Modeling and Worker Motivation in JIT Production Systems

Kenneth L. Schultz • David C. Juran • John W. Boudreau • John O. McClain
L. Joseph Thomas
Kelley School of Business, Indiana University, Bloomington, Indiana 47401
Graduate School of Business, Columbia University, New York, New York 10027
S. C. Johnson Graduate School of Management, Cornell University, Ithaca, New York 14853

This paper concerns the modeling of low inventory lines. Currently, most models assume that processing times are independent. We consider the differences in behavior of workers in low- and high-inventory production lines. Using a laboratory experiment we show that workers speed up when they are the cause of idle time on the line. This means that processing time distributions are not independent of the size of the buffer, of the processing speed of co-workers, or of the amount of inventory in the system. We show that the direction of these effects is predictable and that the magnitude is significant. In particular, there is less idle time and higher output than would be predicted using assumptions of independence. In this experiment the effect completely canceled productivity loss due to blocking and starving. This work is important in understanding both the motivation of workers in low-inventory systems and the implications of models of manufacturing flow lines.
(*JIT*; *Serial Production Lines*; *Independence Assumption*; *Job Design*)

1. Introduction

This paper analyzes the independence assumption commonly used in the modeling of serial production lines. The central question is: are workers in low-inventory situations motivated to act differently than they would with high inventory? Are these differences beneficial, and are they large enough to matter? We will show that the answer to these questions is yes. On low-inventory lines workers adjust their work pace to avoid causing idle time. This leads, in part, to slower workers working faster than in high-inventory systems. Because of these behavioral changes, idle time and processing time are less than expected. Productivity is enhanced to the point where an unbalanced, low-inventory, serial line with significant blocking and starving can match the throughput of the same machines working in parallel.

This work is important to understanding both the motivation of workers in low-inventory systems and the implications of manufacturing flow line models. Just-in-time (JIT) management practices are of increasing importance in manufacturing. A significant part of any JIT system is lower inventory. We hope to improve management techniques through better understanding of the effects of low inventory on worker motivation. Furthermore, we hope to improve the modeling of these systems by analyzing the underlying assumptions.

The operations management literature has a rich history analyzing the role of inventory in dedicated serial production lines with stochastic processing times. The general conclusion of this literature is that lower inventory increases the chances of idle time due to blocking and starving, and therefore decreases the production rate. Blocking occurs when a workstation is idle because there is no room to place output. Starving occurs when a workstation is idle because there is no work in the input buffer to process. Most of this literature uses what we will call the *independence assumption*. It states that the processing time of a machine is independent of past

0025-1909/98/4412/1595$05.00

events and of the current state of the system. We define a buffer as an inventory storage location between two machines. The independence assumption asserts that a worker's processing time is not affected by changes in the size of the buffer, the amount of inventory in the buffer, or the other workers on the line. This allows models to be tractable which can lead to valuable insights. While few would argue that the assumption is strictly true in all cases, there is no current research to suggest a consistent direction or magnitude of error.

This paper looks closely at how this assumption might fail and the implications for management practice and modeling. We present evidence that worker behavior is different, based on the size and status of the buffers. While the independence assumption appears valid in high-inventory situations, in low-inventory systems work speed is not independent of the speed of coworkers or the inventory state of the system. Furthermore, the direction of these differences is predictable and significant. Knowledge of these differences and the reasons for them will improve the understanding of low-inventory systems.

The next section reviews the literature on the role of buffers and inventory in serial production lines and on the behavioral effects of new manufacturing techniques. In later sections we develop hypotheses relating work pace to the inventory state of the systems and describe a laboratory experiment to test these hypotheses. Our findings and a new analysis of data from Doerr et al. (1996), followed by a discussion and conclusions, complete the paper.

2. Literature Review

Much has been written on the mathematical modeling and analysis of manufacturing flow lines. For an excellent review see Dallery and Gershwin (1992). They noted that successive processing times are usually assumed to be independent of one another and independent of processing times at other machines (for one exception, in a different context, see Robinson and Hendricks 1995). Using this assumption, Shanthikumar and Yao (1989) demonstrated that the production rate is an increasing function of the buffer capacities. That is, given two lines that are identical except that every buffer in the second line is at least as large as the corresponding buffer in the first line, the production rate of the second line will be at least as great as the production rate of the first line. Conway et al. (1988) showed that increases in buffer size lead to increases in productivity, but with sharply diminishing returns. They showed that most of the production benefit comes with the first few buffer slots and that this effect is even more pronounced in unbalanced lines. There is consensus, however, on the use of the independence assumption and that lower inventory leads to increased chances of blocking and starving, which, in turn, reduces production rate.

In spite of the above results, many authors (for example, see Schonberger 1982) have observed productivity increases in low-inventory lines. In general they attribute these increases to process improvements. For instance, JIT allows managers to notice and correct poor quality quickly, which leads to less rework and reduced loss. Lower inventory reduces waste due to damage and obsolescence of inventory. It reduces nonvalue-added labor in the physical handling of inventory. It is usually accompanied by continuous improvement and changes in processes that can also increase productivity.

All these explanations are valid. However, they consider only a technological viewpoint. Socio-Technical Systems Theory states that any major change in technology, such as the introduction of JIT, invariably involves changes in the social and cultural systems as well, often with unforeseen consequences (Tichy 1983). We suggest that these social and technical changes lead to changes in worker behavior that counteract the effects of idle time due to blocking and starving.

