Lean Manufacturing Principles: A Comprehensive Framework for Improving Production Efficiency

by

Auston Marmaduke Kilpatrick

B.S. Mechanical Engineering, University of California, Los Angeles
B.A. Philosophy, University of California, Los Angeles

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

Massachusetts Institute of Technology

February 1997

Author.....
Department of Mechanical Engineering
January 17, 1997

Certified by.....
David S. Cochran
Thesis Supervisor

Accepted by.....
Ain A. Sonin
Chairman, Departmental Committee on Graduate Students

Lean Manufacturing Principles: A Comprehensive Framework for Improving Production Efficiency

by

Auston Marmaduke Kilpatrick

Submitted to the Department of Mechanical Engineering
on January 30th, 1997 in partial fulfillment of the
requirements for the Degree of Master of Science in
Mechanical Engineering

ABSTRACT

A framework was created to analyze manufacturing systems and assess the impact of various practices on system performance. A literature review of Lean Manufacturing resulted in the discovery of significant gaps in two areas: (1) modeling the effects of implementing Lean Manufacturing using control theory principles, and (2) a design framework for building Cellular Manufacturing Systems and making the transition from traditional manufacturing to Lean Manufacturing. Work in these areas led to the conclusion that reducing the Order Lead Time until it is less than T_{all}, the allowable customer lead time for post-payment production, would yield tremendous benefits both for individual factories as well as for entire Linear Distribution Systems.

To fill these gaps, a model was created which analyzed the dynamics of Linear Distribution Systems, and showed how Lean Manufacturing represents an opportunity to sidestep many previously insurmountable difficulties that arise as a result of producing to fill inventory levels. The methods for implementation of Lean Principles were explored, from prerequisites for Cellular Manufacturing Systems, to design and optimization of Cells, through exploration of the improvements in quality that are possible in Cellular Manufacturing Systems. A thorough a dissemination of the various contributions to Order Lead Time showed that changeover reduction, information flow, zero defects and cellular manufacturing are all indispensable in achieving the goal of $OLT < T_{all}$. Finally, conclusions were presented which show that achieving this reduction allows for production under an entirely new philosophy that completely eliminates capital investment in inventory.

Thesis Supervisor: David Cochran
Title: Assistant Professor of Mechanical Engineering

Acknowledgments

I would like to thank Professor David Cochran for his leadership and guidance in helping me find my niche in lean manufacturing. Without his help I would not have been able to learn the intricacies of the Toyota Production System. I would like to thank Tom Shields for his continued support and for giving me the opportunity to work on the Lean Aircraft Initiative where many of the ensuing ideas were created. I would like to thank the members of the Production System Design (PSD) laboratory who kept me honest in my manufacturing idealism. I would like to thank Christian Marks, and all my other friends who kept me sane while I was lost in the world of Takt times, Poka-yoke, and System Dynamics. Finally I would like to thank Mike Krawczyk at Briggs and Stratton whose insight and ideas helped me generate the cell design methodology, and helped me to see how to apply lean manufacturing outside academia.

Table of Contents

Acronyms and Symbols

ABC- Activity Based Costing
AD- Annual demand
AICCP- Annual inventory carrying cost percentage
BN- Bottleneck machine
CM- Cellular Manufacturing
CMS- Cellular Manufacturing System
COP- Cost of order preparation
E- A theoretical lot size which minimizes the sum of inventory costs and setup effects
EOQ- Economic order quantity
FMS- Flexible Manufacturing System
GT- Group Technology
LDS- Linear Distribution System
LS- Lot size
LT- Lead time
MCT- Machine cycle time
MLT- Manufacturing lead time
OLT- Order lead time
PV- Priority value
SPF- Single piece flow
TPS- Toyota Production System
UC- Unit cost
WIP- Work in process
X- An order quantity for a single type of part

In Chapter 2's Control model
DU- Defective units
GUP- Good units produced
FGI- Total finished goods inventory (including goods in pipeline)
FGIN- New quantity of finished goods inventory
FGIO- Old quantity of finished goods inventory
IO- Incoming orders
IHO- In-house orders
K_c- Kanban container capacity
K_d- Disturbance coefficient
KQ- Kanban quantity
PRMA- Previous raw materials available
QP- Quantity to produce
Q_{ro}- Reorder quantity
RMI- Raw materials inventory
RQ- Required quantity
SI- Standard day to day inventory quantity
t_r- Time period between removal of containers by shipper

