The Problem with Workarounds is that They Work

J. Bradley Morrison

Brandeis University

International Business School

Waltham, MA

bmorriso@brandeis.edu

(781) 736-2246

February 2011
The Problem with Workarounds is that They Work

ABSTRACT

Resource shortages are a fact of life in most organizations and in the currently challenging economic climate will be even more common. There is little doubt that resource shortages are widespread and that they lead to undesirable outcomes, yet there has been surprisingly little attention to questions about why they occur and even more insidiously why they persist despite the apparently clear adverse consequences. This paper develops a grounded theory that helps to understand chronic resource shortages, drawing on data from field work at a manufacturing firm adopting lean manufacturing. The paper examines how the actions of various groups (e.g., managers, production workers, and other shop floor workers) interact with each other and with the physical characteristics of the workplace to sustain problematic resource shortages. The paper uses a causal loop diagram to highlight some important features of the dynamics and a system dynamics model for simulation analysis and theory building.
The Problem with Workarounds is that They Work

Resource shortages are a fact of life in most organizations. Businesses strive to achieve more with less in order to meet pressures for profitability, satisfy increasing demands from their customers, or respond to new opportunities in the marketplace. And, most companies aim to also build capabilities for improved performance over time. Some undertake proven, planned process improvement initiatives; others build capability through learning by doing. Both producing to meet market needs and investing to build capability require resources, in particular the time and focused attention of people in these organizations. But, these resources are costly and usually constrained in availability, so organizations often try to do more with less. Indeed, in the currently challenging economic climate, we can expect to find more and more organizations cutting costs through staff reductions - more than they reduce expectations for output under the banner of “do more with less”- resulting in even more stress on the workforce.

Despite the widespread prevalence of under-resourcing, scholars have chronicled several perilous consequences. Resource shortages have been implicated in the failure of process improvement initiatives such as total quality management (TQM) implementations (Keating, Oliva et al. 1999), the emergence of erroneous and dysfunctional attributions that give rise to a capability trap (Repenning and Sterman 2002), deterioration of product development processes (Repenning, Goncalves et al. 2001), degradation of quality in service organizations (Oliva and Sterman 2001), and both schedule and budget overruns in project management (Cooper 1980). Moreover, as described in The Overworked American, workplace pressures have led over two decades to American workers each year putting in an average of nine more work hours, or just a little over a workday, compared to the previous year, with consequences ranging from burnout and fatigue to absenteeism, declining health, and strained interpersonal relationships (Shor 1991).

There is little doubt that resource shortages are commonplace and that they sometimes lead to undesirable outcomes. Although scholars have documented their existence and their perils, there has been surprisingly little attention to questions about why they occur
and even more insidiously why they persist despite the apparently clear adverse consequences. Some resource shortages can be explained by the fact that under-resourcing can often invoke better-before-worse trajectories of performance the will be preferred when temporal tradeoffs favor short-term performance benefit at the expense of long-term cost, whether these trade-offs are financial discounting or psychological biases (Sterman 1994). But this explanation alone is not satisfying, especially since it is not adequate to explain the persistence of resource shortages.

This paper develops a grounded theory that helps to understand chronic resource shortages, drawing on data from field work at a manufacturing firm as it attempted to adopt production practices based on lean manufacturing. The paper is organized as follows. The next section describes the research setting and the change initiative. The following section describes the research methods. The next section examines how the actions of various groups (e.g., managers, production workers, and other shop floor workers) interact with each other and with the physical characteristics of the workplace to sustain the problematic resource shortages observed in the source data. The section develops a causal loop diagram grounded in the ethnographic data to highlight some important features of the dynamics. Finally, I conclude with a discussion of implications for managers and for scholars concerned with organizational change.