Much less has been written about the effect of low-inventory lines on workers. Most of this work has concentrated on implementation and human resource management issues. (See, for example, Schonberger 1982, Snell and Dean 1992, Huber and Brown 1991, or Johnson and Manoocheri 1990.) Two articles have looked at job satisfaction with mixed results. Groebner and Merz (1994) found no change in job satisfaction when work changed to a JIT system. Scott et al. (1992) found increases in self-reported satisfaction with increases in self-reported involvement in JIT. While job satisfaction has been linked to some work behavior, such as attendance, it does not appear to increase work effort (see Cranny et al. 1992). Unfortunately, this literature does not help us assess possible changes in work effort that will be associated with JIT implementation.

What little work there is on the motivational effects of low inventory is anecdotal and focused on management use of the work system to improve control of the production process. With small buffers the state of the inventory provides feedback on the relative production rate of the line on both sides of the buffer. If the buffer fills up, the downstream station is too slow. If it empties, the upstream station is the problem. Some authors have argued that management uses this feedback to coerce workers into higher levels of effort. They reasoned that, by reducing inventory, management can better monitor worker performance and concentrate effort on slower workers. Parker and Slaughter (1988) attributed productivity increases at NUMMI to tighter management control. Sewell and Wilkinson (1992) compared JIT factories to optimally designed prisons. They stated that JIT regimes ''both create and demand systems of surveillance that improve on those of the traditional bureaucracy in instilling discipline,'' thereby ''consolidating central control and making it more efficient.'' Delbridge et al. (1992) argued that ''JIT intensifies work as a result of increased surveillance and monitoring of workers' activities.'' These papers made some mention of worker interaction through the mechanisms of peer pressure. In general, however, this body of work is anecdotal and observational in nature and does not measure the motivation effects or relate them to current theory.

An excellent observational study was written by Janice Klein (1987). She discussed the loss of worker autonomy with the introduction of JIT and SPC in an engine plant. She also noted the effects of differences in work speed and described worker perceptions of speed up and stress. She quoted one manager as saying: ''Everyone is impacted, both those who worked faster and slower in the old days. If I was a person who liked to build a lot of engines I could work fast—it was fun, it felt good, I accomplished a lot, and got a lot of satisfaction from it. Today, I am slowed down, so I don't like it because I feel like I'm not producing and it is boring. You take a slow person, one who wasn't pushed at all by the old system, and they have an awful lot of stress because they are really pacing the line, so they have a lot of pressure. The entire thing is very stressful—stressful for the slow person and stressful for the fast person.''

It is unclear whether this manager attributed pressure on slow workers to increased management awareness due to reduced inventory or to other causes. The quote does clearly indicate, however, the belief that low inventory changes workers' behavior.

Closer to the idea of the effects on motivation of low inventory is the work of Brown and Mitchell (1991). They surveyed direct labor employees at a manufacturing firm undergoing a transition to a JIT system over a period of a year and a half. They were particularly interested in performance obstacles in JIT. They considered the concept of the interdependence of work in low-inventory settings and the possibility of slower workers restricting output. Their survey pointed to a significant increase with problems of ''Schedules and Assignments'' and ''Reliance on Co-workers'' in JIT. Schedules and Assignments is a four-item measure, including ''There is too much pressure to do the job quickly.'' Reliance on co-workers is a five-item measure, including ''My co-workers work at a slow output rate'' and ''There is an inadequate match between coworkers' work speeds.'' The paper did not postulate any systematic change in worker performance associated with low-inventory systems. It did, however, demonstrate the importance of relative worker speed and that workers in JIT perceive different sets of performance obstacles.

To our knowledge there is only one paper that directly discusses worker speed in a low-inventory setting. Doerr et al. (1996) looked at the interaction of goals and buffer sizes on productivity in a fish packing plant. When collapsing their data across different goal treatments, they did not observe any significant difference in overall productivity between low- and high-inventory systems, despite increased idle time with lower inventory. While spending more time idle, workers in JIT would work faster when work was available. Doerr et al. (1996) thought that short breaks due to blocking and starving might allow increased effort during work periods. Unfortunately, they were unable to test this suggestion in their paper.

3. Hypotheses

We assert that low inventory motivates workers who are, on the average, slower than their teammates to speed up, and those who are faster to slow down. This causes an adjustment of work speeds toward a tighter band around the average pace of the group and could cause a reordering of relative work speed of members on the team. Low inventory increases feedback and the monitoring of worker speeds, as observed by Parker and Slaughter (1988), Sewell and Wilkinson (1992), and Delbridge et al. (1992). We postulate that this feedback is used by workers to monitor their own behavior and that of their peers. Certainly low inventory increases worker interdependence, as noted by Brown and Mitchell (1991), and this in turn increases the importance of this feedback to workers. This could account for some of the increased pressure on workers to increase speed as observed by Klein (1987), Parker and Slaughter (1988), Sewell and Wilkinson (1992), and Delbridge et al. (1992). However, as alluded to by Klein (1987) and Brown and Mitchell (1991), we expect that slower workers feel more pressure than faster workers. We believe this accounts for increased worker speed, even in the absence of management pressure, as measured by Doerr et al. (1996). Essentially, we argue that workers do not wish to cause idle time and will speed up, either momentarily or on average, to avoid doing so. Furthermore, workers who are about to be idle will slow down either to work at a more comfortable pace or to avoid the appearance of idleness. For a better discussion of the behavioral theory behind these effects see Schultz et al. (1997). We present two general hypotheses:

HYPOTHESIS HA. *Workers will work faster if they are likely to cause idle time*.

HYPOTHESIS HB. *Workers will work slower if they are likely to be idle*.

We will test these hypotheses with dynamic, static, and indirect tests. The dynamic and static tests are direct tests of two different aspects of the general hypotheses. In the dynamic case, we test the assumption that processing times are independent of the current inventory state of the system by looking at the dynamics of the processing time distributions. In the static case, we test the assumption that processing times are independent of the size of the buffers by looking at the means of the processing time distributions. The indirect tests will look at the assumption of independence from the speed of the co-workers and the overall effect on the productivity of the system.

