Self-Confirming Attribution Errors in the Dynamics of Process Improvement

Nelson P. Repenning[1]

John D. Sterman[2]

Sloan School of Management
Massachusetts Institute of Technology
Cambridge, MA USA 02142

Version 1.0

November 2000

Work reported here was supported by the MIT Center for Innovation in Product Development under NSF Cooperative Agreement Number EEC-9529140 and the company described in this paper. Many thanks to Tim Tiernan, Bill Colwell, Laura Cranmer, Dave Lazor, Vic Leo, Frank Murdock, Roger Saillant and Ron Smith for their generous assistance. We thank Bob Cole, Dick Scott, Lotte Bailyn, John Carroll, Lee Fleming, Drew Jones, Steve Graves, Liz Keating, Tom Malone, Rebecca Henderson, Rogelio Oliva, Wanda Orlikowski, Scott Rockart, Julio Rotemberg, Jenny Rudolph, Ed Schein, Peter Senge, and seminar participants at MIT.

For more information on the research program that generated this paper, visit http://web.mit.edu/nelsonr/www/.

1. MIT Sloan School of Management, E53-339, Cambridge, MA USA 02142. Phone 617-258-6889; Fax: 617-258-7579; <nelsonr@mit.edu>.

2. MIT Sloan School of Management, E53-351, Cambridge, MA USA 02142. Phone 617-253-1951; Fax: 617-258-7579; <jsterman@mit.edu>.

Self-Confirming Attribution Errors in the Dynamics of Process Improvement

Abstract

Theorists have long distinguished between two types of organizational change: change focused on improving internal efficiency and change focused on improving fit with the external environment. While many models of change build on this distinction, until recently most scholars have focused on radical, externally-focused change and less effort has been dedicated to understanding its more incremental, internally-focused counterpart. To better understand the factors that both support and inhibit internally-focused change, in this paper we report the results of an inductive study of one firm's attempt to improve two of its core business processes. Our data suggest that the critical determinants of success in efforts to improve continuously are managers' attributions regarding the cause of poor organizational performance. Building on this observation, we propose a dynamic model to capture the mutual evolution of those attributions, managers' actions, workers' responses, and the production technology. Using this model, we show how managers' beliefs about those that work for them, workers' beliefs about those who manage them, and the physical structure of the environment can co-evolve to produce an organization characterized by protracted conflict, mistrust, and increasingly rigid, inflexible control structures that prevent useful change of any type.

Self-Confirming Attribution Errors in the Dynamics of Process Improvement

Introduction

Theorists have long distinguished between two types of organizational change: change focused on improving internal efficiency and change focused on improving fit with the external environment. Manifestations of this distinction include exploration and exploitation (Levinthal and March 1981, March 1991), convergence and reorientation (Tushman and Romanelli 1985), evolution and revolution (Tushman and O'Reilly 1996), and incremental and radical change (Brown and Eisenhardt 1997). While many models of change build on this distinction, until recently most scholars have focused on understanding the causes and consequences of radical change and less effort has been dedicated to understanding its more incremental counterpart. For example, Brown and Eisenhardt (1997:1) conclude in their summary of recent research, "…incremental change is assumed to occur [and] radical change is the focus of interest."

While researchers have dedicated considerable attention to understanding the (in)ability of firms to undertake radical change, accumulating evidence suggests that incremental change can be equally daunting. One manifestation of such difficulties is the inability of many organizations to adopt best practices *within* existing technologies and orientations (Pfeffer and Sutton 2000). Thus, for example, despite compelling evidence that they improve performance, most major US manufacturing firms have not adopted the disciplines associated with TQM (Easton and Jarrell 1998). Similarly, Wheelwright and Clark (1995) lament that, despite the fact that the best practices reported in their earlier book on product development (Wheelwright and Clark 1992) have led to dramatic improvements in isolated projects, firms struggle to implement such changes on an ongoing basis. Little theory exists to explain how and why organizations struggle with incremental changes like adopting best practices (Dean and Bowen 1994).

The purpose of this paper is to examine the factors that conspire to prevent incremental improvements to existing technologies and competencies. Specifically, we study the attempt of one firm to improve its product development and manufacturing processes. Process-focused improvement presents a useful context in which to study the micro-processes that either impede or facilitate competence-enhancing change for a number of reasons. First, while there is some debate about the location of process improvement on the exploration/exploitation continuum, such programs are internally focused and typically do not entail radical changes to the firm's strategic position. Second, such efforts often fail (Dean and Bowen 1994, Easton and Jarrell 1998), thus providing an entry point into understanding the forces that limit their success. Third, improvement programs are a frequently attempted (and frequently studied) type of change, and thus are of some interest to organizational scholars in their own right (Sitkin et al. 1994; Hackman and Wagemen 1995; Westphal, Gulati and Shortell 1997; and Zbaracki 1999 are recent examples of studies related to the use of TQM).

We develop our theory inductively based on two original case studies of process improvement efforts within a division of a US automobile manufacturer. We chose a grounded-theory approach based on its likelihood of producing novel insights into the organizational dynamics surrounding process improvement (Eisenhardt 1989). The focus on novelty resulted from our assessment that the literature had reached some consensus on the forces that drive competence-enhancing change (see Tushman and Romanelli 1985 and March 1991 for general discussions, and Hackman and Wagemen 1995, Dean and Bowen 1994, and Sitkin et al. 1994 discussions specific to TQM) coupled with previous work (Carroll, Marcus and Sterman 1997; Sterman, Repenning and Kofman 1997; Pfeffer and Sutton 2000) suggesting that the inner workings of process improvement efforts are more complicated and problematic than its advocates suggest.

Our theory extends the existing literature in several ways. The data suggest that the critical determinants of success in process improvement programs and related change efforts are managers' attributions regarding the cause of poor organizational performance. Building on this observation, we study the cognitive and social processes through which managers make such attributions. Applying existing theory suggests that managers' attributions of cause are biased towards the people that work for them and away from the production system itself. We push beyond the existing literature by specifying the dynamic processes through which these attributions lead to actions which then create the cues that are used to update causal attributions. The model we propose suggests that initial attributions can be strongly self-confirming. Managers who initially believe that people are the cause of their problems take actions that both force employees to act in accordance with those beliefs and embed those beliefs in the physical structure of the organization. Over time managers' beliefs about those that work for them, workers' beliefs about those who manage them, and the physical structure of the environment can co-evolve to produce an organization characterized by protracted conflict, mistrust, and increasingly rigid, inflexible control structures that prevent useful change of any type.

The contribution of our work to organization theory is to identify a set of processes grounded in basic cognition and the physical structure of production technologies that can enact a highly undesirable environment. A critical implication is that the increasingly rigid routines, structures, and norms normally associated with an "efficient" organization, can result not from efforts to exploit the potential efficiency gains latent within the production technology, but from efforts to exploit the people who work within that process. Our theory also suggests that improvisation and adaptation, usually depicted as beneficial, are general organizational processes that can help or hurt performance depending on the context in which they take place.

The paper is organized as follows. We begin by discussing the research design and the data collection and analysis methods. We then summarize the most striking features of our data and frame the central questions on which we focus our theory development effort. To present our model we first discuss the dynamics at the individual level and then extend it to include both line workers and the managers who supervise them. Finally, we consider linkages to the changing structure of the physical production technology. We conclude with a discussion of implications for practice and future theorizing.

Method

Research Design

The research was conducted within a division of a major US automobile manufacturer. The division designs and manufactures electronic components that are then integrated into vehicles at separate body and assembly facilities. At the outset of the research, a team of people working at various levels within the division was assembled to assist the first author in performing the field research. The team included a division vice-president, people who had (or previously had) line responsibility in assembly plants and product development projects, and members of various internal consultancies that provided services to development projects and plants.

The research design focused on extreme or polar types (Eisenhardt 1989: 537), so the team began by identifying initiatives that were either dramatic successes or failures. We expected that the processes we were most interested in, those preventing competence-enhancing change, would be closer to the surface in dramatic successes and failures, and that comparison of these extremal examples might yield interesting insight. The team reviewed a large number of change initiatives as candidates for study. Additional selection criteria included the size and scope of the effort, the level of investment in the effort by senior leadership, and the subsequent impact on organizing practice.

Extended discussions among team members (who typically had substantial experience with the initiatives under consideration) led to the choice of three initiatives: a cycle time reduction/quality improvement effort in manufacturing (the MCT initiative); a cycle time reduction/productivity improvement initiative in product development (the PDP initiative); and an internal supplier certification initiative patterned after the Baldrige award (the Total Quality Excellence or TQE initiative). The MCT and PDP initiatives provided a unique opportunity to study process improvement for a number of reasons. First, the two initiatives provide a stark contrast in results: the MCT effort was dramatically successful, leading to a twenty-fold reduction in average manufacturing cycle time and savings of hundreds of millions of dollars, while the PDP initiative failed to achieve most of its objectives. Second, despite these differing outcomes, the same senior executive (the general manager of the division) led both initiatives, and the second effort, PDP, was begun in large part due to the success of the first. Their comparison offers a rare opportunity to control for the impact of senior leadership and management style. Third, of the two, PDP received substantially more organizational support, had a dramatically larger communication and training budget, involved a larger group of people supporting and promoting it, and generally conformed more to the conventional wisdom concerning successful large-scale organizational change.

After initial interviews and data collection, the TQE initiative was dropped from this component of the study when it became clear that it was more focused on documentation than actual process improvement (the results of the TQE study are reported in Johnsson 1996).

Data Collection

The primary data collection method was semi-structured interviews. Over sixty interviews were conducted with participants in the two initiatives. The field researcher was given wide access to all levels with the organization; interviews included the most senior person in the division (the general manager), the executives in charge of both initiatives, plant managers whose facilities adopted the MCT program, product development vice-presidents and business unit managers, line supervisors in the manufacturing plants, managers whose projects piloted the PDP effort, operations and manufacturing engineers in plants, product engineers, machine operators, and material handlers. The field researcher conducted studies at two manufacturing facilities in North America, the product development and research center, and the corporate headquarters.

Interviews lasted between 45 and 90 minutes and were recorded on tape. Interview subjects were each presented with a one-page outline of the topics the interviewer wished to cover. Each interview began with the subject describing his or her background with the organization and relevant previous experience. Participants were then asked to give a detailed account of their experience with the initiative. Subjects were asked to assess the key successes and failures of the initiative and to offer hypotheses for their causes. Finally, subjects were asked to describe any lessons learned and to speculate on what they would do differently if they were to participate in a similar initiative in the future. As soon as practically possible (usually the same evening), the interviewer listened to the interview tape, reviewed the notes taken during the interview, and wrote a detailed summary of the interview and his initial reactions. Later the interview tapes were transcribed.

The interviews were supplemented with both extensive collection of archival data and the observation of practice. We were given access to a wide range of promotional and training material associated with each initiative including pamphlets, newsletters, instructional books, and video and audiotapes. Historical performance data were also collected. In the case of the MCT effort, extensive data on actual cycle times, product quality, productivity and other operational variables were available. Fewer data were available for the PDP effort.

Data Analysis

Traditional Approaches

We began our data analysis with the traditional techniques espoused by Eisenhardt (1989) and Miles and Huberman (1984). The analysis began with the first author reading all the interview transcripts, notes taken during the interviews, and the post-interview summaries. A time line for each of initiative was then developed. The time line provided the basic structure for writing two detailed case histories. The cases describe the history of the each initiative drawing on the quantitative data, archival materials, and recollections of participants. After the cases were drafted, all interview subjects were given the case documents to review their quotations for accuracy. They were not allowed to change the content unless there was a factual dispute that could be resolved with additional data. Participants were also asked to review the entire case for accuracy. Case reviews often led to the interviewee providing additional background or history. The research team also reviewed the cases in detail, identifying missing elements of the narrative and suggesting additional interview subjects or data collection. The cases are available from the authors upon request.

Causal Loop Diagrams

We then turned to developing a theory to explain the evolution of each initiative. The primary tool used for theory development was the causal loop and stock and flow diagramming method commonly used in system dynamics (e.g. Richardson and Pugh 1981, Sterman 2000; see Weick 1979, Masuch 1985, and Sastry 1997 for the use of such devices in organization theory). Causal loop diagrams provide a convenient and precise technology for articulating a process theory (e.g. Mohr 1982) describing how a system evolves over time. We began by developing diagrams to describe particular episodes within each initiative, then developed a summary of the main structures that drove the overall success or failure of the initiative. Finally, we integrated the diagrams into a unified framework that could explain both the observed successes and failures. During this phase, as the main elements emerged, we often returned to the cases and the original data to check for and resolve any anomalies or contradictions.