TRMA- Totai raw materials available
T_s - The period of one shift
TSI- Total standard day to day inventory quantity (including pipeline quantity)

Introduction

The difficulties that companies face in today's marketplace are fierce: shifting customer demand, increasing variation in products and demands for perfect quality. Meeting these demands while dealing with complex distribution systems and multi-tiered chains of suppliers is better understood in light of system dynamics (Forrester [1969]) and finding ways to minimize their cyclical nature. The way to escape the pitfalls faced by aerospace companies today requires a redefinition of inventory and a new production philosophy which eliminates the need to produce based on forecasts, or to fill stock levels, and to eliminate rework and acceptance of non-conformances. This thesis presents the tools necessary to make this leap.

Chapter 1 presents a review of the literature followed by a taxonomy which serves to clarify some issues which are integral to understanding lean manufacturing and which have been misunderstood in the past. Chapter 2 introduces the system dynamics problems that are faced by nearly every manufacturing plant in the world, and shows how fluctuations in customer demand create cyclical demand patterns which are amplified at each link in supply chains. It concludes by postulating lean manufacturing principles as a solution to many of the difficulties, and suggests that a truly "lean" factory can deal with the variation in customer demand without the high levels of inventory that are common in most factories today.

Lean Manufacturing is a term popularized by Womack, Jones and Roos [1991] to describe a method for production based on the Toyota Production System(TPS) (Shingo 1989, Ohno 1988). There is a tremendous body of literature available on the details of

the TPS. The purpose of this thesis is to take the principles enumerated in the literature one step further and show how lean manufacturing can address the difficulties of aircraft manufacturing as a linear distribution system (LDS). The LDS is driven by the "evils of inventory", which are enumerated in chapter 3.

Toward these ends, a description of the building blocks for a production system are given. Chapter 4 addresses the issue of setup reduction which is an enabler for cellular manufacturing (CM), the topic of Chapter 5. Cellular manufacturing is described in great detail in the literature and nearly all plant managers will claim to produce product in "cells". However, a comprehensive look at the fundamentals of CM as well as a design methodology to create a cellular manufacturing system (CMS) have not been published. The aim of Chapter 5 is to introduce the reader to the benefits of CM, and show the preliminary steps that are necessary to gain the full benefits of cellular manufacturing. Chapter 6 addresses the issue of quality and shows how to greatly improve the quality of parts that reach the customer and reduce the costs of internal defects. Chapter 7 enumerates how a lean manufacturing system, using CM, a Kanban controlled "pull" system, and zero defects can meet demand in today's customer driven market. The concepts in this thesis can apply to any manufacturing environment. However, here they are tailored to linear distribution systems found in the aerospace industry.

A large body of literature exists on Lean Manufacturing. In the bibliography there is a list of references on lean topics. In general, most texts develop only one aspect of lean manufacturing, and briefly touch on other areas. I have listed two strong texts on Poka-Yoke devices, Shingo [1986] and Hirano [1988], one on single piece flow

manufacturing, Sekine [1990], two on set-up reduction, Shingo [1985], and Smith [1991], one on the effects of inventory, Shingo [1988], and several others on system level approaches to Lean Manufacturing which do not go into detail on individual subjects. In addition, I have listed a number of academic papers published in periodicals which go into great detail on just one aspect of Lean Manufacturing, such as scheduling parts in cellular manufacturing. It is the aim of the author that this thesis will provide a strong introduction to the concepts of lean manufacturing, and encourage the reader to investigate further into each area that they encounter in transforming their factory from a traditional mass producer, or craft shop into a lean manufacturing production system.

Chapter 1. Literature Review and Taxonomy

1.1. Literature Review

A literature review of past and present journals and books (past papers, plus a published article scan for 1993 - 1996) was conducted. The subject and title/keyword searches included: cellular manufacturing, Group Technology, flexible manufacturing systems, single piece flow, and Toyota Production System. Author searches were also conducted for books and articles by Shingo, Monden, Black, and Ohno. The search produced roughly seventy five articles, and 13 books. The topics of the articles and books can be separated into the following categories:

1) General descriptions of Group Technology
2) General descriptions of the Toyota Production System
3) Comparisons of Group Technology and traditional flow shops
4) Control of factories using Group Technology
5) Simulations of production using Group Technology layouts vs. functional layouts
6) Algorithms for scheduling products/parts produced using Group Technology