The dynamic test uses the dynamics of the processing time distributions. We expect that workers will avoid causing idle time by shortening their processing times when they are currently causing others on the line to be idle. We further expect that workers will avoid being idle by increasing their processing times when they are in situations where they are close to becoming idle. In effect we are hypothesizing the existence of state dependent behavior, that processing times are not independent of the inventory state of the system. We use two forms of this case as a direct test of the two general hypotheses.

HYPOTHESIS H1A. *Workers will work faster when they are causing idle time than when they are not causing idle time*.

HYPOTHESIS H1B. *Workers will work slower when they are likely to be idle in the near future than when they are not likely to be idle soon*.

In a low-inventory situation with balanced tasks, the workers who are most likely to cause idle time are the slowest workers on each team. We therefore expect that they will work faster in a low-inventory situation to avoid causing idle time. Similarly, the workers who spend the most time idle are the fastest workers on each team. We expect them to work slower, on average, in a low-inventory situation than in a high-inventory situation. This could result from a general shift in the means of the processing time distributions: they speed up or slow down on all tasks. It could also result from the dynamic effects already discussed. For instance, if the slowest worker speeds up when causing idle time and slows down when about to be idle, his/her average speed would increase because (s)he is causing idle time more often than (s)he is idle. In effect we are hypothesizing that the means of processing times are not independent of the size of the buffers. We refer to these as *static effects* since they affect the means of the processing time distributions.

HYPOTHESIS H2A. *The average of the processing times of the slowest worker from each team on a low-inventory line will be faster than the average of the processing times of the slowest worker from each team on a high-inventory one*.

HYPOTHESIS H2B. *The average of the processing times of the fastest worker from each team on a low-inventory line will be slower than the average of the processing times of the fastest worker from each team on a high-inventory one*.

We also have four indirect tests of the combined effects of Hypotheses A and B. The first two look at idle time and productivity on the line. The second two look at the assumption that processing times are independent of coworkers. In the long run any serial processing line can work only as fast as the slowest machine, the bottleneck. Balance delay is caused when faster workstations become idle waiting for a bottleneck machine to catch up. The amount of balance delay is directly proportional to the difference in average work speed. The effects of Hypotheses 2A and 2B would be to reduce balance delay, since the difference in work speeds would be diminished, either by the faster worker slowing down or by the slower worker speeding up. Furthermore, since the production of a serial line is determined by the slowest worker, output of the low-inventory line will increase to the extent that the slowest workers speed up. Therefore we predict that idle time will be lower and productivity higher than would be predicted with a model using the independence assumption.

HYPOTHESIS H3. *Productivity of a low-inventory line will be higher than that predicted by a model which uses the independence assumption* (*result of general hypothesis A*).

HYPOTHESIS H4. *Idle time on a low-inventory line will be less than that predicted by a model which uses the independence assumption* (*result of general hypothesis A and/or B*).

Our general hypotheses imply that, in low-inventory situations, the average processing time of any one worker is a function of the processing times of the other workers on that team. If people feel pressure to work faster when they are the cause of blocking or starving, then we can expect the slowest member of the team to feel the most pressure and consequently to speed up the most. If people slow down when they are about to be blocked or starved, then we can expect the fastest people to slow down the most. In this way the motivation in a low-inventory situation is a function of speed relative to other workers on the line. One way to test this is by looking at the standard deviation of average processing times within teams. Another test would be to look at the correlation of worker speeds to the speeds of other workers in their team. Loosely speaking this would mean that if one worker on a team is known to have been faster than the overall average then it is likely that the other workers on that team would be also be faster than average.

HYPOTHESIS H5. *The standard deviation of mean processing times about the team average will be smaller in low- than in high-inventory situations* (*result of general Hypothesis A and/or B*).

HYPOTHESIS H6. *In low-inventory situations, a worker's average processing time will be positively correlated with the average processing times of other workers on the same team* (*result of general Hypothesis A and/or B*).

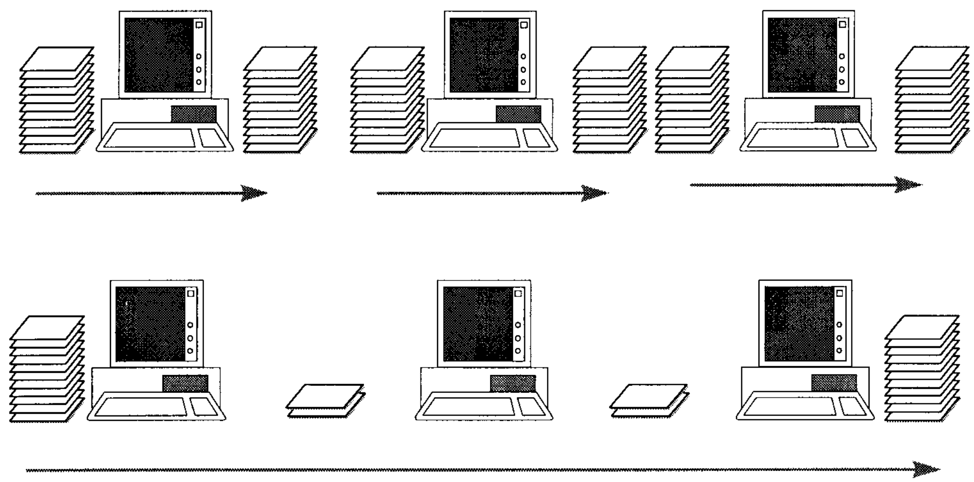
4. Method

We tested 66 subjects in groups of 3 in a laboratory setting. Subjects were volunteers from Ithaca High School and Dryden Central High School. They ranged in age from 14 to 19. High school subjects were chosen rather than college students because they more accurately match the target population of factory workers. We also felt that high school students would have less range restriction in salient behaviors and better reflect the population as a whole.