Connections to the Literature

Finally, as the diagrams that summarized our model emerged, we reviewed each link in the causal maps with an eye towards assessing whether or not the relationship was supported by existing studies. As this phase unfolded, we relied heavily on cognitive and behavioral decision making theory since these frameworks connect peoples' mental models to their judgments and decisions. The hypothesized relationships were strongly supported by this literature. Once integrated into one model, however, the combination of these linkages, by interconnecting physical structure and decision making behavior at the micro level, allows us to explain organizational-level dynamics. The end result of this process is, we believe, a theory that is well grounded in our data and consistent with existing theory in both the management and organization sciences and the experimental literature on human decision making.

Overview of the Two Initiatives

We begin with a brief history of both initiatives.

Manufacturing Cycle Time (MCT)

State of the System Prior to the Initiative

Prior to the MCT effort, the division's plants were operated like those of many companies whose business requires substantial capital investment and labor expense. Line supervisors were charged with keeping each piece of equipment and each laborer fully utilized. The performance measurement and evaluation system emphasized direct labor performance (roughly defined as the number of units produced per person per day) and gave supervisors a strong incentive to keep high levels of work-in-process inventory (WIP) to ensure that breakdowns and quality problems at upstream machines did not force downstream machines to shut down. A large portion of each plant's floor space was dedicated to holding WIP inventory. An operations manager recalled,

> Before [MCT] if you were to walk out onto the floor and ask a supervisor how things were going, he would say "Great, all my machines are running" and you would see tons of WIP sitting around.

High WIP levels hobbled plant performance in several ways. First, carrying WIP was expensive–between sixty and eighty percent of the division's total costs derived from purchased components. Second, a high level of WIP delayed quality feedback—a machine could produce a large batch of defective parts before the defect would be discovered by a downstream operation. Third, it was difficult for the plants to change the production schedule on short notice—high WIP implies a long cycle time. Last minute changes were accommodated through expediting (which destabilized the production floor by forcing operators to do more machine set-ups and change-overs), by reducing lot size, and by increasing production pressure. High WIP levels and expediting were adaptations through which the system had evolved to be tolerant of quality and reliability problems.

Launching the Initiative

A new general manufacturing manager (GM), recently recruited from a leading electronics firm, launched the MCT initiative. He recalls the genesis of the effort:

> We analyzed [for a sample product] the time elapsed between when a part came in the back dock until the time it left the shop floor, and asked the questions “How long did it take?” and “What was the value-added?”. We found out it took 18 days to make the product and we were adding value to the product 0.5% of the time.

Based on this analysis, the GM concluded that substantial improvement could result from reducing the time products spent between operations as opposed to the conventional focus on reducing the time parts spent on a particular machine. He explains:

> Many people thought of cycle time as the cycle time of the equipment. They were looking at reducing the time a part spent on a particular piece of equipment from 20 seconds to 10 seconds. My feeling was when you are at 18 days who gives a rat’s ass about the cycle time of specific machines.

To launch the effort, the GM employed a hands-on approach that entailed spending much of his time visiting the division’s plants. He describes the character of these visits:

> ...[people in the plants] wanted to give me presentations in the conference room, and I would say “no, let’s go out to the floor”... I wanted to show them examples of what I was talking about. I might look at the shipping labels in the warehouse. If it were May, I would usually find parts that had been received the previous August, and I would ask, “if you aren’t using this stuff until May, why has it been sitting here since last August?”

These trips stimulated interest in the effort, and following these visits, a few plants undertook an intense period of experimentation. Early efforts focused on developing appropriate metrics for cycle time and value added percentage. Improvement began almost immediately. As one plant manager recalls. In the first year, cycle time at that plant fell by more than fifty percent.

MCE Analysis

In the middle of the second year, a four-person group was created at the division headquarters to promote the initiative throughout all the plants. The group began by institutionalizing a measurement system based on the experiments performed at the early adopter facilities. Each plant was required to calculate a metric called Manufacturing Cycle Efficiency (MCE), defined as the ratio of value-add time (time in which a function or feature was being added to the product) to total manufacturing cycle time. The early results were not encouraging. Another plant manager recalled, "...when we first started to calculate MCE, the numbers were so low [less than one percent] we really wondered how relevant they were." The process, however, proved valuable. A staff member recalled:

> ...you had to walk through the shop floor and ask the question, "Is this value added?" for every step in the process. By the time you were finished you had flow-charted the entire process and really highlighted all the value add stations....After calculating MCE, we really started to understand the process flow of our products. We knew where value was being added, and, more importantly, where value was not being added.

Within a year, the focus on MCE helped cut the average cycle time for the division to less than five days, down from the initial fifteen day average.

Theory of Constraints

Two years into the initiative, with the MCE analysis well underway in most facilities, the corporate staff focused on shop floor management as the next opportunity for reducing cycle time. The corporate group became interested in the offerings of the Goldratt Institute, which taught the shop floor management philosophy Theory of Constraints (TOC) developed by its founder Eli Goldratt (Goldratt and Cox 1986). The attraction of the Goldratt group was twofold. They offered a proven shop floor management strategy, and, perhaps more importantly, they offered a training program focused on developing intuition through hands-on experience with a computer simulator. The supervisor of the manufacturing simulation group recalled,:

> I called it "Shop Floor Scheduling and Coordination Awareness 101". If you wanted to concentrate in three days everything you would want to understand about the dynamics of the shop floor and how to keep the line running, this was it.

The division made a substantial commitment to disseminating the Goldratt training. Within six months almost every manufacturing engineer and supervisor within the division had participated in a two day TOC class. In the following year, the division developed a hands-on, board game version of the simulator and used it to train almost every operator and material handler within the division. In addition, line supervisors made TOC training a part of their daily operations. One supervisor who experienced substantial success using TOC recalls:

> We started by teaching each of the work teams how to manage their line using TOC...the classes were useful, but I felt the real learning came from working with them on their lines on the floor. I would coach them through making actual decisions. I'd let them make the decisions and then we would talk about the results.

Over time, TOC was widely accepted in the division and continues to play an important role in managing the plants.

Results

By almost any measure, the MCT effort was very successful. Between 1988 and 1995 the average manufacturing cycle time fell from approximately fifteen days to less than one day, product quality improved, and revenue, profit, and cash flow all increased significantly. The manufacturing process also became less elaborate and more flexible. Many facilities are now able to change their production schedule on a daily basis, something that was impossible before MCT. Finally, the reduction in WIP created enough extra floor space within existing plants that two of five planned new facilities were not needed, saving hundreds of millions of dollars in capital expenditures.

Product Development Process (PDP)

Designing a New Development Process

The second initiative, focused on improving the division's product development process, was initiated in large part due to the success of MCT. The general manufacturing manager who launched MCT was promoted to general manager of the division. He launched the PDP initiative by forming a dedicated task force charged with designing and implementing a new development process:

> We need a development process that is fast, is the best in the industry, and it needs to increase throughput by 50% in two years. And everyone must adhere to the same process.

The task force included representatives from the major functions within the organization. Their activities included working with an outside consultant, benchmarking other companies, and documenting the current process. A team member summarizes,

> We spent a substantial amount of time looking at what other people did, how they structured their processes and the problems they had. We looked at...the current state of our process and tried to net out a process that had all the things we wanted and...allowed us to do things much more quickly.

The New Product Development Process

PDP was not the first attempt to improve the development process. Over the preceding decade numerous attempts had been made to speed product development, with mixed results. At the time PDP was launched, two separate improvement initiatives were already in progress. The PDP team consolidated these two efforts along with the benchmarking results and the input of people throughout the company into a detailed new product development process for the division. Three elements distinguished the new process from prior practice.

First, PDP was a "one pass" development process. Historically, projects were started with ambiguous customer requirements and, consequently, numerous physical prototypes were created as the requirements for the final product were updated. Developing multiple prototypes was time consuming and expensive. To combat this "build and bust" cycle, PDP mandated detailed documentation of customer requirements before the design process began. Once the requirements were established, engineers would then do the majority of the design work using computer-aided design tools. The combination of detailed, up-front documentation of customer requirements and use of computer design would allow new products to be developed with one physical prototype and little rework, saving time and engineering resources.

A second goal of PDP was to propagate learning using the "bookshelf." The division did not share technological learning well, causing substantial duplication of effort. The bookshelf was to be an engineering library of technologies, modules, and subsystems. Every time a new technology was used, it was the designer's responsibility to document its uses, capabilities and limitations, then place it on the bookshelf. To complement the bookshelf, PDP also specified a "wall of innovation." Projects using new and unproven technologies often fell behind schedule or suffered from quality problems. The wall of innovation was the point in the development process beyond which every project had to be based on technologies on the bookshelf. Its purpose was to prevent projects from proceeding too far into the development cycle with technologies that had yet to be appropriately tested.

Third, the PDP process was designed to increase discipline. The process was divided into six major phases, and at the end of each phase development teams were required to undergo a "phase exit quality review" before proceeding to the next step. The reviews, conducted by senior managers, required development teams to assemble detailed documentation on the state of the project. Among other things, the phase exit quality reviews were implemented to enforce the wall of innovation: managers were supposed to prevent teams from proceeding to the next phase until each of the technologies they planned to use was documented and placed on the bookshelf. Between reviews projects were to be run using standard project management techniques such as work plans, Gantt charts, and project management software. By using project management tools, engineers would be more accountable, efficient, and better able to meet critical milestones.

Pilot Development Projects

The design team tested the new process on a number of pilot projects. The pilots served two purposes. They provided an opportunity for the team to identify and correct problems in the process. And if they were successful, the pilot projects could be used as examples to drive the process through the organization. The first pilot project was a high profile product critical to the corporation’s image and financial success.

The pilots suffered, however, because much of the support infrastructure required for the new process was not in place. Engineers did not have computers powerful enough to use the new CAD/CAE/CAM software, and once the new computers arrived, the rest of the organization was not able to accept their output due to software incompatibility. In addition, learning how to use the tools imposed a substantial burden on the already overworked engineers. One engineer recalled:

> ...I had some background in CAD/CAE from my master’s program and I still stayed at work until midnight every night for a month learning how to use the tools and trying to figure out how to get my work done....Some of the older engineers, even with training, they just have a [computer] sitting on their desks gathering dust.

While another said:

> ...the value of the tools was way overestimated...we never had time to take the courses and get the equipment we needed to really make this stuff work....it was really exhausting trying to learn how to use the tools and do the design at the same time.

The project also required the use of new and unproven technologies. As the first test of the new process, the bookshelf of documented designs was nearly bare, and, consequently, engineers were not able to achieve the "one pass" design dictated by the PDP process. Instead, much of the design was substantially reworked late in the development cycle, increasing work pressure and stress on members of the project team.

To meet the project schedule and specifications many of the engineers working on the pilots abandoned much of the methodology. One recalled, “...we crashed through the wall of innovation and never looked back.” The effect of these problems on the morale of the engineers was significant. Every interviewee reported being frustrated with the process. Many felt that management had defined a development process and then immediately gave the engineering staff a project and a time line that could not be accomplished using it. A common sentiment was expressed by an engineer who said, “...I believe PDP is a good process. Some day I’d really like to work on a project that actually follows it.”

Results

Evaluating the success of the PDP initiative is difficult. There is little quantitative data with which to evaluate the success of the initiative. Further, the parent company and the division have undergone numerous reorganizations since the initiative. The lack of data caused by the long cycle times for development projects is a key feature of the feedback structure governing the success of the program and not just a problem for researchers. Without rapid feedback on results, people formed judgments about the effectiveness of PDP through personal experience, anecdote and rumor. Indeed, despite the lack of hard data, many people developed strong feelings about the impact of the program. The vast majority believed that the process as designed was good, but that the division as a whole did not follow it. The GM rated the effort as a fifty-percent success. The executive in charge of PDP believes that they achieved eighty to ninety percent of their objective for the use of new tools but less than twenty percent of their objectives for documentation of customer requirements, using project management, and developing a more rigorous and repeatable process. Members of the design team also believe the effort failed to achieve its objectives, but hoped it would provide a catalyst for future improvements. Among the engineers interviewed, not one believed the initiative had materially influenced his or her job.

Conflicting Objectives and Conflicting Attributions

We begin our analysis by summarizing the most striking features of the data, then propose a model to explain the phenomena we observed. We first specify the model at the individual level of analysis, then extend it to capture group dynamics. Finally, we consider dynamics around the structuration of the environment.

Despite the large differences in context, the interviews revealed some striking similarities between the MCT and PDP initiatives. First, while the two environments are quite different, when asked why they didn't engage in the activity suggested by the initiative (e.g. run experiments to reduce cycle time in MCT, or place designs on the bookshelf in PDP), almost every participant described a basic trade-off between doing their "real" work and the additional improvement work required by the initiative. Participants also universally reported being under tremendous pressure to meet short run objectives, thus leaving little time to engage in improvement work. For example, a staff member located at a manufacturing plant describes the challenges she faced in getting line supervisors to participate in improvement activities:

> In the minds of the [operations team leaders] they had to hit their pack counts. This meant if you were having a bad day and your yield had fallen ... you had to run like crazy to hit your target. You could say "you are making 20% garbage, stop the line and fix the problem", and they would say, "I can't hit my pack count without running like crazy." They could never get ahead of the game.

