt;br&gt;&lt;<em>Response</em>: Overtaken by current actions</td>
</tr>
<tr>
<td><strong>No Novelty</strong></td>
<td><strong>Don’t adjust</strong>&lt;br&gt;&lt;<em>Shock</em>: Similar to prior experience&lt;br&gt;&lt;<em>Response</em>: Addressed with routine responses</td>
<td><strong>Don’t search much</strong>&lt;br&gt;&lt;<em>Shock</em>: Perceived as threats or crises&lt;br&gt;&lt;<em>Response</em>: Threat-rigidity effects</td>
</tr>
</tbody>
</table>
| No Cognitive Restructuring | **Don’t sweat it**  
*Shock:* Below the implausibility threshold  
*Response:* Explained with a current belief | **Don’t acknowledge**  
*Shock:* Lofty or authoritarian pronouncements  
*Response:* Made sensible as compliance |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Find new meaning for novel action</td>
</tr>
</tbody>
</table>