The task consisted of entering data, in the form of parts orders, into a computer. Subjects worked seated at a computer terminal as diagrammed in Figure 1. Each subject would begin an order by picking up a booklet of three pages from a buffer to their left. Turning to the page appropriate for their work station, they would enter the information, and put the booklet in an output buffer to the right. The data to be entered included an eight-digit order number and, randomly, either four or seven lines of information. Each line required selecting a name and a size with the mouse, and typing in a one- or two-digit quantity. When the subjects completed one line they would click a button on the screen and start on the next line. Lines were only slightly different among the three machines.

Macintosh computers were programmed using HyperCard to measure elapsed time for each subject to enter each order number (Order # Time), and each line (Line Time). By using only these two measures of processing time we avoid confounding the speed of entering data with the idle time between orders. These measures exclude small parts of the work process such as picking up and putting down a set of pages. Nevertheless they account for almost all of the nonidle time in the experiment. The computers also recorded the total number of orders completed by each workstation in

Figure 1 **Inventory Treatments**



each session. Because the three computers were synchronized at the beginning of each run we can use recorded data to determine the buffer contents at each point in time. For example, in the low-inventory case at the instant in time that workstation one has just finished the tenth unit, if workstation two is processing his/her eighth unit, then there are three units in the buffer and workstation one would be blocked (the buffer began the experiment with one unit). We will use these buffer states as independent measures in certain cases.

Personality and demographic variables were also measured. Personality was measured using the NEO Five Factory Inventory published by Psychological Assessment Resources (Costa and McCrae 1992). This is a self-report questionnaire in which subjects express their level of agreement, on a 5-point scale, with each of 60 statements. It reflects a five-factor model of personality description that has gained eminence in recent psychological research (see Goldberg 1990, for example). The five factors are neuroticism, extraversion, openness, agreeableness, and conscientiousness. Demographic variables included age, gender, and intention to pursue a college degree.

Each experimental run included three work periods: an individual practice to familiarize the subjects with the task, a treatment practice to familiarize the subjects with inventory procedures, and a production period. At the end of each of these work periods the computer would match data entered against a master list and provide subjects with a count of the number of pages done and the number done correctly.

There were two treatments: low and high inventory. The task and arrangement of machines were exactly the same in both treatments. All treatments also had, essentially, an infinite source of work in front of the first workstation and an infinite sink for completed work after the third workstation.

The two inventory treatments differed in the size of the intermediate buffers. In the low-inventory treatment, we restricted the intermediate buffers to a size of two, with blocking after service. That is, work stops on a machine upon placing a third item into the output buffer. In the high-inventory treatment intermediate buffers were, for all practical purposes, infinite and unshared. While we told subjects that the output from machine one would eventually become the input for machine two, there was no transfer during the course of the experimental run. Eleven runs of each of the two treatments were done.

Subjects were paid either a flat rate of $25 or a variable rate based on group performance. The variable rate was selected, based on pretests of the experiment, to average $25 per subject. Actual variable pay averaged $29 per subject. Each treatment had exactly the same number of subjects paid by each of the pay schemes. There are no significant findings with respect to pay in this paper.

5. Analysis and Findings

We will consider the overall effect on output first. The first section will look at Hypotheses 3 and 4, dealing with productivity and idle time. Section 5.2 will deal with Hypotheses 5 and 6, relating to the independence of the processing time means within workers on each team. The following section will deal with Hypotheses 1A and 2A, that workers will work faster under certain conditions. Finally, §5.4 tests Hypotheses 1B and 2B, that workers will slow down under certain conditions.

Personality and demographics were tested for correlation with both the treatments and processing time measurements (alpha = 0.1, two tailed). Intention to pursue a college degree and the personality measurements, neuroticism, extraversion, openness, agreeableness, and conscientiousness did not significantly correlate with either processing time or treatment. Gender and age were correlated with treatment but not with processing time measures; therefore, these measures were not used as covariates.

Prior to each t-test, normality was tested using both the Kolmogorov-Smirnov test with the Lilliefors statistic and the Shapiro-Wilks test. If the tests rejected normality then a log-transformation of the data was done. In cases where log-transformed data failed to reject normality, transformed data were used for the appropriate tests while descriptive statistics show data in their original form.

5.1. Effects on Productivity and Idle Time

Hypothesis 3 states that productivity of a low-inventory line will be higher than predicted by a model that uses the independence assumption. Productivity was measured as the average number of orders completed per unit time. A model was constructed using the independence assumption and a bottleneck assumption. Under the independence assumption, processing times are the same under high- and low-inventory conditions. This implies that we may use times from the high-inventory line to model the low-inventory situation. The bottleneck assumption states that the slowest worker is always a bottleneck. That is, the variation of processing times of the workers is never enough, compared to the difference in mean processing times, to cause idle time for the slowest worker. This is a deterministic model which predicts that the cycle time of the line equals the processing time of the slowest worker. Table 1 shows the mean and standard deviation of these predictions and the actual results for the low inventory workers. Not only is the difference statistically significant, in support of Hypothesis 3, but also it is large: this model underestimates the productivity of the low-inventory workers by 19%. *Production line models that use the independence assumption can significantly underestimate the productivity of low-inventory systems*.