Similarly, an engineer from a PDP pilot project explains the difficulty he experienced in trying to use the project management techniques required by the PDP process:

> ...under this new system the engineer was responsible for doing the physical design work using the new tools...however, none of our old tasks went away, so the new workload was all increase...in some cases your workload could have doubled...many times you were forced to chose between doing the physical design and doing the project and administrative work. To be successful you had to do the design work first, but the system still required all this extra stuff.

While another from a different pilot project was more blunt:

> People had to do their normal work as well as keep track of the work plan. There just weren't enough hours in the day, and the work wasn't going to wait.

Importantly, while both engineers felt that they did not have time to complete all the activities required by PDP, neither contested its benefits. In fact, all of those interviewed agreed it constituted a better way to develop products than the old process. For example, one engineer summarized his experience by saying, "The tenets of PDP are good; using them is a different story."

Nearly every line-level worker explained her failure to use new processes and tools by invoking the basic trade-off between investing in long-run improvements and hitting short-run production targets. Further, interviewees also universally reported being under intense pressure to achieve their production objectives, thus causing them to spend little time on improvement. Before the MCT effort, manufacturing line supervisors and machine operators, fearful of missing their throughput objectives, would not stop their machines to do preventive maintenance, fix problems that led to low yield, or run experiments that might lead to yield improvements. Similarly, product development engineers would not use project management tools, invest in learning computer-aided design and development tools, or take the time to document their uses of new technology for fear of missing delivery dates for their "real" design work. To act otherwise was universally seen, as one manager put it, as "…a career limiting decision." One engineer said, "The only thing they shoot you for around here is missing product launch. Everything else is negotiable." Thus, we were initially intrigued by the question of how the organization had evolved to this state of paralysis by production pressure, where aggressive short-run objectives coupled with intense pressure to achieve them prevented any form of learning or improvement.

We are, of course, far from the first scholars to identify the trade-offs between improving and working. It appears in many previous studies of process improvement (e.g. Carroll, Marcus, and Sterman 1997), has been the subject of various formal, rational-actor models (e.g. Fine 1986), and has become a staple of pop culture personal and business oriented self-help books (e.g. Covey 1989, Senge 1990). The interesting question for organizational theory is why, despite ample evidence suggesting that improvement and learning are worth substantial investment, and the vast array of scholars, consultants, and self-help gurus who preach its virtues, many people still grossly under-invest in such activities (see Easton and Jarrell 1998 for a discussion of TQM and Pfeffer and Sutton 2000 for a more general discussion). The answer can't be as simple as "high discount rates."

The second interesting observation arose initially from the PDP effort. While the engineering staff universally reported facing conflicting objectives and having little time for improvement activities, more senior managers attributed the failure of the initiative to "lack of discipline" in the development process and among the engineers. For example, the executive in charge of the PDP design team recalls the results of its initial assessment:

> ...we found...[the existing development process] was...poorly documented and poorly disciplined....Engineers by trade, definition, and training, want to forever tweak things.... It was a little bit of a Wild West culture.

Similarly, a chief engineer characterizes his view of the process before PDP:

> We went through a period [prior to PDP] where we had so little discipline that we really had the "process du jour." Get the job done and how you did it was up to you....It allowed many of the engineering activities to go off on their own and as long as they hit the key milestones, how they got there wasn't that important.

Based on these assessments, the executive in charge declared that the prime objective of PDP was "to instill some rigor, repeatability, and discipline into the process."

Despite their considerable efforts, however, the consensus of those in leadership positions was that PDP was not successful in achieving this objective, and the source of difficulty lay with the engineering staff. One chief engineer said:

> …it was fair to say that a lot of engineers viewed this as a neat way to get some fancy tools and to hell with process.

Similarly, the leader of the effort recalls:

> A lot of the engineers felt that [PDP] was not value added and that they should have spent all their time doing engineering, not filling out project worksheets. It's brushed off as bureaucratic.

There was remarkable consensus on the part of engineers, line supervisors, and operators as to why improvement efforts failed in both areas, but it was not shared by executives in the product development arena. They attributed the failures much more to general stubbornness on the part of the engineering staff.

Given this gap in the assessments of engineers and managers, we turned our attention to the question of whether such a contrast existed in manufacturing either prior to or during the MCT initiative. Operators and line supervisors clearly highlighted the basic trade-off between improvement and work in manufacturing and reported being under intense pressure to reach objectives, but assessing the attributions of more senior managers prior to the initiative was more difficult. While managers typically did not reveal much information about their pre-MCT assessments (which, given the subsequent success of the effort, were subject to hindsight bias) they did give us detailed information about the state of the measurement and performance evaluation schemes used in the plants.

Before the MCT effort the plants operated under a very constraining measurement scheme designed, in the words of one supervisor, "...[to] make sure every worker and every machine was busy all the time." Line supervisors were evaluated on their labor efficiency–roughly the number of units produced per person–on a daily basis and their performance was often scrutinized by managers at the highest levels within the organization. For example, the general manager who started the MCT initiative describes his first experience with senior executives at his new employer:

> I [first] came as [the] general manufacturing manager [responsible for all manufacturing in the division]…[in] the meetings that I was going to they were asking about the machine utilization of equipment in [one facility], you know, a particular piece of equipment.…I couldn't figure out why, you know? To me, I thought it was nonsensical, but to the executive VP who was asking…he didn't view it as nonsensical. They wanted to know about machine utilization and up times and stuff like that.

The executive vice-president was a very senior manager with responsibility for the entire division, which, at that time, had revenues of over $5 billion per year.

Accompanying the detailed focus on utilization was an incentive scheme that strongly discouraged missing daily production targets. As one manager recalled, "... supervisors who missed their targets knew they were going to get 'beat up' by their managers." Consequently, line supervisors focused their attention on meeting those objectives. An operations manager at one plant recalls, "...supervisors would always hit their exact targets, if the goal was 200, they would pack [produce] 200, never 198, never 202." Another noted "...[supervisors] would make sure everybody was busy all the time to make labor efficiency."

Thus, the manufacturing environment before MCT was quite similar to that in product development. Employees in both areas faced a strong trade-off between doing their daily work and engaging in process improvement activities. Further, the incentive and measurement schemes in both systems had evolved to the point where workers did not feel that they could risk missing their daily objectives by engaging in improvement efforts. Managers evaluated the efforts of machine operators in manufacturing on a daily basis and imposed severe penalties for low performance, while engineers were required to produce detailed documentation concerning their development efforts. These observations turned our attention toward explaining why the system evolved so that engineers and operators, despite feeling it was not the best thing to do, focused exclusively on their self-described "real work" and never invested in process improvement, while, at the same time, managers did not recognize this trade-off and, instead, felt that their attempts at improvement failed due to lack of discipline on the part of the workforce.

The Model

The model integrates the physical setting and technology, organizational structures and routines, and the mental models and behaviors of the workers and managers. Their interaction generates the dynamics of the organization and feeds back to affect all three: technology, organization, and mental models.

The Physical Structure of Improvement

The first construct in our model is *Net Process Throughput*. *Net Process Throughput* is the rate at which inputs are successfully converted into outputs (e.g. saleable products manufactured per day or usable product designs completed per month), and represents the "real work" of the organization. *Net Process Throughput* is determined by three variables: *Gross Process Throughput*, the total quantity of new work accomplished (widgets per day or product designs per month); *Defect Introduction*, the quantity of new work that is not usable because it was done incorrectly (defective widgets per day, flawed designs per month); and *Defect Correction*, the quantity of previously defective work that has received additional attention and is now usable. *Defect* is used as a generic term for any undesirable outcome of a conversion process (Schneiderman 1988). For example, a product produced correctly, but delivered late, is defective if timely delivery is a desired attribute of the conversion process. Figure 1 shows the basic physical relationship between these variables in the form of a causal diagram (Forrester 1961, Weick 1979, Richardson 1991, Sterman 2000).

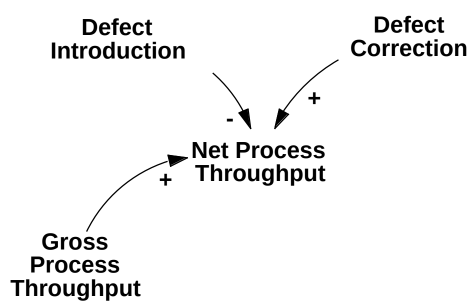

Figure 1. Net Throughput for any process is Gross Throughput less the rate of Defect Introduction plus correction of any previously defective work. Arrows indicate the direction of causality. Signs ('+' or '-') at arrow heads indicate the polarity of relationships: a '+' denotes that an increase in the independent variable causes the dependent variable to increase, ceteris paribus (and a decrease causes a decrease). That is, $X \xrightarrow{+} Y \Leftrightarrow \partial Y/\partial X > 0$. Similarly, '-' indicates that an increase in the independent variable causes the dependent variable to decrease; that is, $X \xrightarrow{-} Y \Leftrightarrow \partial Y/\partial X < 0$. See Sterman 2000. In this case Net Process Throughput = Gross Process Throughput – Defect Introduction + Defect Correction.

An increase (decrease) in gross throughput or defect correction causes an increase (decrease) in net throughput (ceteris paribus). Similarly, an increase (decrease) in defect introduction, ceteris paribus, causes a decrease (increase) in net throughput. Causal diagrams provide a compact and precise representation of interdependencies and are useful in describing the feedback structure of systems.[1]

Having divided the field of activity into three possibilities, we turn to the physical and decision making structures that link these options. The most basic linkage is arises from the fact that the rate of defect correction is ultimately constrained by the rate of defect introduction (you can't repair defective pieces that were never produced). To capture this connection, we introduce the level variable *Defects* connecting the *Defect Introduction* rate and the *Defect Correction* rate. A level variable (or stock) is the integration (accumulation) of its inflows less its outflows, denoted by the straight arrows with valves. Unlike a flow, which represents action or activity in a system, a stock represents the accumulated impact of those activities. Stocks are critical in creating the dynamics of systems: stocks characterize the states of the system which give rise to the information upon which decision are based; these decisions then alter the rates of flow which drive the stocks, thus closing the feedback loops in the system (Forrester 1961).

Capturing the linkage between the other elements is more complicated. In particular, what determines the rate of defect introduction? In manufacturing, defects are often created by physical features of the machinery (e.g. a dull cutting tool), and will continue to be produced until machines are stopped and the defect-causing elements eliminated. To capture this notion of permanence, we draw on a fundamental contribution of the founders of the quality movement who recognized the distinction between correcting defects that have already been produced and preventing them from occurring (Deming 1986). We label the causes of defects *Process Problems* (also known as root causes in the quality literature, Ishikawa 1985). Process problems

---

[1] Causal loop diagrams are not intended to provide mathematical specification of the relationships, which may be linear or non-linear, or of any time delays between cause and effect. Specifying a formal mathematical model is often the next step in testing the theories embodied in causal diagrams. For examples of formal feedback models of process improvement programs, see Repenning (2000b) and Sterman *et al*. (1997).

are the features of the process, either physical or behavioral, that generate defects. The stock of process problems determines the Defect Introduction rate.

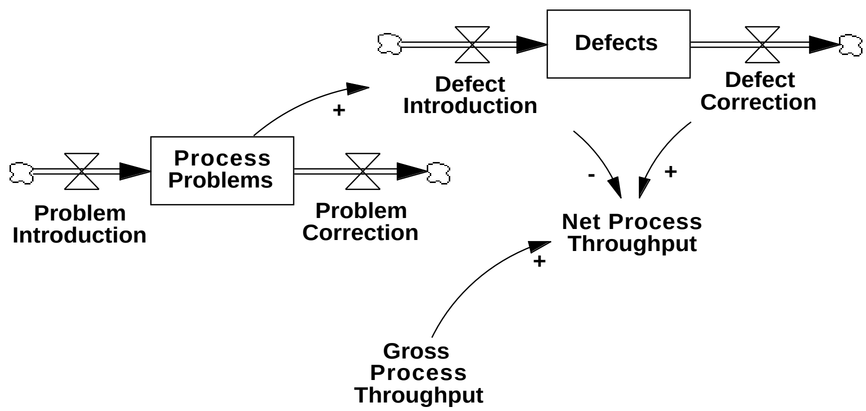

Figure 2. The stock and flow structure of defects and process problems. Boxes represent Stocks (or level variables) are represented by boxes while arrows with "valves" represent flows (or rate variables) are represented by arrows with "valves". A stock is the accumulation of the difference between its inflows and outflows. Formally, Defects (t) = $\int_t$ [Defect Introduction(s) – Defect Correction(s)] ds + Defects (t_0).

The stock of process problems is increased by *Problem Introduction* and reduced by *Problem Correction*. Process problems arise as equipment ages and wears, and as changes in products, processes or customer requirements create conflicts with existing procedures, skills, and equipment. Process improvement is the result of learning that leads to problem correction. Improvement activities or the adoption of new tools results in problem correction, reducing the stock of process problems, decreasing the defect introduction rate, and, ultimately, improving net process throughput.