**THE RESEARCH SETTING**

The site of the research is a plant that manufactures parts and assembles engines that are installed in another of the company’s plants into vehicles the company sells. The general manager of the plant became interested, in his words, in “taking [the company] to a new level” of manufacturing performance by adopting the practices of “lean manufacturing” (Womack, Jones et al. 1990) or the Toyota Production System (Monden 1983). I had the opportunity to observe the effort over a period of 20 months, beginning when the general manager was beginning to think about how to introduce new production practices, a long enough time horizon to observe some illuminating dynamics in the first sections of the plant where implementation was attempted.
The engine plant employs a unionized workforce. Management and union leaders alike attest to a strong, cooperative relationship between the union and the management. For several years, work practices have been guided in part by the collective bargaining agreement, which details elements of the organization structure and responsibilities of the union leaders and workers and the salaried management. Elements of the agreement include work groups organized around production units, “work group advisors” rather than supervisors, joint representation in many decision processes, and a high degree of information sharing with employees, as well as other practices typically characterized as high-involvement work practices (Pil and MacDuffie 1996).

METHODS

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

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

**Data Analysis**

As Barley (1990, p. 234) notes, "the analysis of field data actually begins during a study's observational phase." The ongoing analysis of field data offers the possibility of developing interim hypotheses and directs attention to relevant data (Glaser and Strauss 1967). My data analysis followed traditional qualitative methods (Miles and Huberman 1984; Eisenhardt 1989; Yin 1994). I also made frequent use of causal loop diagrams
because they are especially useful in the analysis of complex systems (Weick 1979; Masuch 1985; Sterman 2000). During the course of my fieldwork, I made frequent sketches of such causal loop diagrams, often in the evenings during my trips to the research site. The diagrams were one explicit means by which I was "always trying to make sense of one's data and thinking about what more one can find out" (Feldman 2000, p. 615).

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

THE IMPLEMENTATION CHALLENGE

Managers in the plant recognized a need to improve the manufacturing capabilities in their organization. They discussed the idea with union leaders at the plant and agreed to proceed. As one union leader said:
I came to understand that we are archaic in the way we do our manufacturing here. We all know that competition is always knocking on our door. We know that if we're going to be competitive, especially here at this plant, because we are such a low-budget, low-profit margin, here, that we’re going to have to continuously improve to be competitive, and keep the [product] at a price where we can get people, … first time buyers, into our family.

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

Enthusiasm was high, as one hourly employee said:

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

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

They got people looking at the machines, finally fixing them. We got the layout running. You got a pull system in place. And – start looking at some of the overall numbers – they're outstanding. Scrap has come down [thousands of dollars]. Performance went up from 70% to 94% [in-stock levels of finished goods inventory meeting or exceeding target minimums].

Yet several months later, some new work practices had been abandoned and performance had deteriorated. One informant described the situation as “the wheels are coming off.” Another said, "If you go over there a couple of months later, after all this stuff, the
started to deteriorate." Several respondents reported similar characterizations of the improvement activity in this work group: an early phase of improved performance followed by a plateau and then decline, the start and fizzle pattern. The purpose of the following analysis is to develop a dynamic model that captures the ongoing interaction between the work of improvement activity and the organizational context to shed some light on why implementation of process improvement displays these perverse dynamics.

In what follows, I draw on my field data to induce a model that describes how the choice, policies, and actions of various organizational members interact with the physical characteristics of improvement activity. I represent the model using causal loop diagrams to describe the feedback structure enacted by the organization and discuss how this structure gives rise to the observed dynamics. Simulation analysis is a useful method to understand the dynamics of feedback systems and to advance to process of theory building (Sterman 2001; Davis, Eisenhardt et al. 2007). To further illuminate how the feedback structures shown generate specific patterns of behavior over time, I also use a mathematical model to simulate the logical consequences of the causal relationships presented. Complete details of the models used are available in a technical appendix.

Expanding on a format typical of grounded theory, the following exposition moves back and forth among the motivating source data, the emerging theory (as represented by the causal diagrams), and the simulated behaviors that the feedback structures generate.

**Interaction with the Managers**

The approach to capability building that is at the core of the Toyota Production System, the framework on which the focal improvement initiative was based, is participatory improvement – front lines workers are empowered and engaged not only to do the salient work of producing their marketable output (e.g., parts or engines) but also to continuously improve the process of doing their work (Ohno 1988; Spear and Bowen 1999). In the research setting, close-in observation of how this improvement activity was actually done showed that front-line workers generated improvement ideas, but these ideas resulted in task work assigned to other people (the support personnel), such as
maintenance workers, engineers, and other salaried employees. As events unfolded, support personnel were in great demand, and their limited availability was an issue. Thus, I turn my attention to the managers (or others who have influence on the support personnel in question) and how they dealt with this issue.