Balance delay is defined as idle time due to differences in mean processing times (bottlenecks). Processing times from our experiment had a coefficient of variation of 0.3. If the line were balanced then, with three workstations and buffers of size two, we could expect idle time due to variation (variance delay) to be on the order of 5% with no balance delay. Since our line is unbalanced, due to differences in worker abilities, we can expect variance delay of 1% or less but significant balance delay. (Estimates are from Conway et al. 1988.) Using the bottleneck assumption, let the average processing time of the slowest worker, W_S, be the cycle time of the line. Using W_F as the fastest worker's average processing time, the predicted percent idle time (balance delay) of the fastest worker will be $(W_S - W_F)/W_S$. With a similar expression for the middle speed worker and zero idle time for the slow worker, the total fraction idle time for a team is:

$$\frac{(W_S - W_M) + (W_S - W_F)}{W_S}. \tag{1}$$

As explained, there was very little variance delay in the experiment and this computation, although a lower limit on idle time, should closely estimate total delay.

Table 1 **Predicted and Actual Productivity of a Low-Inventory Line**

Source	Mean	Std. Dev.	n
Predicted	74.7	16.0	11
Actual	92.0	19.0	11

Note. One-tail t-test with unequal variance: $t = 2.31$, $p = 0.02$.

Table 2 shows the mean and standard deviation of idle times predicted by this model using the high-inventory workers, compared to the actual results for the low-inventory workers. A log-transformation of the data was used. Again, the difference is quite large and statistically significant, supporting Hypothesis H4. The difference would increase only if variation delay were included in the computations. *Production line models that use the independence assumption can significantly overestimate the amount of idle time in low-inventory systems*.

Note that the actual idle time in the low inventory experiment is 9.7%, which is still a substantial loss of productivity. However, the slowest workers in the low inventory case also exhibited higher average processing speed, as will be seen later. Table 3 shows that, as a result, the low inventory lines achieved the same average output rate as the high inventory lines. (They differ by 0.7% in favor of the low-inventory line, but the difference is not statistically significant.)[1]

There are obviously something at work here. Through the remainder of this paper we present evidence as to what that may be.

Table 2 Predicted and Actual Idle Time of a Low-Inventory Line

Source	Mean	Std. Dev.	n
Predicted	16.1%	9.97	11
Actual	9.7%	7.33	11

Note. One-tail t-test, unequal variance, log-transformed data: $t = 1.96$, $p = 0.03$.

Table 3 Actual Productivity of High- and Low-Inventory Lines

Treatment	Mean	Std. Dev.	n
High-Inventory	91.3	24.2	11
Low-Inventory	92.0	16.0	11

Note. Two-tail t-test with unequal variance: $t = 0.09$, $p = 0.93$.

5.2. Interdependence of Mean Processing Times

Hypothesis 5 states that we expect low-inventory teams to have less within-team variation in average processing times. As a measure of within-team variation we use the standard deviation of the average processing times of the three workstations in each team. The hypothesis test is based on 11 observations of within-team variation for each of the 2 measures of processing time (Line Time and Order # Time) in each of the 2 treatments. A log-transformation of the data was used. As shown in Table 4, low-inventory teams have significantly less variation than high-inventory teams. Note that the "average" in this table is the average of the standard deviation of 3 average processing times for each of 11 teams. The term Std Dev refers to the standard deviation of those 11 measures. These data support Hypothesis 5.

Hypothesis 6 is based on the correlation of average processing times of workers on the same team. Table 5

Table 4 Average Between-Worker, Within-Team Standard Deviation

	Low-Inventory		High-Inventory		
Measure	Average	Std Dev	Average	Std Dev	Sig.
Line Time	47.4	29	133.1	126	0.01
Order # Time	115.1	52	176.9	95	0.05

Note. $n = 11$/cell, significance levels are one tailed, based on t-tests assuming unequal variance and using log-transformed data.

Table 5 Correlation and Significance (p values) of Average Processing Times

	High Inventory		Low Inventory	
	Order # Times	Line Times	Order # Times	Line Times
W1 with W2	−0.030	0.185	0.586	0.710
Sig. (one tail)	(0.47)	(0.30)	(0.03)	(0.01)
W1 with W3	−0.550	0.275	0.471	0.921
Sig. (one tail)	(0.04)	(0.21)	(0.07)	(0.00)
W2 with W3	0.015	−0.373	0.379	0.589
Sig. (one tail)	(0.48)	(0.13)	(0.13)	(0.03)

Note. $n = 11$/condition, e.g., per workstation. Numbers in parentheses are p values of the significance tests of the correlations.

[1] High inventory productivity data failed to reject normality using the Kolmogorov-Smirnov test but did reject normality using the Shapiro-Wilks test. A log-transformation failed to improve the distribution of the data, and the data in their original form were used.

shows the correlations of average processing times of workers on one machine with workers on another machine on the same team. For instance "$W1$ with $W2$" is the correlation of the processing times of workers on the first machine with workers on the second machine. Results are provided for the high-inventory case to provide a contrast. The results clearly reject independence for the low-inventory case. All six of the correlations are positive, as predicted, and four out of the six are significant at a level of 0.05 while one is significant at a level of 0.1. On the other hand, the test fails to reject independence for the high-inventory case. Exactly half of the correlations are positive and only one is statistically significant. An examination of scatter plots of the data from the high-inventory treatment also revealed no patterns. Additionally, when workers were asked if they adjusted their work pace to keep up with their co-workers, 58% said yes in the low-inventory situation as opposed to 13% in the high-inventory treatment. Together these analyses show that average processing times are not independent within teams in low-inventory situations, but may be with high inventory. This supports Hypothesis 6.