Explicitly representing the key stocks in the system leads to a number of important insights into the dynamics of process improvement. Foremost among these is the distinction between defect correction and defect prevention and the value of engaging in improvement activities. Notice that the stock of process problems determines the flow of defects. One process problem creates a continual inflow of defects, forever reducing net process throughput unless each and every defect is corrected. Once a process problem is corrected, however, the stream of defect introduction is forever reduced. The challenge of process improvement is to shift attention from reducing the stock of defects to reducing the stock of process problems.

Responding to Throughput Pressure

We turn now to the decisions made within the physical structure of the system. Integrating the stock and flow structure with the behavioral processes governing the flows closes the feedback loops that determine the system's dynamics (Forrester 1961, Weick 1979, Richardson 1991). Consider the processes through which managers regulate throughput. Managers determine the adequacy of current throughput by comparing it to Desired Throughput and assessing the Throughput Gap (Figure 3). The demand for the organization's products or services (which we take as exogenous for the purpose of our analysis) determines desired throughput.

First-order Improvement

When faced with a throughput shortfall, workers and managers have several options to close the gap. Figure 3 shows three: expand capacity, utilize existing capacity more intensely, or repair defective output. Each option forms a negative or balancing feedback loop with the goal of eliminating the throughput gap by raising net process throughput towards the desired rate. First, managers can expand production capacity by hiring more workers and purchasing additional plant and equipment, boosting gross process throughput through the balancing *Capacity Expansion* loop (B0). However, expanding capacity takes time, is costly, and is generally not an option for managers responsible for day-to-day operations. We treat the capital stock and work force as exogenous since these decisions were beyond the authority of the participants in the improvement programs we studied.[2]

---

[2] . For models of the interactions between process improvement and capacity see Sterman *et al.* (1997) and Repenning (2000a).

Second, to increase net process throughput workers can *Work Harder* (balancing loop B1), increasing the utilization of existing resources. Effort can be increased through greater focus on task, shorter breaks, reduced absenteeism, and overtime. In the pre-MCT period, line supervisors primarily relied on the work harder loop to achieve their production objectives. Third, managers can allocate resources to correct existing defects (the balancing *Rework* loop B2). In the product development organization, the rework loop was used extensively to complete projects on time.

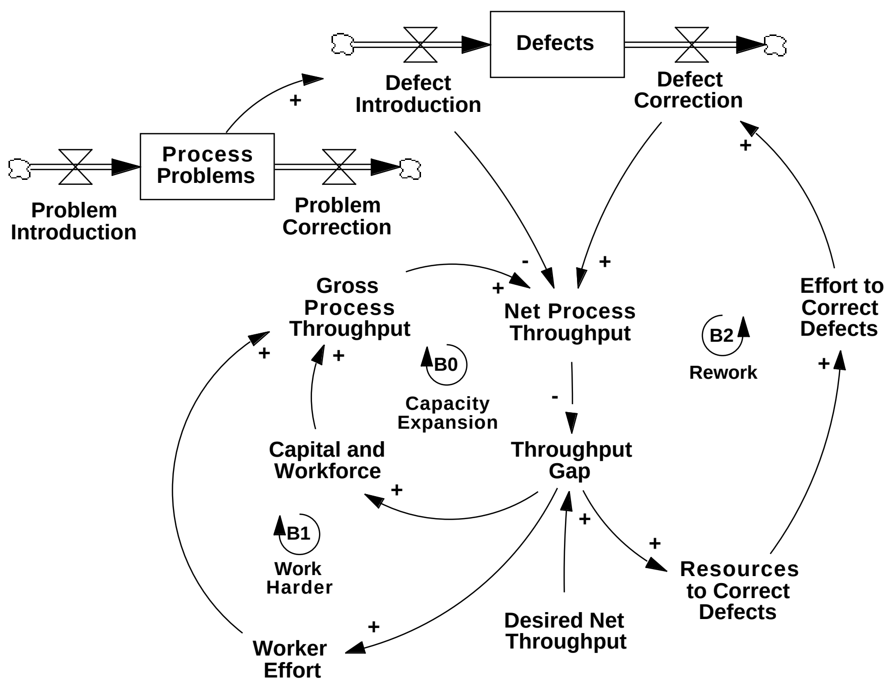

Figure 3. Negative feedbacks controlling throughput. The loop identifiers (e.g. B1) indicate whether a loop is a negative (balancing) feedback or a positive (self-reinforcing) feedback. See Richardson and Pugh 1981 or Sterman 2000.

Second-order Improvement

Each of the first-order improvement loops can close the throughput gap, but only at significant and recurring cost. A more effective solution is to eliminate the process problems that generate defects (Deming 1986). Such "second-order" improvements create the negative *Work Smarter* loop B3 (Figure 4), which closes the throughput gap by eliminating process problems. Making such fundamental improvements requires managers to train their work force in improvement techniques, release those workers from their normal responsibilities, and, most importantly, give them the freedom to deviate from established routines so they may experiment with potential solutions (Deming 1986, Wruck and Jensen 1994).

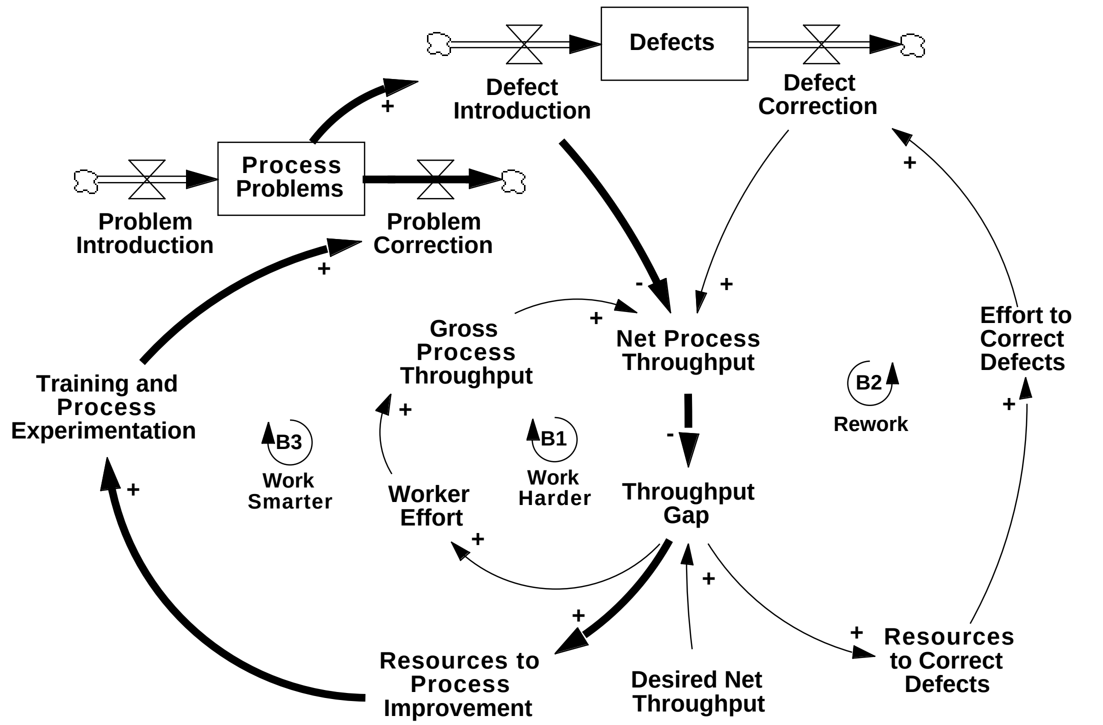

Figure 4. Second-order improvement: Investing in improvement activity creates the negative *Work Smarter* loop (B3) which enhances net throughput by reducing process problems.

Interactions of Physical Structure and Decision Making at the Individual Level

We now use our model to understand how the system evolved to the undesirable state described in the previous section, first pursuing this question at an individual level of analysis and then extending our model to include both managers and workers. We begin with the question of how participants choose to close a persistent throughput gap.

Cognitive and Perceptual Biases Against Fundamental Improvement

The high leverage point for increasing net process throughput is allocating more effort to reducing the stock of process problems. There are, however, at least four reasons, rooted in basic perceptual and cognitive processes, why activities targeted at gross process throughput and defect correction often take precedence over defect prevention. First, the output of a process (the actual products that are produced) is more salient and tangible than process problems. People have repeatedly been shown to over-weight available and salient features of the environment (Kahneman, Slovic and Tversky 1982; Taylor and Fiske 1975). In the manufacturing environment, for example, defective products are physical objects, easily seen and touched. They accumulate in a pile sitting somewhere on the production floor. They are literally in the way. In contrast, process problems are often invisible. Processes consist of the activities and relationships that create tangible products, and cannot be easily discerned from the products themselves (Orlikowski 1995). Indeed, many quality improvement tools are designed to ferret out root causes from observations of the defects they create. The situation is similar in product development. A project, particularly when it is close to its launch date, is very tangible and salient. In contrast, a project in its early phases is just a concept or an idea and has few, if any, physical manifestations. A chief engineer makes this point when describing his organization's unwillingness to allocate resources to the early phases of a development project as the PDP process dictated:

> ...as an example, if you are engineer with a project close to job one, say 3 months, and for whatever reason you are not ready to ship your product, that's a very visible and apparent problem. There is no question the project is in trouble and needs attention. Now, [suppose] you are thirty months ahead of job one, 2.5 years away, and you are lacking customer definition. You are trying to get the same level of attention, "I've got to have this or I can't move ahead." There is much more of a tendency to say "Come on, quit crying and get on with it."

Second, defect correction and process improvement work at different speeds. Process improvement requires time to document the current process, diagnose root causes, experiment with possible changes, implement solutions, train participants in the new procedures, and so on. The delays between the start of an improvement program and results are long and variable, ranging from months to years, depending on the complexity of the process (Schneiderman 1988). Defects, however, are often easily identified and quickly repaired. Choosing to eliminate process problems often entails a short-term reduction in throughput as resources are reallocated from throughput and defect correction to improvement. Faced with this worse-before-better trade off, managers and workers under pressure to close a throughput gap are likely to choose correction over prevention even if they understand that doing so suppresses the symptoms without curing the disease. The executive in charge of PDP discusses the problem created by the long time delays inherent in successful process improvement:

> Imagine at the end of the year the general manager…going up in front of the president and saying, "We missed our profitability numbers because we spent extra money developing our new design process that won't be fully deployed and rolled out till five years from now but wasn't that a good move?"

Third, first-order production efforts have a more certain outcome than second-order prevention efforts. It is usually clear when a defect has been corrected. Similarly, it is relatively easy to assess the benefit of an extra hour dedicated to increasing gross process throughput. In contrast, process problems are more complex and their characterization more ambiguous. It is often unclear whether and how a proposed process change will result in fewer defects. Risk aversion is a basic feature of human decision making, and people have also been shown to be ambiguity averse (Einhorn and Hogarth 1985). Faced with a throughput gap, most managers prefer the more certain gain of production and correction efforts to the ambiguous, uncertain and delayed yield of an investment in prevention. The executive in charge of the PDP effort also alludes to this dilemma:

> ...taking the risk and spending incremental money that’s not in your budget and taking the hit for over running your budget– even though in the long run it would have been the right thing to do–is a difficult thing to do...the budget is something that’s easy for your boss to tell you whether or you hit it or not... anybody can hold it up in your face.

Fourth, eliminating process problems, while preventing future defects, does nothing to eliminate the stock of defects already generated. The stock of defective outputs represents a substantial and tangible investment in materials, labor and capital. Most accounting systems report the value of the inputs to each product, making it is easy to assess the benefit of investing in correction: if the value of a repaired product is $y and its scrap value is only $x, it is worth investing anything up to $y-x to correct the defect. In contrast, assessing the value of defect prevention is more difficult. As one manager in our study said, "...nobody ever gets credit for fixing problems that never happened." The well-known sunk cost fallacy (Arkes and Blumer 1985, Staw 1976, 1981, Thaler 1980) reinforces the bias towards correction. Decision makers often continue a project beyond the economically rational point when they have already made a substantial investment of time, money and emotion. Here, the sunk cost fallacy means managers often favor defect correction rather than defect prevention, to, as they see it, recoup past investments in defective outputs, even though these investments are sunk costs.

Linkages between First- and Second-order Improvement

The discussion so far suggests that, at the level of the individual, the physical structure of most production processes leads to significant biases against improvement activities. The situation is, however, more complicated than the relatively simple task environments found in most laboratory decision making experiments because first and second-order improvement processes can be strongly coupled. Interconnections arise for two reasons. First, resources are finite. Line workers have limited time, which they must allocate among production, defect correction, and process improvement. Similarly, engineers face a trade-off between doing new designs, correcting old ones, and investing in improved process capability. Managerial attention is also limited and must be allocated to competing activities (March and Simon 1958/1993). Improvement activities require management time to motivate employees, guide training, review results, and mediate conflicts. Process-oriented improvement programs often cut across traditional organizational boundaries, intensifying demands for senior management attention.