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

We always had updates to the leadership group on how things were going. … They asked what we needed, and our request was we need a list of people that we can go to, maintenance, facilities, that we can call on the phone or whatever and get things done.

I begin the explanation of the model development by representing the intended rationale of this request in Figure 1. The request from the implementation team to management was to increase the Assigned Workers. The diagram shows that assigning workers would lead to an increase in the Work Rate. After some time for managers to notice, the Managers Work Rate Estimate would adjust towards the higher Work Rate, thus easing the Pressure to Assign Workers, which arises from an imbalance between the Desired Work Rate and the Managers Work Rate Estimate. This causal pathway closes a feedback loop that works to reduce the gap between the Managers Work Rate Estimate and the Desired Work Rate by adjusting the quantity of Assigned Workers. The feedback loop is a balancing loop, labeled B1, because it acts to offset or bring into balance a change that takes place. For example, imagine an increase in the Desired Work Rate, which would induce some Pressure to Assign Workers, which would cause manager to allocate more Assigned Workers thus increasing the Work Rate and after a short perception delay also increasing the Managers Work Rate Estimate. Pressure to Assign Workers is thus reduced.

[Insert Figure 1 about here.]

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

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

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

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

Table 1 summarizes some of the interview data that reveal the chronic and salient nature of the resource shortages.

[Insert Table 1 about here.]

The persistent resource shortfall raise two questions. First, under these circumstances of resource starvation, how did the workforce respond? Second, why, in the face of the clear requests for more, did the managers not allocate more resources? Were the managers simply being unresponsive to the requests from the workforce? Or, is there another explanation? To explore these questions, consider how the actions of the work group and implementation team may have influenced the managers.

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

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

So, the support resources find other ways of getting things done. Some scholars view this process as one of improvisation or bricolage - the "use of whatever resources and repertoire one has to perform whatever task one faces." (Weick 1993, p. 352) The mounting workload challenges the support personnel, the bricoleur in our example, to somehow increase the Work Rate. Under the prevailing condition of fixed resources, the options available to the overworked person facing this challenge must take the form of somehow finding a way to accomplish more tasks in the same amount of time.

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

You'd kind of lay it out as an engineer yourself and come down and talk to the operators and they'd sign it and it's a done deal. [With] more input [from the work group], of course the process is going to be a lot longer. The process is longer to implement than actual layout moves. Theoretically, in the past, it was not a long process.

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

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

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

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

Working in a less collaborative manner is one possible solution, a type of workaround or shortcut that will get the work done to alleviate the dominant pressures although doing so
may require compromising some standard. Other researchers have documented the use of workarounds in response to time pressure. For example in a commercial bank, lending officers facing work pressure from a backlog of orders cut corners to reduce the time they spend on each customer order (Oliva and Sterman 2001).

The use of workarounds to increase the *Work Rate* creates another balancing feedback loop, also shown in Figure 1. As *Pressure on Employees* increases, the support personnel respond with *Workarounds*, such as conducting their work in a less collaborative manner. Because the workaround is quicker, the result is an increase in the *Work Rate*, which eventually increases the *Front Line Work Rate Estimate* and thus relieves some of the *Pressure on Employees*. The balancing loop thus formed is designated Loop B2 in Figure 1.

The structure depicted in Figure 1 then shows two balancing feedback loops that describe two responses, one by the managers and the other by front-line workers, to address a shortfall in production (i.e., work rate) relative to the target (i.e., desired work rate). To investigate the interplay of these responses, I turn to a mathematical model of the feedback structure in Figure 1. The mathematical model formalizes the relationships shown in the diagram by defining one equation for each variable in the diagram. For example, $Pressure \text{ to Assign Workers} = \frac{Desired Work Rate}{Managers Work Rate Estimate}$. The two feedback loops B1 and B2 are formulated as precisely analogous equations with only one difference: the perception delay between the *Work Rate* and the estimated work rate is longer for the managers than for the front-line workers, reflecting the longer delays for managers to become aware and update their own formal or informal estimates of what is happening on the shop floor. The model draws heavily on standard model system dynamics formulations, familiar fragments of model structure that frequently occur (Sterman 2000) and is fully documented in the technical appendix. I use the model to conduct experiments to see how a system characterized by the posited feedback structure will behave under various conditions.
Figure 2 shows the simulating results of two scenarios. In both cases (and all of the experiments shown in the remainder of the paper), the firm begins in equilibrium conditions such that the actual work rate is equal to the desired work rate. That is, work practices and the allocation of resources are such that the rate at which work is getting done is just exactly equal to the rate at which new work needs are arriving. While such an equilibrium is hardly representative of real world conditions, it is good modeling practice to begin from such baselines in order to best understand the system behavior observed as a result of some change. The tests shown here introduce an increase in workload, such as might arise from embarking on a new improvement initiative or from raising a performance goal.