The largest portion of literature in the earlier papers was dedicated to first glances at Group Technology: attempted formulations of the guiding philosophies of group technologies, metrics to measure production in Group Technology layouts versus functional layouts, and comparisons of simulations. In general there were several different pictures of what "Group Technology" represented. None of the papers surveyed

grasped the multitude of views. It is the author's opinion that a review article to assemble them, and point out the differences , as well as create a taxonomy so that a common language is used in future publications would be very beneficial. Several articles made requests for a common taxonomy, but none put one forth. A literature review that outlined the current opinions as to what Group Technology encompasses, as well as the metrics that are commonly used to compare Group Technology layouts to functional layouts would be useful in making data from present and future publications at least similar in nature. The past literature used a multitude of many different measurables which made comparison of simulations either difficult or impossible.

A common experimental methodology would also improve the quality of simulations. For example, some authors ran simulations until they appeared to reach steady state (starting from an empty factory) before taking data, while others chose arbitrary starting points, either after 100 days or at startup, or with an arbitrary initial loading sequence. Another disparity among many of the simulations is whether or not the authors considered the length of changeover time to be sequence dependent, which is clearly the case when using Group Technology methods for individual parts, or families of parts. The number of machines and different parts used in each simulation also varied widely making the results of each simulation incomparable with the others (although several attempts were made). In general, since simulation technology is rapidly developing, allowing more and more complexity to be added, it would be counter-productive to fix the simulation parameters (number of machines, number of part types, operation length, etc.) but the past publications' habit of cross-comparing simulations that

are not comparable does little except add confusion to the Group Technology-job shop debate.

When simulating production using Group Technology and comparing it to a shop with a functional layout, one must be careful to create realistic working conditions in order to see which layout produces better results. Several journal articles have based their simulations on actually existing factories. This serves two purposes: 1) it makes the simulation believable, rather than just being an example created to support the author's personal beliefs, and 2) it will be useful to the actual company whose factory is being simulated. This leads to a further point, and one that would be of great value to manufacturing as a whole. Many articles are concerned with simulations and some even have links to real factories, however, none of the articles presented before and after production data from a company as well as data from simulations based on the same company. In general, all the simulations lacked supporting evidence in the form of comparisons to actual production data.

A common complaint from authors of case studies was that the changes in various production metrics could not be directly traced to their source (i.e. group technology, layout changes, etc.) This is a result of the lack of understanding of the causal mechanisms between implementation of lean principles and system response. Chapter two will address this issue. In the future, case studies with simulations that show the effect of changing a) the layout, b) the scheduling, c) the control philosophy and technology and d) the grouping of parts/machines, individually and in various combinations would illustrate the effects of each and serve to show causation rather than correlation. Also, a simulation that produces results which are reasonably close to actual

production data from a factory in the given loading conditions would lend great support to that simulation, as well as to the methods used to create that simulation.

From the literature, it appears that it is unclear as to which are the important elements that must be included in simulations in order to make them valid in their characterization of production systems. A standard from which everyone could start would prove invaluable in improving the quality of simulations. To improve the validity of simulations (which have made up a large portion of the published articles on Group Technology and cellular manufacturing in the past), a comparison to production data from the plants upon which the machine layout is based should be made.

A further difficulty with the simulations that have been produced is the following: most authors, even those who advocate implementing Group Technology, do not suggest that all machines should be dedicated to part families (a maximum figure of 60 - 70% is suggested by several authors). However, none of the simulations that were surveyed compared a factory using a job shop layout to a factory with 60% of the machines dedicated to product families and the others left in a functional layout. It is highly likely that this hybrid factory would behave quite differently compared with a job shop or a flow shop (or a shop exclusively using cells). It is possible that the complexity of simulating a hybrid layout has deterred previous efforts, but certainly given the computational ability available today, a simulation of this sort is quite feasible This is a further effort to make the factory simulations more closely linked to the operation of the actual factories, since the purpose of simulations is often to decide if a company should adopt the Group Technology approach to manufacturing and to what extent.