Resource constraints coupled with the previously discussed biases against improvement lead to two negative links: as *Worker Effort* rises, *Training and Process Experimentation* suffer. Likewise, *Resources to Process Improvement* fall when people increase *Resources to Defect Correction* (Figure 5).

The second interconnection arises because learning and improvement are not free. Instead, while such activities have a permanent, long run, benefit, they also typically come at the cost of a temporary reduction in throughput. In manufacturing, machines must usually be taken off-line to conduct experiments, many of which fail. The time spent planning, doing, and interpreting these experiments cannot be devoted to production. Similarly, when engineers attend training, practice with new development tools, or experiment with new technologies they are not completing designs or managing other aspects of a given project. We capture these short-run costs with a negative link from *Training and Process Experimentation* to *Gross Process Throughput*. The strength of this link depends on the available slack. If experiments can be run when machines are normally idle, and engineers engage in improvement activity when their services are not required on development projects, then the link is weak and the marginal cost of improvement is low.

The new links close three important feedback loops. The first is the balancing *Focus on Throughput* loop (B4). Process participants can close the throughput gap, at least in the short run, not only by *Working Harder* (B1) and by doing more *Rework* (B2), but also by *Focusing on Throughput* (B4), that is, by reducing the time spent on process improvement, creating another balancing feedback which helps operators and engineers reach their goals. Of course, any gains in throughput made by *Focusing on Throughput* (B4) are temporary. As participants allocate more effort towards work and away from improvement, net process throughput will initially increase, but then later decline as process problems accumulate.

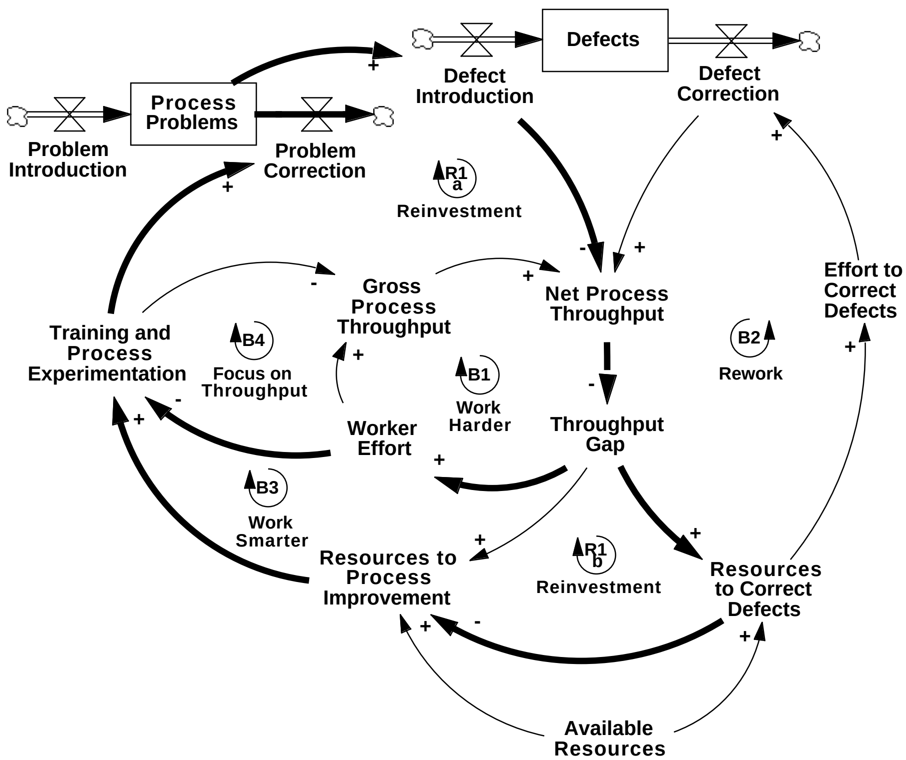

Figure 5. The reinforcing feedback created by finite resources: as more time is devoted to defect correction, less is available to correct process problems, leading to still more defects and still less time for improvement. Note also the balancing Focus on Throughput loop: workers can meet throughput goals, at least in the short turn, by cutting back on improvement activity.

The second and third loops are the self-reinforcing *Reinvestment* loops R1a and R1b (Figure 5). Unlike the loops described so far, the *Reinvestment* loops are positive feedbacks that tend to reinforce whichever behavior currently dominates. Successful improvement increases net process throughput by reducing defect generation. As the throughput gap falls, workers have more time to devote to training and experimentation, leading to still more improvement (loop R1a). Similarly, if the organization succeeds in reducing defect generation, less time and effort are needed for gross process throughput and defect correction, freeing resources for fundamental improvement, speeding the elimination of process problems, and driving defects down still further (loop R1b). Here the loops operate as virtuous cycles. Conversely, if defects increase, worker effort rises and more resources are allocated to defect correction. Improvement effort falls, process problems accumulate at a faster rate, leading to still more defects, and the reinvestment loops operate as vicious cycles. For example, deferring preventive maintenance to repair unexpected equipment breakdowns can lead to more breakdowns and still greater pressure to reassign maintenance mechanics from preventive to reactive work (Carroll et al. 1997). Similarly, in product development, allocating resources to fix problems late in a development cycle reduces the resources available to projects in earlier phases, leading to future problems and still fewer resources for new projects (see Repenning 2000b for a formal model).

The Capability Trap

The feedback processes created by the interactions between judgmental biases and the physical structure creates a phenomenon that we call the *Capability Trap*. Consider a machine operator or a design engineer faced with a shortfall. Using the improvement offset loop, she can close the throughput gap by reducing the time she spends on learning activities and reallocating that time to work effort. Such behavior is tempting since it has an immediate positive impact and the total amount of time dedicated to work does not change. Unfortunately, while such actions improve throughput in the short run, the reduction in time dedicated to learning causes process capability to decline. Soon our hypothetical worker finds herself again falling short of her throughput target, forcing a further shift towards working and away from improving, thereby trapping her in a downward spiral of declining process capability, increasing work hours, and little time for fundamental improvement.

The capability trap played an important role in both the pre-MCT and the PDP efforts. For example, a supervisor describes the state of manufacturing prior to MCT:

> ...supervisors never had time to make improvements or do preventive maintenance on their lines...they had to spend all their time just trying to keep the line going, but this meant it was always in a state of flux, which in turn, caused them to want to hold lots of protective inventory, because everything was so unpredictable. It was a kind of snowball effect that just kept getting worse.

In this case, supervisors rely on the *Improvement Offset* loop to hit their throughput objectives by "…spending all their time keeping the line going." The lack of attention to improvement causes the *Reinvestment* loops to operate as vicious cycles thereby trapping the line at a minimal level of capability and preventing supervisors from engaging in improvement activities.

Similarly, in product development, the capability trap prevented the organization from developing new processes that would have increased productivity. Consider the efforts to implement the bookshelf, as described by an engineering manager:

> An engineer might not take the time to document her steps or put the results of a simulation on the bookshelf and because of that she saved engineering time and did her project more efficiently. But in the long run it prevented us from being able to deploy the reusability concepts that we were looking for.

Thus, just as machine operators faced a basic trade-off between producing and improving, development engineers were forced to trade off completing their assigned tasks against documenting what they learned so that others might benefit. Engineers could make more rapid progress towards their objectives by using the *Focus on Throughput Loop* and forgoing the bookshelf, but doing so prevented them from initiating the self-reinforcing reinvestment loops that would have led to improved process capability. And, whereas the positive loop trapped the manufacturing operation in a state of low process capability, so too it prevented the PD organization from realizing potential productivity gains. Figure 6 shows both situations represented in our framework.

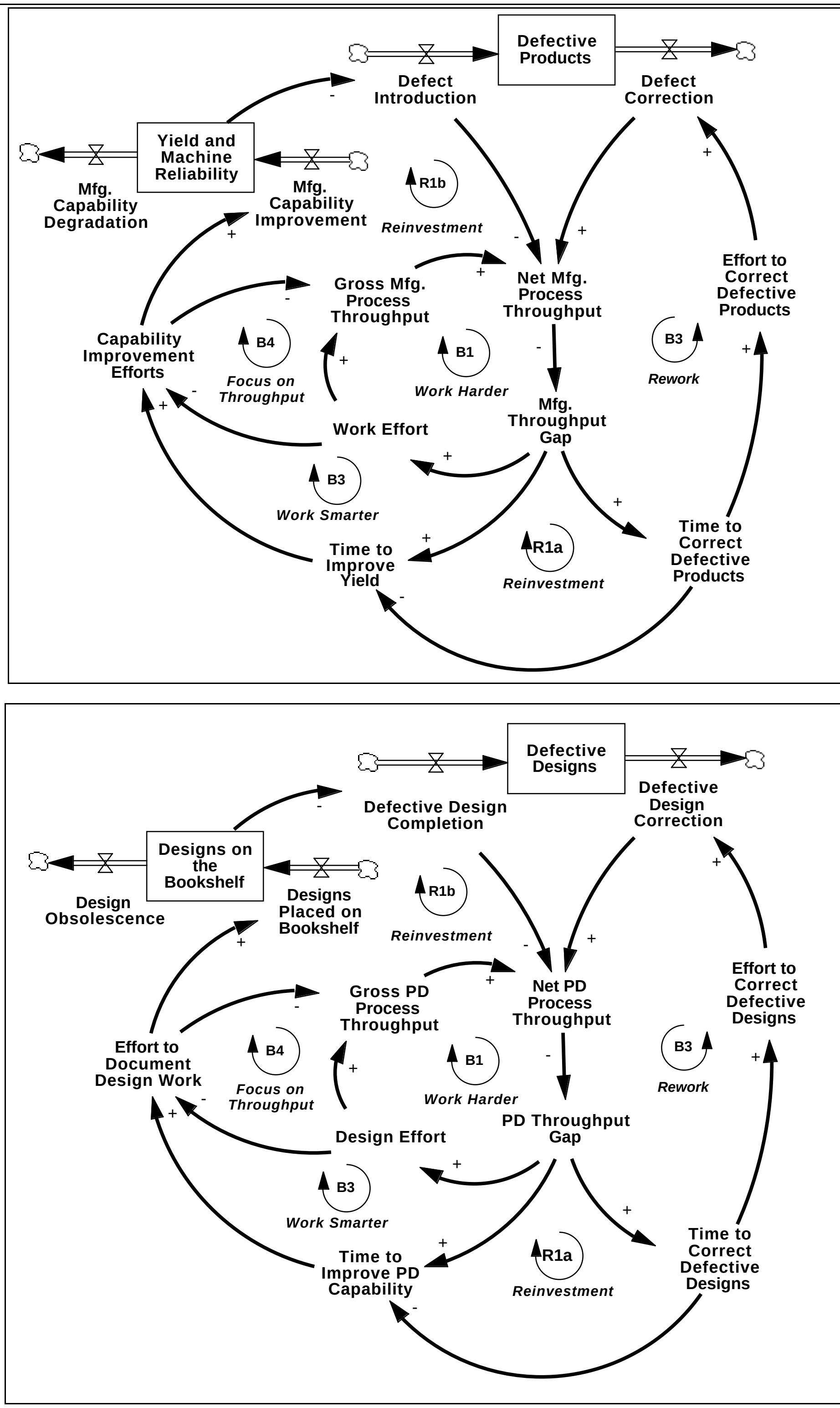

Figure 6. The Capability Trap in Manufacturing and Product Development.

Misperceptions of Feedback

There are only two ways out of the capability trap: sacrifice short-term throughput objectives in favor of investments in process capability or add additional resources for improvement. While the existing cognitive literature strongly supports the behavioral linkages in the model, wouldn't participants quickly learn to overcome these dynamics by making the appropriate short-term sacrifice? Studies of human performance in dynamic decision making tasks suggest that the answer to this question is an emphatic no. Consider the outcome feedback received from a decision to spend more time working and less time on improvement. Performance quickly increases, producing a clear, salient, unambiguous outcome. In contrast, the negative consequences of this action—the decline in process capability—happen with a delay, are hard to observe, and may have ambiguous interpretations. In experiments ranging from running a simulated production and distribution system (Sterman 1989b) to fighting a simulated forest fire (Brehmer 1992) to managing a simulated fishery (Moxnes 1999), subjects have been shown to grossly overweight the short-run positive benefits of their decisions while ignoring the long-run, negative consequences. Participants in these experiments produce wildly oscillating production rates and inventory levels, allow their fire-fighting headquarters to burn down, and find their fleets idled after overexploiting their fisheries. Such "misperceptions of feedback" have been repeatedly observed in a variety of systems with even modest levels of dynamic complexity (Sterman 1994). These and other studies also show that learning in such systems is very slow (Paich and Sterman 1993), and suggest that once caught in a set of pathological reinforcing dynamics, people tend to persist in their chosen course of action rather than revisit the assumptions that led to it (Staw 1976, 1981). Thus, once caught in the capability trap, individual workers and managers are unlikely to learn to overcome it.