The two charts in the left column of Figure 2 show the results of the first scenario in which only the managers’ adjustment is active. The ability for workers to use workarounds is disabled in this scenario. The top graph compares the Desired Work Rate and the Work Rate. The lower graph compares the two responses measure as indices compared to the initial values. The resource index shows the relative quantity of Assigned Workers, and Workarounds show the relative intensity with which the front line personnel are using workarounds. The simulations show two important features of the dynamics. First, in response to the increase in Desired Work Rate, the managers respond by increasing the allocation of resources, and the simulation ends with more resources in place, as shown by the graph of the resource index. Second, the adjustment is effective in bringing the work rate to equal the desired work rate. Through a period of somewhat turbulent adjustment, the work rate does indeed grow to match the new higher target.

The second column of Figure 2 shows the scenario in which both the managers’ adjustment and ability for workers to use workarounds are active as responses to the increased workload. Now we see that once again, the system responds effectively to the change, and the work rate grows to match the desired work rate. Indeed, as the top graph shows, in comparison to the first scenario, the convergence to the desired rate takes place more quickly. At least from the manager’s perspective, this appears to be a smoother
transition to the sustained higher level of output. A second important difference is that the adjustments in this scenario show that workarounds have become an important part of the solution. Not only are workarounds a prominent feature of the new way of doing work, the workarounds have accounted for a larger portion of the gap-closing adjustments the organization has made. This consequence, that the workers have in effect borne more of this burden, is a result of the asymmetry in the perception delays. Because the (industrious and well-intentioned) workers quickly realize the overload condition and act to do something to make due under the circumstances, their solution (workarounds) becomes dominant solution.

These signs of progress were a source of feedback to managers reduced the sense of urgency to allocate more resources. Because the worker’s industrious and creative solutions in the form of workarounds as in loop B1 are indeed effective, the manager’s response in the form of B2 is not needed, or at least not needed as much. And, because there are important delays in the manager loop B1, such as for managers to perceive the need, make decisions to act, find appropriate resources, relieve those resources of other responsibilities and reassign them to the focal effort, loop B2 will contribute a far greater portion of the solution to the need for a higher \textit{Work Rate}. The very success the support personnel have in accomplishing these tasks with the workarounds they improvise, such as doing things less collaboratively, sends a signal to managers that the work is indeed getting done. Seeing the work get done, the managers feel less pressure to allocate more personnel, so the support personnel get locked in to a pattern of working with less collaboration. The more productivity gained from reducing collaboration, the stronger the boost in getting tasks done will be. The more effective the shortcuts are at increasing productivity, the more they undermine the fundamental solution to the problem they are addressing - excess work to do.

**Learning by Doing**

There are some benefits to doing things in the prescribed manner that are unfortunately circumvented when workarounds are used. A consequence of working collaboratively is that it leads to experience collaborating which increases the productivity of the time spent
collaborating. Moreover, workers build knowledge of the production process through involvement in improvement activity, which fosters better idea generation in the future. Figure 3 introduces the variable *Proficiency with New Skill* to summarize the lasting benefits, accumulated through learning by doing when appropriate methods are used. The use of workarounds compromises this learning. These relationships form a different type of feedback loop, a reinforcing loop labeled R3 in Figure 3. For example, as *Pressure on Employees* increases stimulating an increased use of *Workarounds*, the *Proficiency with New Skill* will be lower than otherwise, so *Productivity* declines and so too does the *Work Rate*, all else equal. The *Front Line Work Rate Estimate* will be even lower and thus the increase in *Pressure on Employees* is reinforced and continues to grow. This loop may often be working as a vicious cycle, pushing the system to further lock in to the workaround method of doing things.