Numerous algorithms for scheduling a factory using Group Technology have been published. Very few of these articles make comparisons to other existing algorithms, making it difficult for members of industry to make sense of which one produces the best results. For example, if all algorithms were put to a standard test (a sample factory, or a number of sample factories representing various sizes and manufacturing techniques) then one could see which produced the best results for the sample factory that most closely matched that manager's factory. Some common measurables that have shown to be of interest are throughput, average tardiness, overtime required to finish a given quantity of parts, standard deviation of job tardiness, and machine utilization. Many others were given in the literature pointing to the need for a standard set of metrics. Another common complaint of many of the authors was that some positive effects of Group Technology, such as improved quality, ability to track defects to their source, and simpler scheduling of Group Technology layouts, are not included in simulations or comparisons of job shops and factories using Group Technology methods. These advantages are often stated, but rarely quantified, and none of the articles reviewed made any efforts to include these factors in their simulations. With the latest development in cost of quality (COQ) systems, it is becoming possible to quantify the effects of improved quality. A method which quantifies the advantage of being able to track defects back to their source in manufacturing cells (and find their cause) based on the number of defects that would have resulted had the faulty machine continued producing would serve this purpose. In addition, an effort to quantify the value of implementation of poka-yoke devices and other mistake-proofing methods which cannot be done in job shops, but are very feasible using Group Technology methods of manufacturing would be beneficial.

In conclusion, there exists a large body of literature from the last five years comparing Group Technology, and cellular manufacturing, to traditional or job shop production. However, as a whole, the effort is disjointed. There is a lack of a common taxonomy. For example, in many articles Group Technology is equated with cellular manufacturing while in others cellular manufacturing is presented as just one aspect of GT. In addition, while numerous simulations of factories have been published, none show comparisons to actual manufacturing data or make use of real data (machine changeover times, machine cycle times for given parts, etc.). In short, there is much improvement that can be made in the efforts to explore the benefits of these new production methods. It is the hope of the author that a widely published call for solidarity can improve many of these shortcomings.

1.2. Taxonomy

Cellular Manufacturing (CM)- Processing of parts or part families in a single cell, with no backtracking. Each part follows a predetermined part path, and is said to be trackable. Individual cells may be built around Group Technology formed part families or based on a single product line (which facilitates APC methods). In CM the operators move independently of the part processing time. A cell processes a subset of the total operations in the production system. Cells process parts sequentially through a series of machines or manual stations, whereas a job shop processes parts in parallel processing through duplicated machines. A *cellular manufacturing production system* is network of logically linked cells (Cochran [1994]). Hence throughput of cells can be dependent on

(a) a single machine's cycle time (the bottleneck), or (b) the speed of the material handler or operator.

Changeover Time- The length of time a machine or processing area is down (not producing parts) in order to change over from one part type to another. It begins when the machine finishes producing the last part of type A and ends at the beginning of the production of the first good unit of type B.

Cost of Order Preparation- The costs incurred by a factory in the process of receiving an order, releasing it to the floor and closing the order upon receipt by the customer.

Customer Lead Time- The length of time beginning when a customer places an order and ending when the customer receives goods to fill that order. It encompasses the Production Lead Time plus the time required to transmit the order to the factory and the time to ship the product to the customer.

Defects vs. Errors- a defect is a product or service's non-fulfillment of an intended requirement or reasonable expectation for use. An error is the result of improper processing of the part, which leads to a defect only when the part is inspected and fails. Thus, errors can be fixed before they become defects.

External Setup Operations- Operations in changeover from one part type to the next which are done while the machine is still producing parts.

Final vs. Absolute Defect Rate- the final defect rate reflects only the number of parts that fail the final inspection. However, the absolute defect rate is a measure of the number of defects for the entire process from start to finish (see Shingo [1986])

Flexible Manufacturing System (FMS)- A group of numerically controlled machine tools interconnected by a central control system. The various machining cells are

interconnected via loading and unloading stations by an automated transport system. Parts are generally processed in parallel rather than serially as in a Cellular Manufacturing System. Operational flexibility is enhanced by the ability to execute all manufacturing tasks on numerous product designs in small quantities. However, there are often limitations in flexibility due to the hard tooling that is required for material transport and fixturing in each machining cell. Capital investment is very high as compared to a manned cell.

Group Technology (GT)- A part classification technique used to categorize parts according to one of two possible similarities: (1) part geometry, or (2) processing similarities. GT is generally used to create part families which can then be processed in single cell.

Informative vs. Subjective Inspections- An informative inspection yields information about a part, such as a dimensional characteristic. where as a subjective inspection is a pass-failure inspection where the inspector must make a subjective judgment about the quality of a part.

Internal Setup Operations- Operations in changeover from one part type to the next which are done when the machine is stopped (is not producing parts).

Job- Work order; an order from a customer which comprises a fixed number of parts of a specified type.

Lot Delay- In batch production, the time spent while the part is waiting for other parts in the batch to be processed.