Interactions of Physical Structure and Decision Making at the Group Level

Differences in information availability, salience and time delays can bias individuals against fundamental improvement, and the tight coupling between first and second-order activities can create a vicious cycle of declining learning and intense work pressure. The dynamic is prevalent in numerous situations ranging from the product development and manufacturing settings discussed here to junior faculty members who never take the time to learn productivity-improving software because they feel pressure to produce additional research papers. Process improvement is, however, not an individual activity. Those assessing the throughput gap and giving directives are typically managers, while those actually allocating their time between working and improving are engineers and machine operators. The addition of these organizational and group dynamics intensifies the bias towards working harder.

Biased Attributions About the Causes of Low Throughput

Extending the analysis to the organizational context raises additional complications because managers influence both the total *quantity* of time spent on work and its *allocation* between the first and second-order activities. Thus, when choosing to pursue first or second-order improvement managers must make a judgment about the causes of low process throughput. If managers believe the cause of low performance lies in the physical structure of the process, they are likely to focus their efforts on process improvement. If, however, low throughput is thought to result from lack of worker effort or discipline, then managers are better off focusing their attention on increasing the total quantity of work.

How do managers make such attributions? The cues people use to make causal attributions include temporal order, covariation, and contiguity in time and space (Einhorn and Hogarth 1986). Attributing low throughput to inadequate worker effort is consistent with all these cues: worker effort immediately precedes the production of an item, production is highly correlated with worker effort, and workers and the items they produce are highly contiguous in time and space. In contrast, process problems typically precede low throughput with much longer and often unobservable delays, the correlation between process problems and low throughput is frequently imperfect or unobservable, and process problems can be far removed in time and space from the detection of the defects they create. Thus, managers are likely to attribute a throughput shortfall to the attitudes and dispositions of the workers even when the true causes are systemic features of the environment such as process problems. Many studies show that attributing the cause of a problem or behavior to individuals rather than the systems in which they are embedded is a pervasive and robust phenomenon – the so-called "fundamental attribution error" (Ross 1977). Thus, while the individual-level analysis indicates a bias against fundamental improvement, extending the model to include both managers and workers suggests the problem becomes worse because managers are not simply assessing the relative benefits of both activities; they are also assessing the attitudes and character of the workers they direct.

Attribution Errors and The Capability Trap

To understand the dynamic implications of this bias, consider what happens if managers conclude that the work force is underutilized and/or undisciplined (thus attributing low performance to people rather than process problems). If they reach such a conclusion, then the intendedly rational response is to *Squeeze out Slack* by increasing *Production Pressure and Worker Control* (loop B5 in Figure 6). The construct Production Pressure and Worker Control aggregates a number of behaviors observed in our field studies. Production pressure includes increasing the penalties assessed for failing to achieve objectives, as in the pre-MCT manufacturing setting when "…supervisors who missed their objectives knew they were going to get beat up by their managers" and in product development where "…they shoot you for missing product launch." More subtly, it also entails changes such as increasing the frequency and granularity with which measurements are reported. For example, prior to the MCT effort, the measurement system evolved to the point where utilization was often reported on a per machine, per shift basis. Similarly, a project manager we interviewed recalled that when a subsystem for which he was responsible fell behind schedule his boss required him to call in every hour with a status report until the prototype met the specifications. Finally, production pressure and worker control also include even more direct attempts to restrict the activities of the workforce such as documentation standards that require engineers to justify their activities on a daily basis.

What happens if managers observe a throughput gap, make an Attribution of Low Effort and consequently increase Production Pressure and Worker Control? Figure 7 suggests two immediate impacts. First, worker effort and gross process throughput immediately rise, thereby closing the throughput gap via the *Squeeze Out Slack* loop (B5) and generating the outcome the manager intended, an increase in the *quantity* of time dedicated to work. Second, as slack declines, workers are less able to achieve their objectives by simply increasing the time they spend working. To continue to achieve their ever-increasing objectives, they may be forced to reduce the time spent on improvement activities. As Time Spent Training and Improving declines, gross process throughput also rises (at least in the short run), creating another balancing feedback. In contrast to the *Work Harder* loop, this loop increases throughput by shifting the allocation of time from improvement to production.

But, in contrast to the *Squeeze out Slack* loop, the *Focus on Throughput* loops (B4a and B4b), while having the desired effect in the short run, yield a long-run side effect. With less effort dedicated to process improvement, fewer process problems are corrected, the defect introduction rate rises, and net process throughput falls, thereby offsetting the initial gains. Paradoxically, by continually increasing throughput objectives in the pursuit of improved system performance, managers who mistakenly attribute low performance to the attitudes and dispositions of their workforce, inadvertently force the system into the capability trap, dramatically limiting success.

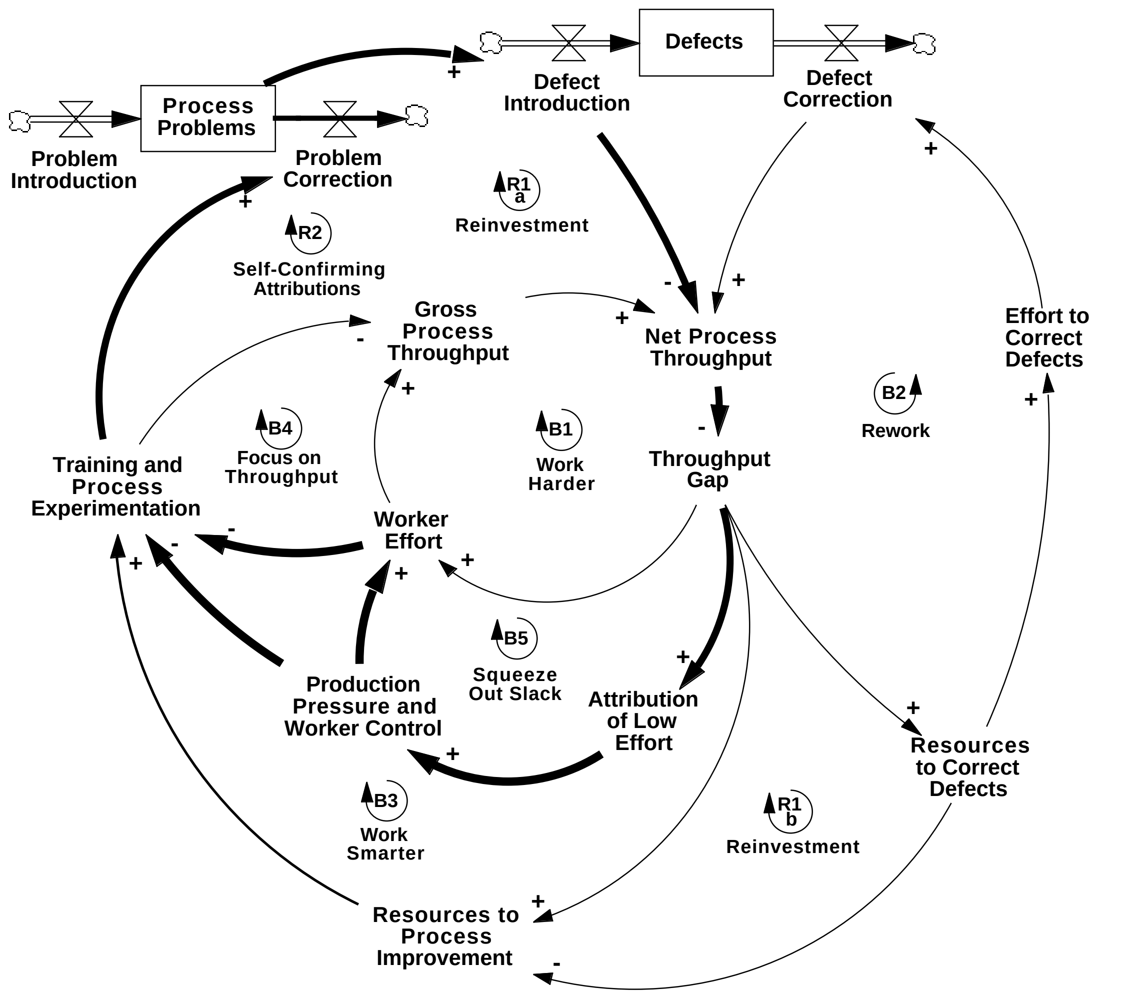

Figure 7. Managers who attribute the throughput gap to worker shirking will increase production pressure and monitoring in an attempt to *Squeeze out Slack* (B5). Throughput rises in the short run, but production pressure pulls resources away from improvement activity, leading to erosion of process capability and still lower throughput (reinforcing loop R2).

The Self-Confirming Attribution Error

As workers spend more and more of their time on throughput and cut back on fundamental improvement, wouldn't managers realize the true cause of low process throughput was inadequate process capability rather than lazy employees? We previously discussed how, at the individual level, people receive feedback that is likely to increase rather than offset the bias towards working harder. These dynamics are intensified work when managers, not workers, assess the benefits of increased production pressure.

In most settings, managers cannot observe all the activities of the workers. Hence, after applying production pressure, they cannot easily determine how much of the additional throughput is due to increased work effort and how much results from cutting back on training, improvement or maintenance. Suppose, for example, that there is a throughput gap requiring six hours of additional productive effort per person per week. Managers, believing employees are simply not working hard enough, increase production pressure and monitoring. Workers will focus their activities on production, cutting their breaks and other nonproductive time. Suppose these responses yield only two hours per person per week in effective work effort. To close the remaining throughput gap, workers may gradually reduce the time they spend on process improvement, training, and experimentation until they free up the needed four hours per week. Managers observe that throughput rises by the equivalent of six hours of productive effort.

Because managers do not fully observe the reduction in training, experimentation, and improvement effort (they fail to account for the *Focus on Throughput* loop), they overestimate the impact of their get-tough policy; in our example by as much as a factor of three. To the extent managers are unaware of how workers reallocate their time to meet their goals, the throughput gains resulting from production pressure provide powerful evidence confirming the managers' suspicions that workers were not giving their full effort. Managers quickly learn that boosting production pressure *works*; throughput rises when they turn up the pressure.

These links create the *Self-Confirming Attribution* loop (R2). Like the self-reinforcing *Reinvestment* loops discussed earlier, the *Self-Confirming Attribution* loop can drive the organization to higher levels of production pressure with fewer resources dedicated to process improvement. Far more importantly, however, it also changes the mental models of the managers by providing them with increasingly compelling evidence that the source of low throughput can be found in the poor attitudes and weak character of the workforce. While individuals may get caught in the capability trap due to their inability to anticipate complex dynamics, they are likely to be aware of the changing allocation of their own time. They must seek some resolution to the dissonance created by their declining performance (e.g., I didn't get enough support from management, the training was poor, this program is a fad, etc.). In contrast, managers, at least partially ignorant about both the quantity and allocation of the workers' effort, may easily confound the two and attribute low performance to low work quantity rather than poor allocation.

Unlike the individual case where people must seek out explanations for their low performance, management pressure to elicit full effort from slothful workers generates powerful evidence to reinforce managers' initial, but incorrect, attribution that the workers just need a kick in the pants. No additional resolution is needed. Recall the project manager discussed above who was required to provide hourly status reports on a balky prototype: Soon afterward the problem was solved, confirming the boss's belief that he had acted appropriately, indeed had decisively taken charge of the situation, even though the team was already working around the clock and his interference drained precious time from their efforts to solve the problem.

Even more insidiously, the long run effects of production pressure also reinforce managers' belief that workers are the problem. The delay between increased production pressure and worker control and increased throughput (via the *Work Harder*, *Focus on Throughput*, and *Squeeze Out Slack* loops) is much shorter than the time required to detect the resulting erosion in process capability as the reinforcing *Reinvestment* and *Self-Confirming Attribution* loops lead to more process problems, lower throughput, more shortcuts and less improvement. The erosion of process capability caused by production pressure is delayed, gradual and diffuse. It is distant in time and space from its cause. Managers are unlikely to attribute the cause of a throughput gap to the pressure they placed on workers months or even years before. They are likely instead to conclude that the workers have once more become lazy and require another increase in production pressure.

Note how workers may unwittingly conspire in strengthening managers' attributions. Faced with intense production pressure, workers are naturally reluctant to tell supervisors they can't meet all their objectives. The more effectively workers are able to cover up the process shortcuts they take to meet their throughput targets (loops B4a and B4b), the less aware managers will be of the long run costs of production pressure. Unaware that improvement activity, maintenance, and problem solving have been cut back, throughput appears to rise without requiring any sacrifices, reinforcing management's attribution that the workers really were lazy.