[Insert Figure 3 about here.]

Figure 4 shows the results of the same experiment conducted in Figure 2 when the model includes the feedback structure shown in Figure 3. Now, the proficiency with new skill is a quantity that varies based on the history of learning by doing. The graphs in Figure 4 show that the system responds to address the need for a higher work rate, closing the gap so that work rate equals desired work rate. As before, the adjustment by the workers through the use of workarounds is the larger component of the solution. But, this simulation highlights another important feature of these dynamics. The increase work rate comes at the expense of a decline in proficiency. The high previous level of proficiency is in a sense a form of organizational slack. The workers have adopted practices that rely on workaround that have the dual feature that they do indeed increase output, but only at the expense of hollowing some of the fundamental capability of the organization.

[Insert Figure 4 about here.]
The simulation in Figure 5 are similar to those in Figure 4, but now the increase in desired work rate is only temporary – for a period of 30 simulated weeks starting in week 10 of the simulation. The graphs in the left column show the results from a modest temporary increase in workload. Proficiency declines for a period of time while the work demand is higher and for a short time afterwards, but then the system is able to restore the Proficiency to its initial levels once the temporary challenge has abated. The graphs in the second column, however, tell a far different story. In this scenario, the increase in desired workload is again temporary but the magnitude of the increase is (only slightly) larger than that in the test for the left column. The outcome now is a permanent decline in proficiency. The system has crossed a tipping point. The workers are locked in to a new (presumably inferior) method of working based on a higher proportion of workarounds, which are necessary because they have lost so much proficiency that in order to achieve even the original rate of working, they must take shortcuts. The system has been overwhelmed by the larger temporary increase – but notice that the situation may not appear so horrific to a manager who is focused only on the work rate. The behaviors in the “Adjustment Index” graphs may not be visible to some managers, as the workers use workarounds. Moreover, the decline in proficiency may be even less salient. More insidiously, such deterioration in proficiency may lead to seriously dysfunctional attributions (Repenning and Sterman 2002).

[Insert Figure 5 about here.]

An important consequence of the workarounds was that in most cases these efforts worked and therefore generated signals that improvements were being made and even led to positive business results (e.g., improvements in cost, quality, or delivery). The most salient performance problem is a stock-out of parts when needed by the assembly line. The situation in the work group before the improvement initiative started, as described by the production manager was that:

We were not doing very well in [this manufacturing cell]. It was almost a daily phone call from [the other engine plant about] the fact that we shut them down for some reason or other because of a quality issue or lack of product. So we weren’t feeling really good about what was going on down in [that manufacturing cell].
But, the efforts in the work group had begun to significantly improve the situation. As one union member described:

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

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

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

Commenting that the cost of personnel should be obvious, a production manager said:

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

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

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

The challenge is more than just making more support personnel time available. The challenge is to make the right person with the right skills available at the right time.

When asked why the managers did not allocate more support personnel, one answer from a production manager is simply that it is difficult to anticipate the needs:

We [management and front-line workers] are not as good at defining what resources are needed. … I don't think that we do a very good job of being clear about what is the work we need to do … or when does it need to be
done …and who are the resources to get it done [and] how we support that work to get it done when we said it was going to get done. Around here it seems to be it'll get done when it gets done.

Moreover, in many situations the key person is a "shared resource" with duties that span many areas in the plant. As another production manager explains:

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

**Additional Challenges**

There are many reasons to explain why the managers favored the salient and optimistic interpretation that resources are not urgently needed. First, the optimism is plausible. Managers were exposed to signals that may give them a mistaken impression about the progress of the improvement initiative. In light of a support personnel shortage, the people focused on the improvements chose tactics that circumvented the high involvement approaches that were the intent of the program. These efforts did indeed generate some positive and salient results, even if temporary. Managers were thus given a false sense of the progress of the program. Second, accomplishing tasks is more salient than doing work collaboratively. Research has repeatedly shown that people overweight salient and available information (Taylor and Fiske 1975; Tversky and Kahneman 1982), a bias that has been implicated in failures to learn and in the pathology of process improvement (Sterman 1994; Repenning and Sterman 2002).