5.3. Workers will Work Faster if they Cause Idle Time

We now look at direct tests of general Hypothesis A. As described earlier we use both static and dynamic tests. Our first test looks at the people who will feel the most pressure for causing work flow disruptions, the slowest person on each team. Hypothesis 2A states that the slowest person on a low-inventory team will be faster, on average, than the slowest person from a high-inventory team. We tested for lack of normality as in §5.1. Normality was not rejected for Order # Times, but a log-transform was required to avoid rejecting normality for Line Times. Results are presented in the first line of Table 6. Line Times were, on average, 21% faster for the slowest workers in low-inventory compared to the same group in high-inventory. The difference for Order # Times was 12%. Both of these differences are significant and in the predicted direction. These data support Hypothesis 2A. *Processing times are not independent of the size of the buffers*.

Hypothesis 1A states that individuals work faster when they are causing idle time. For Line Times we compared average processing times when causing the physical idleness of a co-worker with the average processing times when not doing so. The approach for Order # Times was different. Because order numbers are the first thing done on a new page, they are never done while actually causing a neighbor to be idle. Therefore, for Order # Times we compared those times, which immediately followed the causing of idle time with those when the worker had not just caused idle time. Because of unequal cell sizes and large differences in variance a t-test was not used. However, a sign test is appropriate and was used. As can be seen in the last line of Table 6, results were significant. This supports Hypothesis 1A. *Processing times are not independent of the inventory state of the system*.

Table 6 Differences in Processing Times when Causing Idle Time

	Order # Times			Line Times		
	% Differ	Sig.	n	% Differ	Sig.	n
H2A: Static Test	12%	0.07	22	21%	0.03*	22
H1A: Dynamic Test	11%	0.00	30	3%	0.01	22

* Indicates log transformed data. H2A uses a one-tailed t-test assuming unequal variance. H1A uses a sign test.

5.4. Workers Will Work More Slowly if they Are Close to Being Idle

Finally we do direct tests of general Hypothesis B. We test these in the same manner as Hypothesis A, static and dynamic. Again the static test is a t-test, while the dynamic case uses a sign test. Normality tests, as described in §5, fail to reject normality for the static test. The results, shown in Table 7, are not statistically significant. Hypotheses 1B and 2B are not supported.

Table 7 Differences in Processing Times when Close to Being Idle

	Order # Times			Line Times		
	% Differ	Sig.	n	% Differ	Sig.	n
H2B: Static Test	0.1%	Not Sig.	22	4%	Not Sig.	22
H2A: Dynamic Test	8%	Not Sig.	13	(4%)	Not Sig.	31

Note. ( ) Indicates change is not in the predicted direction. H2A uses a one-tailed t-test assuming unequal variance. H1A uses a sign test.

6. New Interpretation of Data from Doerr, Mitchell, Klastorin, and Brown

A reanalysis of the data from the paper by Doerr et al. (1996) also supports general Hypothesis A. Their experiment used factory workers on a production line. While the low-inventory line had significant idle time, overall productivity did not vary significantly from the high-inventory line. This supports Hypothesis 3. Their summary data are shown in Table 8. *High* and *low* refer to high-inventory and low-inventory treatments, respectively. A line consists of one slitter, one gutter, and three slimers. Therefore, in order to present average cycle times per fish, processing times for the slimers are average processing time divided by three. Notice that, like our experiment, this line also is not balanced. For instance, slitters consistently take less time than slimers.

Doerr et al. proposed that workers might work harder to earn a break by being fast enough to become idle due to blocking or starving. If this is the case, we would expect to see the biggest reduction in processing time with those workers who are most likely to gain idle time when they successfully work fast. Slimers do not fit this description because, due to being the bottleneck in the line, they are unlikely ever to be idle. Yet it is exactly the slimers who show the largest decrease in processing times from high- to low-inventory settings. Similarly, in our experiment, the person who has the least chance of gaining idle time by working fast is the slowest person on the line. Yet it is exactly these people who show the greatest gain in speed. We conclude that working to earn idle time is not the primary motivation in this situation.

The data from Doerr et al. (1996) are consistent with Hypothesis A. Average processing times are consistently lower in low-inventory situations than in the corresponding high-inventory ones. The workers who speed up the most are those with the longest cycle time, the slimers (supporting Hypothesis 2A). Those who speed up the least are those with the shortest processing times, the slitters. In one case the workers with the shortest processing times actually take longer with lower inventory (slitters, group goal). For every goal treatment the standard deviation in low-inventory situations is less than the standard deviation for the corresponding high-inventory one (supporting Hypothesis 5). These data support hypothesis A, that workers who cause idle time speed up in low-inventory situations. They also give support to the inference of the hypotheses: in low-inventory situations workers adjust their speeds toward that of their co-workers. We also note that workers in the fastest position on the line appear to speed up more often than they slow down. We do not know if this is in contradiction to general Hypothesis B or is due to other causes, such as occasionally being the cause of idle time. All the findings shown are based solely on the summary data provided in Doerr et al.

7. Discussion

The overall effect of the changes in behavior associated with low inventory is to increase the productivity of the system above predictions made using the independence assumption. Under certain conditions processing times are not independent of the size of the buffers, the contents of the buffers, or the speed of co-workers. These effects are predictable, significant, and—as we will

Table 8 **Cycle Time Data from Doerr et al. (1996)**

	Total		No Goal		Group Goal		Individual Goal	
Position	High	Low	High	Low	High	Low	High	Low
Slitter	5.1	5.0	4.9	4.3	5.1	5.3	5.4	5.1
Gutter	7.4	6.9	6.8	6.4	7.1	6.9	8.4	7.2
Slimers	8.8	7.9	10.9	8.8	7.1	6.7	9.1	8.0
Average	7.1	6.6	7.5	6.5	6.4	6.3	7.6	6.8
Std. Dev.	1.9	1.5	3.1	2.3	1.2	0.9	2.0	1.5

argue—generalizable. While our experiment did show large amounts of blocking and starving with low inventory, idle time was less than expected. We also found compensating behavioral effects that decrease processing time and counteract the effect of the remaining idle time.