Evidence of these self-confirming attributions appears in two different guises in our data. As discussed earlier, senior managers in the PDP effort repeatedly attributed the difficulties they experienced to the undisciplined engineering staff. When pressed to explain further, many managers resorted to even deeper attributions. For example, when asked to explain why engineers were so resistant to using project management, the manager in charge of the initiative replied:

> Program management and the disciplines associated with it continue to be a problem in my opinion in most western cultures. The people that are particularly rigorous and disciplined, the Japanese and the Germans, tend to be so by cultural norms. I can't tell you if it's hereditary or society or where it is they get it but the best engineers are those that tend to be the most disciplined, not as individual contributors but as team based engineers. So there's a strong push back from the western type of engineer for much of this.

Such attributions, here generalized to entire nations and ethnic groups, are typical of the fundamental attribution error. As these attributions are shared and repeated they become institutionalized. They become part of the corporate culture, and, as suggested by the quote above, can strengthen widely held stereotypes and prejudices in society at large.

In the MCT effort, the data also supported our core hypothesis. Whereas in the PDP effort these attributions appeared in explanations for failure, in the MCT effort, many managers attributed the success of the effort to *overcoming* their tendency to blame workers. One supervisor, when asked about the success of the MCT effort, articulated this point clearly:

> There are two theories. One says "there's a problem let's fix it." The other says "we have a problem, someone is screwing up, let's go beat them up." To make improvement we could no longer embrace the second theory, we had to use the first.

Similarly, as the end of the first round of interviews approached, the first author was given the opportunity to interview the general manager responsible for launching the MCT effort. Among the many things discussed, the following question was posed:

INT: How would you define leadership? What does the leader do?

VP: Sets the vision; sets the direction; makes the tough calls; does what's right for the business and the people, not an either/or; is motivational when he has to be or she has to be; sticks his neck out and protects his people; values -- thinks the world of people and their capability. If you don't do that you're in deep trouble.…99% of the people want to do well and have the capability and come to work really wanting to contribute and use their capabilities.

Following this, in response to a subsequent question, the GM continued by explaining the experience through which he had come to this view:

> I'll tell you where I got it… [At a previous job] I became a plant manager… I'll never forget as long as I live… [The previous plant manager] started blaming his people for all his problems... He was really on one guy very badly... His view was the whole reason for the poor performance of the plant was because of this guy. So I didn't say anything or do anything and I took my time and I gave it about two or three months. Then I gave the guy more responsibility…as much responsibility as he'd take. He ended up being one of the best people in the plant. I guess that was probably the turning point…giving people more responsibility, showing you have confidence in them, having a view that they're going to succeed and not fail, trusting them.

Thus, while many managers in the PDP effort attributed the initiative's failure to the engineers working within the process, managers in the successful MCT effort often attributed its success to *changing* their own attributions, from people to the production system.

Adaptation, Ad Hoc Changes, and Structuration

The model outlined so far suggests that an organization can evolve to the point where workers are under high degrees of pressure to achieve their throughput objectives while managers attempt to control their activities in some detail. Our data suggest that eventually these dynamics lead to a set of mutually incompatible objectives. Caught between ever-higher throughput goals and the need to comply with stricter controls, workers must make situation-specific adaptations to appear as if they are meeting all of their objectives. Such adaptations include manipulating measurement systems, short cuts and ad hoc, undocumented, and sometimes even surreptitious changes to the process.

The organizational literature contains many examples, ranging from simple “work-arounds” on the manufacturing floor (Orlikowski and Tyre 1994) to changing the standards for O-ring tolerance on the space shuttle (Wynne 1988). Clearly, not all such changes are harmful. Pressure can sometimes spur a creative solution to vexing problems. But to the extent they face time pressure and multiple, incompatible objectives, workers will be tempted to erode standards, cut corners, fail to follow-up and resolve problems, and neglect to document their work. Even if creative workarounds solve the initial problem, they can create new ones when downstream processes are not updated to reflect the new upstream process. As shown in figure 8, such ad hoc changes increase the number of process problems.

Often, workers will keep their “work arounds” secret from management and manipulate metrics to appear to comply with objectives when, in fact, they are not. These links create two additional positive feedbacks, the *Process Integrity* and *Double Bind* loops (R3 and R4) which inadvertently erode production capacity by introducing new process problems as a side effect of management’s attempt to boost throughput. The irony of such a situation is that, while localized adaptations relying on situation specific, often tacit, knowledge are the hall mark of successful process improvement (Wruck and Jensen 1994), in the situation described here, the high throughput pressure and restrictive controls redirect this potentially productive activity towards circumventing management.

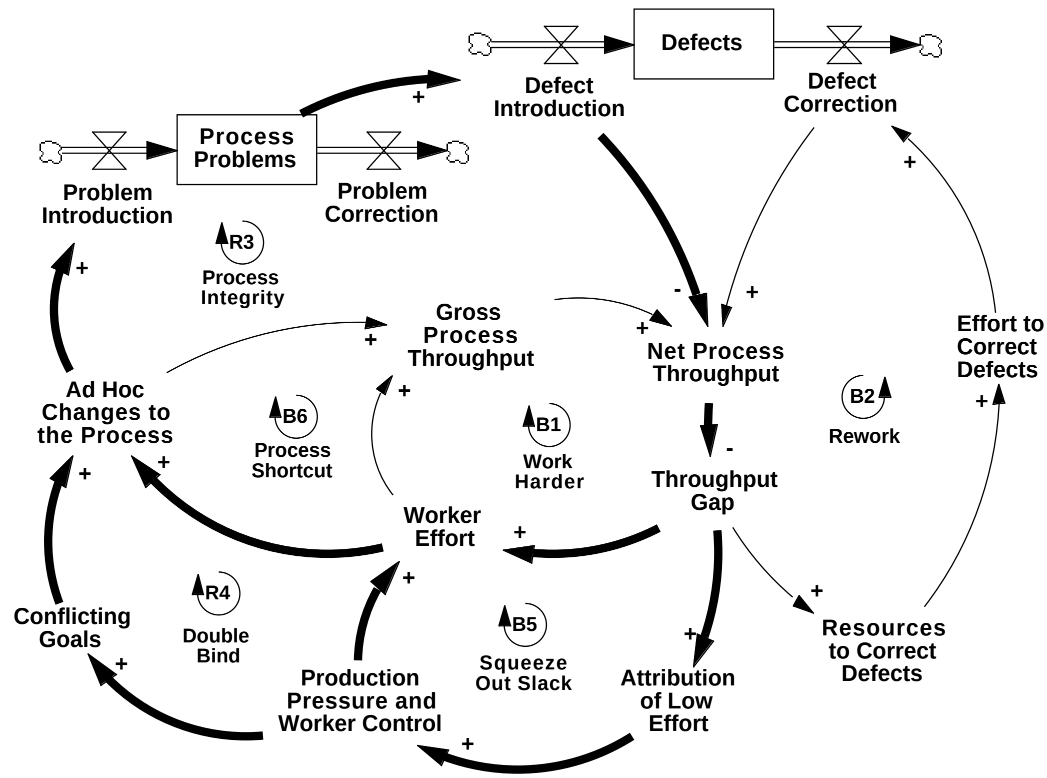

Figure 8. Production pressure and control over worker effort conflict, forcing workers to find "work arounds," eroding process integrity and leading to still more production pressure and still tighter controls (reinforcing loops R3 and R4).

These dynamics were particularly clear in the pre-MCT period in manufacturing. Prior to the improvement effort, manufacturing supervisors and operators worked under the increasingly constraining measurement system discussed above. Previous initiatives targeted at reducing WIP inventory created a direct conflict with the objectives of high machine and labor utilization. Operators and supervisors reacted by making ad hoc changes to the manufacturing process that allowed them to appear to satisfy both objectives. Many surreptitiously accumulated secret work-in-process inventories so they could keep their machines running, even if the outputs weren't needed. A manager explains:

> Supervisors at that time were evaluated on labor performance on a daily basis. It didn't take long for them to develop a buffer in front of their line so that if the schedule called for 700 and their line was fully utilized at 800, they could still run 800 units every day, and still make their labor performance.

Similarly, in product development the PDP process required development teams to pass a series of check-points, or stage-gates, and provided detailed reports on their progress. Engineers reported spending substantial time preparing for these reviews and reporting that steps were completed when, in fact, they were not. One engineer recalls his experience:

> ...we scrambled to get ready for the definition phase exit review, get all the paper work done and not miss the exit. We put on this show that says we exited this phase on time... Well 3 or 6 months later the customer is still making changes to the features of the product...so we are supposed to have, say, a year for the detailed design phase, and it ends up we have more like six months.

Ad hoc process changes strengthen the dynamics we've already discussed. As increased production pressure and ad hoc work-arounds inadvertently create new process problems, net throughput falls. Faced with a persistent throughput gap, managers may feel compelled to further increase production pressure and worker control. However, the stress of constant crisis, extended overtime, ever more aggressive throughput objectives and conflicting goals eventually causes fatigue and burnout among workers, lowering productivity and quality. For example, development engineers working on PDP pilot projects reported working extremely long hours. One explains how his group tried to hit its delivery deadline:

> How do we catch up? We stayed late. Most of the team was working from 7:00 a.m. to 8:00 p.m. and on weekends. A lot of people worked right through the Christmas vacation.

Absenteeism and turnover rise, eroding skills and lowering gross throughput still more. Workers grow to resent the control exerted by management and the lack of trust behind it, leading to an increasingly hostile and adversarial relationship between superiors and subordinates. Workers ultimately have no choice but to evade or subvert management's controls, play games with performance metrics, and shirk to relieve an intolerable workload. Once discovered, these "work arounds" provide powerful evidence confirming managers' initial assessments. Thus, what begins as a false attribution by management that workers are slothful, undisciplined, and untrustworthy becomes reality. Managers' fears are realized as a consequence of their own actions.

Over time the physical environment adapts to both reflect and perpetuate these self-reinforcing attributions. Managers who have come to believe that production pressure is an effective way to improve throughput will often resort to technology to further increase their control over the work force. Such technological solutions can take the form of time cards, detailed work reporting systems, video surveillance, or software that measures the keystroke rate of data entry operators. Workers develop increasingly sophisticated means of circumventing technological controls, which, when eventually discovered, further confirm managers' belief that the controls were necessary and, perhaps, even need to be augmented – another reinforcing feedback.

So it is that initially erroneous attributions about the capabilities and motives of the work force can soon become embedded in the routines, culture and even the physical structure of the organization, thereby perpetuating the cycle. Consistent with structuration view (Giddens 1984, Barley 1986, Orlikowski 1992), mental models, behavior, and the physical structure of the system mutually reinforce one another to generate organizational dynamics.

Discussion

Four main implications emerge from our analysis. First, and most fundamentally, it suggests that theorists should be careful to distinguish between efforts to exploit technology and efforts to exploit people working within that technology. March (1991:71) writes, "Exploitation includes things captured by terms such as refinement, choice, production, efficiency, selection, implementation, [and] execution." and later equates such activities with the "…refinement of existing technology." While widespread within the organization studies literature, this usage stands in stark contrast to that originally espoused by Marx (1887) and later used by sociologists like Perrow (1986). In this view, exploitation is largely targeted at appropriating people's time and energy, not at technological improvement and learning. For example, in his discussion of exploitation, Marx reported the following conversation with a mill owner concerning his attempt to circumvent regulations governing the length of the work day, "'If you allow me,' said a highly respectable master to me, 'to work [people] only ten minutes in the day over-time, you put one thousand a year in my pocket. Moments are the elements of profit.'" As this contrast suggests, we find at least two types of exploitation in our data: that directed at finding improvement opportunities latent within the organization's technology and that directed at appropriating peoples' time and energy.

Second, distinguishing between two senses of exploitation is important because they interact: structuring an organization to extract maximum effort from people can effectively prevent the exploitation of the production technology. As the MCT effort shows (and other advocates and theorists have argued, e.g. Deming 1986, Orlikowski 1996, Wruck and Jensen 1994), improving and refining a complex technology is a distributed activity requiring the active participation of numerous people. To be successful, participants must have considerable latitude in their ability to try new methods and experiment with new solutions. When managers focus their attention on appropriating peoples' time, however, they typically put severe limits on the range of acceptable activity and thus curtail the organization's ability to find and capitalize on opportunities to improve its technology. For example, before the MCT initiative, in an attempt to improve throughput, managers implemented highly bureaucratic procedures and set aggressive throughput objectives. Such actions were quite effective in increasing the effort devoted to throughput, but in so doing prevented the workers and supervisors from spending time running experiments and trying new ideas that would have improved quality and productivity. There is a fundamental incompatibility between the two types of exploitation.

Third, while a focus on exploiting people effectively curtails improvements in the underlying technology, it does not eliminate localized adaptation and improvisation. Instead, as an organization increasingly resorts to production pressure, surveillance, and bureaucratic procedures to control and exploit its workforce, workers' creativity and improvisational activities are increasingly directed at circumventing those structures rather than at improvement and learning. We found operators and engineers responding to increasingly stringent control mechanisms by cleverly creating work-arounds such as holding secret inventory, "crashing through the wall of innovation", or skimping on documentation. These innovations, often elaborately kept secret from management, allowed them to (or at least appear to) meet their targets, but seriously compromised overall system performance. Localized adaptation and improvisation under these constraints structure a system that conforms to management's expectations and confirms the benefits of a focus on exploiting people. If, like Weick (1998), we posit improvisation as a core process of organizing, then efforts to restrict it in favor of more output really amount to redirecting its focus from improving performance to circumventing control structures.