A second reason is that reporting practices often exacerbate this optimistic interpretation of progress. As Figure 6 shows, when *Salient Results* are accomplished, workers may be less likely to make *Requests for Help*, forming balancing loop B4. Moreover, the positive attitude, enthusiasm, and pride that often accompany successful implementation, even if accomplished because of effective workarounds, may foster a “can-do” attitude that further suppresses the *Requests for Help*, as shown in loop B5.

[Insert Figure 5 about here.]
In the initiative studied, the work group and implementation team themselves often reported with pride the accomplishments they were making. Indeed, in the very presentations during which they asked the managers to assist with the shortage of support personnel availability, the implementation team reported continued progress to the managers. Several months after the start of the initiative, the work group and implementation team presented a progress report to the plant management. Using 23 PowerPoint slides, seven production workers conducted a presentation complete with before and after pictures, graphs and charts of progress. They reported improvement activity and claimed savings from reductions in floor space, set-up time, overtime and scrap materials and benefits from increases in in-stock percentages of finished goods. The presentation emphasized progress, success, enthusiasm, and positive business results.

The group did raise the issue of the need for support personnel - in one bullet point on one of the 23 PowerPoint slides that was discussed for less than 2 minutes out of the 30-minute session. The plant management was so impressed with the progress that they arranged for the work group to make the presentation to three other groups of managers from the company, including the senior executive team of the manufacturing division. By the time the work group presented to these senior executives, the presentation had grown to 42 slides. The slide that noted the need for support personnel had been dropped, and a new slide showed an email to the plant managers and the work group from their main customer in the other production facility:

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

Moreover, the managers themselves frequently described the successes of the improvement program in presentations to peers, conversations with plant employees, and discussions with each other. Both the individuals doing the work and the managers put a positive spin on the results. Balancing loop B6 in Figure 6 reflects the idea that these habits of accomplishing more with less can over time foster the emergence of a culture of achievement that still further suppresses the Requests for Help that might send corrective signals to management that more resources are indeed needed.
DISCUSSION AND IMPLICATIONS

In the preceding analysis, I have drawn from a field study of process improvement at a manufacturing plant to elucidate key feedback relationships and induce a model of organizational change. Following Repenning and Sterman (2002), I have posited a feedback structure that calls attention to critical interactions between characteristics of the workplace and the behaviors of the agents acting in the system. I offer a more finely grained analysis that examines the interactions among the work of several groups of personnel on the shop floor and the work of managers. The model explains how people's well-intentioned actions interact to determine the observed pattern of behavior at the organizational level. I find that a high involvement improvement program, such as lean manufacturing, constitutes a feedback structure with the potential to be self-limiting. The generation of ideas through high involvement activities generates demands on key personnel and induces the people involved to find ways to work around the resource constraints they face, but the workarounds undermine the effectiveness of future implementation. I have examined four consequences of these workarounds, the ways people make do with what they have available to do the best they can under the circumstances. First, the efforts do generally yield results - that is, the intent of getting the work done is realized. Second, the successes they generate with their ways of working around the support personnel constraints send signals to management that more personnel are not needed and thus counteract their requests for more support personnel. Third, the approaches they take include ones that circumvent the intended high involvement of the work force, thus precluding the development of resources in the form of human and social capital that benefit future implementation. Fourth, circumventing the high involvement of the workforce also compromises worker understanding of the new production processes.