The effects of changes in inventory state on processing times are predictable in direction. Workers speed up when they are the cause of idle time. Apparently, they quickly and intuitively grasp the dynamics of blocking and starving and are conscientious enough to try to avoid being the cause of disruptions in the flow of work, even in the absence of management pressure.

The effects of changes in buffer size on line productivity are significant. Current production line models using the independence assumption predict significantly more idle time and lower productivity than actually occurred in our experiment. The magnitude was at least 19%. Due to the way the experiment was organized, we compared three machines running in a tightly coupled low-inventory system with the same machines running (in effect) in parallel. The production rates were essentially equal, despite the fact that the low-inventory system was highly unbalanced.

We believe that the effects described in this paper are generalizable. We used high school students as subjects because they were the available pool that most closely matched the general population. The behavioral theory here is projected to apply to the population as a whole, irrespective of age (see also Schultz et al. 1997). While the reaction might develop more slowly in older adults we believe the direction and general effect will be the same. Subjects in the Doerr et al. (1996) study were adult factory workers, and the findings were similar to ours. We feel that statistically significant results emerging after only 90 minutes of production speaks to the strength of the human reactions being measured here. However, further work is needed to demonstrate this effect in other populations.

8. Further Research

While we have shown how behavior can differ as a result of changes in inventory policy many questions remain unanswered. For instance, we have been interested in idle time and have made no distinction between blocking and starving. It is possible that the psychological effect of these two are different. Our results are inconclusive and of insufficient strength to present here. Workers causing blocking and starving seemed to react more to their input buffer, while those being blocked or starved seemed to react only to their output buffer. It is possible that the most salient buffer is a function of the relative speed of the worker. We have insufficient data to test this possibility.

Another unanswered question is the best line position for slower workers. Do we surround slower workers by the fastest workers available? Or does low inventory guarantee that effects of differences in work speed are felt more or less equally throughout the line? This is tied, in part, to the question above, because worker order will influence whether blocking or starving predominates. Bartholdi et al. (1995), for a special type of low-inventory system called bucket brigade lines, showed that the sequencing of workers in order from slowest to fastest provided certain benefits. McClain et al. (1997) found instances, in systems similar to bucket brigades, where slowest to fastest ordering is not optimal. Neither of these papers addresses behavioral issues or lines with workers dedicated to particular workstations, so the question remains unanswered for the general low-inventory case.

We could also ask if there is greater productivity from heterogeneous or homogeneous groups. Bartholdi et al. (1995) recommended heterogeneous groups for bucket brigade lines so that order from slowest to fastest is maintained despite some variance in worker processing times. The argument for heterogeneous groups in our context would be that they induce the most productivity increase from the slowest workers. However, when workers are actively trying to avoid being the cause of idle time, homogeneous groups might lead to higher output. Groups selected to minimize differences in worker processing speed would increase the chance that any worker could be the cause of the disruption. One way to avoid causing idle time in a stochastic environment is to work faster than average. If all workers follow this strategy in homogeneous groups, then productivity should rise.

It would be interesting to know how much blocking and starving are necessary to induce the behavioral benefits we have seen. Is there a way to design lines such that the idle time due to blocking and starving is minimized while the pressure to avoid causing idle time remains high? In the two experiments we have seen, the behavioral effect canceled the effects of blocking and starving. There is no theoretical reason why this should be the upper limit. It is theoretically possible to build a coupled, low-inventory line that actually outperforms a parallel arrangement. However, it may be that the threat of idle time, when rarely realized, loses its motivational effect. An experiment using different buffer sizes could be designed to help answer this question.

A larger issue deals with the long-term nature of the phenomena observed here. After months in a low-inventory work setting, do workers continue to adjust their work pace based on the state of the system? Can the slowest workers maintain the increased pace? This question should be addressed by a longitudinal study in a working factory.

9. Conclusions

A three workstation serial processing line with small buffers can have equal output to a line with the same workstations organized in parallel despite significant differences in idle time and imbalance in processing times. We present evidence that this is the result of the motivational effects of low inventory. In certain circumstances average processing times on low-inventory lines are not independent of the size of the buffers, the speed of co-workers or of the state the system.

Modelers and researchers of serial line systems should understand the differences in worker behavior between low- and high-inventory lines. The independence assumption appears valid in high-inventory situations and, of course, where worker motivation does not affect processing time. It appears invalid, however, in low-inventory situations where processing times are readily affected by worker motivation. We now know that worker speed depends on the state of the system, as represented by the workers' input and output inventories. Furthermore, their mean processing times depend on with whom they are teamed. Unfortunately, the magnitude of the effect is likely to be heavily influenced by the idiosyncrasies of each situation; therefore no quantitative relationships have been established that could be used in a simulation or stochastic models. Nevertheless, modelers need to be aware of the limitations of the independence assumption and analyze their conclusions appropriately.

We have not proved that factory workers on actual manufacturing lines will act differently in low- vs. high-inventory settings. However we have presented evidence to support this conjecture. Socio-Technical Systems Theory argues that changes in the technical aspects of work, such as a reduction in buffer size, can affect the cultural and political environments. We have argued that inventory reduction increases feedback and interdependence, which can effect worker motivation. We have behavioral theory that supports this hypothesis (Schultz et al. 1997). Finally, we have two examples where these effects were measured (this paper and Doerr et al. 1996). Opposed to this evidence we know of no behavioral theory or empirical evidence which demonstrates that workers do not change their behavior. Nevertheless, we believe that both conditions exist, and we look for further research to define the conditions that separate the two.