Fourth, once the focus of managerial attention turns to exploiting people, there is little chance for recovery without an external intervention or a major crisis. The focus on exploiting people is strongly self-confirming: it structures an environment in which such actions make sense. Beliefs, attitudes, behavior, and organizational results all reinforce one another. Prior to the MCT initiative, managers seeking productivity gains targeted the effort and productivity of individual workers as the path to further improvement. With time, the information system became increasingly granular (sometimes focusing on specific machines), information was reported increasingly frequently (in most cases daily or per shift), and the pressure to hit targets became increasingly intense (people who missed targets would get "beat up"). While such actions may have improved performance in the short run, they also restricted experimentation and reduced the search for new opportunities. Further, these actions also caused participants to modify (sometimes secretly) the system in ways that allowed them to achieve management's increasingly specific objectives (for example by holding secret inventory). All of these actions combined to confirm management's belief that there were few remaining opportunities for improvement. After all, the system was predictable—objectives were always achieved—and there was little evidence that opportunities for change existed since few improvements were being made. Many, if not most, managers came to believe that they worked with a fully exploited technology, thus justifying the focus on the people rather than the process.

However, while—indeed, because—the organization had fully appropriated its members' time and energy, they were literally tripping over unrealized opportunities to improve the operation of the physical system. Thus, while researchers and managers often assume that a highly-decomposed, tightly controlled process is the consequence of having fully exploited the opportunities for learning (e.g., Sitkin et al. 1994), our analysis suggests that the causality can also run in the opposite direction: the *belief* that a technology is well understood and fully exploited can result from tight control over the activities of the workforce. A focus on exploiting people can be strongly self-confirming, structuring an environment in which such actions appear to be the only way to improve performance.

Conclusion

In this paper we have explored the factors that conspire to either support or prevent internally focused change. Our data suggest that successful internally-focused change requires that managers give line-workers enough slack and authority to improvise and adapt to local contingencies. Our data also suggest that to create an environment conducive to such change, managers must believe that improvement opportunities reside within the production technology. Thus, managers' attributions regarding the cause of poor organizational performance play a critical role in determining the success of an internally-focused change effort. Building on this observation, we propose a dynamic model that captures the co-evolution of beliefs, actions, and the state of the physical system. Our model suggests that managers who believe people are the cause of low performance take actions that embed those beliefs in the physical structure of the organization and force employees to act in accordance with those beliefs. Over time managers' beliefs about those that work for them, workers' beliefs about those who manage them, and the physical structure of the environment can co-evolve to produce an organization characterized by protracted conflict, mistrust, and increasingly rigid, inflexible control structures that prevent useful change of any type.

There are numerous implications of our analysis for future research in organization theory. Most important, while the notion of feedback loops and the use of causal loop diagram mapping tools have a long history in organizational theory (e.g. Weick 1979, Masuch 1985, Richardson 1991, Sastry 1997), they have largely been on the periphery of mainstream theorizing. Many recent trends, however, suggest that organizations are far more dynamic than existing frameworks appreciate (Brown and Eisenhardt 1997). Unfortunately, existing modeling tools and empirical methods often induce blind spots that prevent scholars from fully appreciating the dynamic and sometimes transient character of organizations and the processes of organizing. Proponents of theoretical perspectives that emphasize the dynamic character of organizations (e.g. structuration) have sought to overcome these limitations by accumulating detailed, longitudinal data sets that describe the evolution of organizing practice. To date, however, little effort has been dedicated to developing more general theories of how organizations might evolve. We believe the time is ripe to supplement these rich accounts with modeling tools like causal loop diagrams that are capable of capturing the rich and evolving set of interdependencies that constitute an organization. Structuration theorists like Giddens, Barley and Orlikowski have convincingly argued that, to understand, organizations we must focus on the mutual evolution of participants' mental models, their actions, and the state of the physical technology. Coupling these rich and detailed descriptions with modeling tools that can easily and flexibly capture both the physical structure of a system and the behaviors of those working within it represents one important opportunity for continuing this line of inquiry.

References

Arkes, H. R., & Blumer, C. (1985). "The psychology of sunk cost." *Organizational Behavior and Human Decision Processes*, 35, 124-40.

Barley, S.R. 1986. "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Department," *Administrative Science Quarterly*, 31(1): 78-108.

Brehmer, B., (1992). "Dynamic Decision Making: Human Control of Complex Systems," *Acta Psychologica*, 81,211-241.

Brown, S. L. and K. M. Eisenhardt (1997). "The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations," *Administrative Science Quarterly*, 42: 1- 34.

Carroll, J., J. Sterman, and A. Markus (1997). Playing the Maintenance Game: How Mental Models Drive Organization Decisions. R. Stern and J. Halpern (eds.) *Debating Rationality: Nonrational Elements of Organizational Decision Making*. Ithaca, NY, ILR Press.

Covey, S. (1989). *The Seven Habits of Highly Effective People.* New York: Simon and Schuster.

Dean, J. W. and D. Bowen (1994). "Management Theory and Total Quality: Improving Research and Practice Through Theory Development, *Academy of Management Review*, 19(3): 392-418.

Deming, W. E. (1986). *Out of the Crisis.* Cambridge: MIT Center for Advanced Engineering Study, Cambridge, MA.

Easton, G. and S. Jarrell (1998). The Effects of Total Quality Management on Corporate Performance: An Empirical Investigation, *Journal of Business*, 71, 2.

Einhorn, H. J. & R. M. Hogarth (1985). "Ambiguity and uncertainty in probabilistic inference," *Psychological Review*, 92, 433-461.

Einhorn, H. J., & Hogarth, R. M. (1986). "Judging probable cause." *Psychological Bulletin*, 99(1), 3-19.

Eisenhardt, K.M. (1989). "Building Theories from Case Study Research," *Academy of Management Review*, Vol. 14, No. 4, 532-550.

Fine, C. (1986), "Quality Improvement and Learning in Productive Systems", Management Science, 32,10.

Forrester, J. W. (1961). *Industrial Dynamics.* Cambridge, MA: The MIT Press.

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structure*. Berkeley, CA: University of California Press.

Goldratt, E. M. and J. Cox (1986). *The Goal: A Process of Ongoing Improvement* (revised edition). Croton on Hudson, NY: North River Press.

Hackman, J. and R. Wageman (1995). "Total Quality Management: Empirical, Conceptual, and Practical Issues," *Administrative Science Quarterly*, 40: 309-342.

Hogarth, R. M. (1987). *Judgment and Choice: The Psychology of Decision* (2nd ed.). New York: Wiley.

Ishikawa, K. (1985). *What is Total Quality Control?* Englewood Cliffs, NJ: Prentice Hall.

Johnsson, Fredrik (1996). Sustainable Improvement Programs: Total Quality Excellence. Unpublished MS Thesis, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Ma.

Kahneman, D., Slovic, P., & Tversky, A. (Ed.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge Univ. Press.

Levinthal, D. and J. March (1981), "A model of adaptive organization search," *Journal of Economic Behavior and Organization*, 2, 307-333.

March, J.G. (1991). "Exploration and Exploitation in Organizational Learning," *Organization Science,* 2, 1:71-87.

March, J. G. and H. A. Simon (1958/1993) *Organizations*. 2nd ed. Oxford: Blackwell.

Marx, K (1887), *Capital*, Ch. 10, section 2, available at http://www.marxists.org/archive/marx/works/index.htm

Masuch, M. (1985). "Vicious Cycles in Organizations," *Administrative Science Quarterly*, 30:14-33.

Miles, M. and A. M. Huberman (1984). *Qualitative data analysis*, Newbury Park, Ca: Sage.

Mohr, L. (1982). *Explaining Organizational Behavior*. San Francisco, Josey-Bass

Moxnes, E. (1999). "Misperceptions of Bioeconomics," *Management Science*,44,9:1234-1248

Orlikowski, W.J. (1992). "The Duality of Technology: Rethinking the Concept of Technology in Organizations," *Organization Science*, Vol. 3, No. 3.

Orlikowski, W.J. (1995). "Action and Artifact: The Structuring of Technologies in Use." MIT Sloan School of Management Working Paper WP 3867-95.

Orlikowski, W.J. and M.J. Tyre (1994). Windows of Opportunity: Temporal Patterns of Technological Adaptation, *Organization Science*, 5, 1:98-118.

Paich, M. and Sterman, J. (1993). "Boom, Bust, and Failures to Learn in Experimental Markets." *Management Science*, 39(12), 1439-1458.

Perrow (1986). Complex Organizations: A Critical Essay. New York, McGraw Hill.

Pfeffer, J. and R. Sutton (2000). *The Knowing-Doing Gap*. Boston, Harvard Business School Press.

Plous, S. (1993). *The Psychology of Judgment and Decision Making*, New York, McGraw-Hill.

Repenning, N. (2000a). "Drive Out Fear (unless you can Drive It In): The role of agency in process improvement efforts." *Management Science*, 46,10.

Repenning, N. (2000b). "A Dynamic Model of Resource Allocation in Multi-Project Research and Development Systems," *System Dynamics Review*, 16, 3.

Richardson, G. P. (1991). *Feedback Thought in Social Science and Systems Theory*. Philadelphia: University of Pennsylvania Press.

Richardson, G. P. and Alexander Pugh (1981). *Introduction to System Dynamics Modeling with DYNAMO*. Cambridge: MIT Press.

Ross, L. 1977. "The intuitive psychologist and his shortcomings: Distortions in the attribution process." In L. Berkowitz (ed.), *Advances in experimental social psychology*, vol. 10. New York: Academic Press.

Sastry, M. A. (1997), “Problems and Paradoxes in a model of punctuated organizational change," *Administrative Science Quarterly*, 42, 2, 237-275.

Schneiderman, A. (1988). Setting Quality Goals, *Quality Progress.* April, 55-57.

Sitkin, S., K. Sutcliffe, and R. Schroeder (1994). "Distinguishing Control from Learning in Total Quality Management: A Contingency Perspective," *Academy of Management Review*, Vol. 19, No. 3:537-564.

Staw, B. M. (1976). "Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action." *Organizational Behavior and Human Performance*, 16, 27-44.

Staw, B. M. (1981). "The escalation of commitment to a course of action." *Academy of Management Review*, 6, 577-587.

Sterman, J.D. (2000), *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Chicago, IL: Irwin/McGraw Hill.

Sterman, J.D. (1994). “Learning in and about Complex Systems”, *System Dynamics Review*, Vol. 10., No. 2-3:291-330.

Sterman, J. D. (1989). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* 35 (3): 321-339.

Sterman, J., N. Repenning, and F. Kofman (1997). Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement. *Management Science*, April, 503-521.

Taylor, S. E. and S. T. Fiske (1975). "Point of view and perceptions of causality." *Journal of Personality and Social Psychology*. 32, 439-445.

Thaler, R. (1980). "Towards a positive theory of consumer choice." *Journal of Economic Behavior and Organization*, 1, 39-60.

Tushman, M. L. and C. A. O'Reilly (1996). "Ambidextrous Organizations: Managing Evolutionary and Revolutionary Change0", *California Management Review*, 38, 4, 8-30.

Tushman, M.L. and E. Romanelli (1985). "Organizational Evolution: A metamorphosis model of convergence and reorientation." In L.L. Cummings and B.M. Staw (eds), *Research in Organizational Behavior*,7:171-222.

Weick, K.E. (1998). "Improvisation as a Mindset for Organizational Analysis," *Organization Science*, 9, 5: 543-555.

Weick, K.E. (1979). *The Social Psychology of Organizing*, Second Edition, New York, Random House.

Westphal, J. D., R. Gulati, and S. M. Shortell (1997). "Customization or Conformity? An Institutional and Network Perspective on the Content and Consequences of TQM Adoption," *Administrative Science Quarterly*, 42: 366-395.

Wheelwright, S. and K. Clark (1992). *Revolutionizing Product Development: Quantum Leaps in Speed Efficiency and Quality*, The Free Press, New York, NY.

Wheelwright, S. and K. Clark (1995). *Leading Product Development*, The Free Press, New York, NY.

Wruck, K.H., and M.C. Jensen (1994). "Science, Specific Knowledge, and Total Quality Management", *Journal of Accounting and Economics*, 18, 247-287.

Wynne, B. (1988). “Unruly Technology: Practical Rules, Impractical Discourse and Public Understanding”, *Social Studies of Science*, 18, 147-167.

Zbaracki, M.J. (1998). "The Rhetoric and Reality of Total Quality Management, "*Administrative Science Quarterly*, 43: 602-636.