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

The notion that building organizational capability requires an investment of resources is not new, but the study here adds two points to suggest that managers may be unwittingly undermining their own futures. First, the self-reinforcing nature of alterations in how work gets done recasts the magnitude of the concern. Initial changes in work practices to meet the challenges of constrained personnel can easily become locked in. Second, because managers are likely to be slow to recognize personnel shortages, the shortage conditions are likely to persist. Work does get done, and changes do get implemented, so managers see tangible evidence that things are not that bad. Negative consequences of personnel shortages are often less salient, less certain, and result only after a significant delay compared to the vivid, certain, and immediate outcomes they notice. Researchers have documented the tendency of people to overweight salient and available information, to fail to recognize important delays, and to exhibit an aversion to risk (Kahneman, Slovic et al. 1982; Dawes 1988; Sterman 1994). Repenning and Sterman (2002) discuss how these effects yield a preference for work for which the outcomes are salient and are achieved with short delays and greater certainty. Taken together, these additional insights suggest that much of organizational life may take place in the aftermath of personnel shortages.
This work has implications for how we study change in organizations, because it focuses our attention on the interactions between managers, shop floor workers, and other groups of employees. The groups, apparently acting in their own best interest, generate outcomes that are in conflict with their apparent objectives. Similarly, this work has implications for practice, since it calls attention to some of the potential pitfalls in a high involvement improvement program. In particular, this work highlights the critical importance of support personnel. Moreover, the feedback structure posited here describes some potential barriers to the organization's ability to sustain momentum for change. The interactions are subtle, masked by the apparent progress that effective shortcuts generates, and embedded in complex feedback processes. Powerful biases favor interpreting the course of events as the best of times, even when the critical resource of social capital is deteriorating, so also enacting the worst of times. Managers and researchers both stand to gain from a shift of attention away from the salient content of improvement activity and toward the way in which work gets done builds the capability for future success.
REFERENCES


So we went up there and did the analysis, got some of the tools ordered, and did some of the basic things. Then we were starting to require more and more engineering. … That’s where it really fell down. I think probably the biggest thing we were missing on that team is we had no one to go to. (Implementation Team Member).

The workers will always be able to generate ideas faster than we can implement them. (Production Manager).

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

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

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

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

It wasn’t really [anything] overt that we’re not going to do this and not going to do that, but it was just like [pause] you’re pulling an ox cart through a mud pit is basically what you felt like. (Implementation team member).

A guy like [our engineer] and those guys, those guys are busy. They've got so many things on their platter that they can't devote the time that you need for that process. (Union official).

Any time there was a problem it fell back on the engineers. … Everything came down to the engineers. (Implementation team member).

You couldn't get anything fixed. Do you hear what I am saying? I mean, simple things like putting lines on the floor and getting things moved, it just didn't happen. (Production worker).

So, [imagine] I'm in a work group. I need to get something done on my machine, and I ask maintenance to go deal with that. And I want it done right then. Right? I am empowered to do that. I am behaving in a way that is consistent with this principle. Maintenance then says, "OK, I have got this work group that has asked just me for some help, but I have a down machine over here. I've got to make a choice as to where I am going to work, because I am a scarce resource. (Production Manager).
FIGURES

Figure 1: Two Solutions to the Pressure to Get More Done

Figure 2: Dynamics of Manager’s and Front-line’s Adjustment to Increased Workload
Figure 3: Feedback Structure with Learning by Doing of New Capabilities

- Desired Work Rate
- Assigned Workers
- Pressure to Assign Workers
- Proficiency with New Skill
- Workarounds
- Pressure on Employees
- Managers Work Rate Estimate
- Front Line Work Rate Estimate
- Learn by Doing
- B1: Manager’s Adjustment
- B2: Front Line Adjustment

Diagram showing feedback loops involving work rate, assigned workers, pressure to assign workers, proficiency with new skill, workarounds, and pressure on employees.
Figure 4: Response to Increased Workload with Variable Capability

Response to Increase in Work Arrival

- Work Rate
- Adjustment Index
- Proficiency

Graphs showing changes in work rate, adjustment index, and proficiency over time with different indices and workarounds.
Figure 5: Response to Increased Workload with Variable Capability

Response to Temporary Increase in Work Arrival

Work Rate

Adjustment Index

Proficiency

Response to Slightly Larger Temporary Increase in Work Arrival

Desired Work Rate

Work Rate

Resource Index

Workarounds

Capability
Figure 6: Additional Feedback Structure in the Implementation Challenge

- **Desired Work Rate**
- **Pressure to Assign Workers**
- **Assigned Workers**
- **Pressure on Employees**
- **Workarounds**
- **Proficiency with New Skill**
- **Learn by Doing**
- **Habits of Mind**
- **Self Efficacy**
- **Satisficing**
- **Requests for Help**
- **Pride and Spin**
- **Salient Results**
- **Pdy**
- **B1** (Manager's Adjustment)
- **B2** (Front Line Work Rate Estimate)
- **B3** (Front Line Work Rate Estimate)
- **B4** (Manager's Work Rate Estimate)
- **B5** (Habits of Mind)
- **B6** ("Can-Do" Culture)