Mathematical modeling of production lines has led to important insights. The independence assumption has been an essential part of this work, and we do not advocate discontinuing its use. Rather we believe that by understanding its implications we can improve our models and better translate model implications into reality. Frey (1996) compares the art of modeling to an hourglass. ''Reality is big and broad and messy. By necessity, it (modeling) narrows that reality through assumptions and simplifications, and, at the neck of the hourglass, preliminary modeling results and insights are developed. The modeling effort should not be left at that point, but expanded back out to the reality in which the results and insights must be implemented.'' Independence is, and will continue to be, a valuable tool for reducing reality to a manageable size. We believe this paper will add insight into how to translate those results and insights into usable answers for managers.

Managers should be aware of the full range of motivating factors in the workplace. These factors include social pressures and the influence of peers. It appears that, even in the absence of management pressure, workers do not wish to be identified as being the cause of idle time on the line. There are many factors present in factories which could overshadow the effects noted here, worker antagonism with management goals or strong informal leadership on the line for instance. Nevertheless, by understanding the underlying motivations induced by the work system, managers can make better decisions.[2]

[2] The authors wish to thank the Clifford H. Whitcomb Faculty Fellowship and the students and staff of Ithaca and Dryden Central High Schools for their support of this research.

References

Bartholdi, J. J., L. A. Bunimovich, D. D. Eisenstein. 1995. Dynamics of 2- and 3-worker ''Bucket Brigade'' production lines. Working Paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Brown, K. A., T. R. Mitchell. 1991. A comparison of just-in-time and batch manufacturing: the role of performance obstacles. *Acad. Mgmt. J.* **34** 906–917.

Conway, R. W., W. L. Maxwell, J. O. McClain, L. J. Thomas. 1988. The role of work-in-process inventory in serial production lines. *Oper. Res.* **36** 229–241.

Costa, P. T., Jr., R. R. McCrae. 1992. *Revised NEO Personality Inventory* (*NEO PI-R*). Psychological Assessment Resources Inc., Odessa, FL.

Cranny, C. J., P. C. Smith, E. F. Stone. 1992. *Job Satisfaction: How People Feel About Their Jobs and How it Affects Their Performance*. Lexington Books, New York.

Dallery, Y., S. B. Gershwin. 1992. Manufacturing flow line systems: a review of models and analytical results. *Queuing Systems* **12** 3–94.

Delbridge, R., P. Turnbull, B. Wilkinson. 1992. Pushing back the frontiers: management control and work intensification under JIT/TQM factory regimes. *New Tech. Work and Employment* **7** 97–106.

Doerr, K. H., T. R. Mitchell, T. D. Klastorin, K. A. Brown. 1996. Impact of material flow policies and goals on job outcomes. *J. App. Psych.* **81** 142–152.

Frey, S. C., Jr. 1996. Darden School of Management, University of Virginia, Charlottesville, VA. (private communication).

Goldberg, L. R. 1990. An alternative ''description of personality'': the big-five factory structure. *J. Personality and Social Psych.* **59** 1216–1229.

Groebner, D. F., C. M. Merz. 1994. The impact of implementing JIT on employees' job attitudes. *Internat. J. Oper. Production Management* **14** 26–37.

Huber, V. L., K. A. Brown. 1991. Human resource issues in cellular manufacturing: a sociotechnical analysis. *J. Oper. Management* **10** 138–159.

Johnson, T. W., G. H. Manoochehri. 1990. Adopting JIT: implications for worker roles and human resource management. *Indust. Management* **32** 2–6.

Klein, J. JIT, SPC and teams. Working Paper 87-058. Harvard Business School, Boston, MA.

McClain, J. O., K. L. Schultz, L. J. Thomas. 1996. Worksharing Systems. Working Paper 96-05. S. C. Johnson Graduate School of Management, Cornell University, Ithaca, NY.

Parker, M., J. Slaughter. 1988. *Choosing Sides: Unions and the Team Concept*. South End Press, Boston, MA.

Robinson, L. W., K. B. Hendricks. 1995. Using state-dependent processing rates to emulate SPT queue discipline in an FCFS queueing network. *Trans.* **27** 530–541.

Schonberger, R. J. 1982. *Japanese Manufacturing Techniques: Nine Hidden Lessons in Simplicity*. The Free Press, New York.

Schultz, K. L., D. C. Juran, J. W. Boudreau. 1997. The effects of JIT on the development of productivity norms. Working Paper 97-17, Center for Advanced Human Resource Studies, Cornell University, Ithaca, NY.

Scott, A. F., J. H. Macomber, L. P. Ettkin. 1992. JIT and job satisfaction, some empirical results. *Production and Inven. Management J.* 36–41.

Sewell, G., B. Wilkinson. 1992. Someone to watch over me: surveillance, discipline and the just-in-time labour process. *Soc.* **26** 271–289.

Shanthikumar, J. G., D. D. Yao. 1989. Monotonicity of concavity properties in cyclic queueing networks with finite buffers. *Queueing Networks with Blocking*. H. G. Perros, T. Altiok, eds. North-Holland, Amsterdam, 325–345.

Snell, S. A., J. W. Dean, Jr. 1992. Integrated manufacturing and human resource management: a human capital perspective. *Acad. Management J.* **35** 467–504.

Tichy, N. 1983. *Managing Strategic Change: Technical, Political and Cultural Dynamics*, Chapter 1. John Wiley, New York.

Accepted by Abraham Seidmann; received January 21, 1997. *This paper has been with the authors for* 2 *revisions.*