Lot (or Batch) Production- Producing parts in lots or batches which are greater than one part.

Machine Utilization- The ratio of time during which the machine is operating to the total time available. The operating time should only include time in which product is being produced (setup time is excluded).

Machine Cycle Time- The length of time required for a machine to process one part, not including loading and unloading time. It can be measured as the length of time beginning when the start button is pressed and ending when the part can be removed.

Manufacturing Lead Time- The length of time from the first operation for a given order until the entire order has been transported to the shipping area.

Order (or Production)Lead Time- The length of time beginning with receipt of an order from a customer and ending when product for that order has been manufactured and transported to the shipping area. It includes the Manufacturing Lead Time plus the time required to process the order and begin production.

Process vs. Operation- The distinction between a process and an operation is not one of time scale; the two have different subjects of study. A process is a flow of *product* from raw materials to finished parts. Operations are the actions of *man*, or *machine*, and what they do to the product.

Process Mapping- A systematic analysis of a production system using Time Division Analysis to determine which operations are integral to a process, and which must be eliminated.

Processing Delay- The time spent while a part is in storage, either in queue behind other batches of parts, or waiting due to machine downtime.

Production Balancing- Efforts to make the processing time for a given part at each station in a process equal. In a perfectly balanced process, no stations will be idle when producing parts using single piece flow.

Production Leveling- Efforts to increase the product mix and decrease the batch size in a manufacturing system. In *Level Production*, assuming there is no variation in processing, at least one unit of each part type is produced each day.

Production Synchronization- Efforts to synchronize the start, stop and transport of product at each machine, station and process. In *Synchronous Production*, the upstream process produces goods at the same rate (or Takt time) as the downstream process.

Self (or Source) Inspection- Inspection carried out by the person (or factory) which produced the product.

Single Piece Flow- A term describing the processing of parts in a batch size of one. The only processing lot size in which the lot delay is reduced to zero.

Statistical Quality Control (SQC)- The application of statistical techniques to control quality. Generally refers to monitoring or inspecting some percentage (less than 100%) and inferring the quality of the rest of the parts from a sample. SQC is subject to α and β errors.

Successive (or Post-receipt)Inspection- Inspection by a person or factory of goods that were produced at an upstream station or plant. Traditionally thought to be more objective than self inspections since there is no incentive to pass defective parts through the inspection. However, this inspection method can only discover defects after they have been made (see defects vs. errors).

System- A regularly interacting or interdependent group of items forming a unified whole toward the achievement of a goal.

Takt Time- A production rate determined by customer orders (or sales) which specifies the interval of time between production of successive parts. Is determined from the following equation:

$$\text{Takt Time} = \frac{\text{total time available for production per shift (in sec)}}{\text{required number of parts per shift}}$$

Time Division Analysis- A technique similar to process mapping where a process is analyzed by tracking a part as it flows from raw materials to the finish crib and constantly specifying whether a part is being processed, transported, stored or inspected. Storage is further broken down into Lot delay, Transportation delay, and Processing delay.

Transportation Delay- The time spent waiting for the means of transportation (i.e. waiting for a forklift).

Value Added vs. Non Value Added Processing- Value added processing refers to processing steps that will add value to a part, viewed from the eyes of the customer. Non value added processing includes all processing that does not add value. Thus a cutting operation on a machine is value added processing only if it creates a feature in the part that is of value to the customer.

Volume vs. Part Types Flexibility- Volume flexibility refers to the ability of a manufacturing system to produce parts at different rates. Part type flexibility refers to the ability to produce a number of different part types in a given period of time.

Chapter 2. Linear Distribution Systems: How Lean Manufacturing Can Improve System Performance

2.1. Introduction

Linear distribution systems (LDS) can be used to model (1) a single part flow path in a plant (2) the link between two manufacturing plants, (3) an entire manufacturing system from mining for raw materials to sale to the customer at a retail store, or (4) any segment of (1)-(3). The difficulties associated with trying to meet demand and manage inventory for each member of the system are described in the literature (Forester [1965], Forester [1968], Senge, Sterman [1992]) but not in the context of lean manufacturing. This chapter will address the most common difficulties of system dynamics, and show how implementation of lean manufacturing principles can aid in dealing with these problems.

To begin we will define a linear distribution system (LDS). A LDS is a group of production or distribution organizations in which each organization depends on the upstream supplier for product, and sells product to the downstream customer(s) (see Figures 2.1 and 2.2).

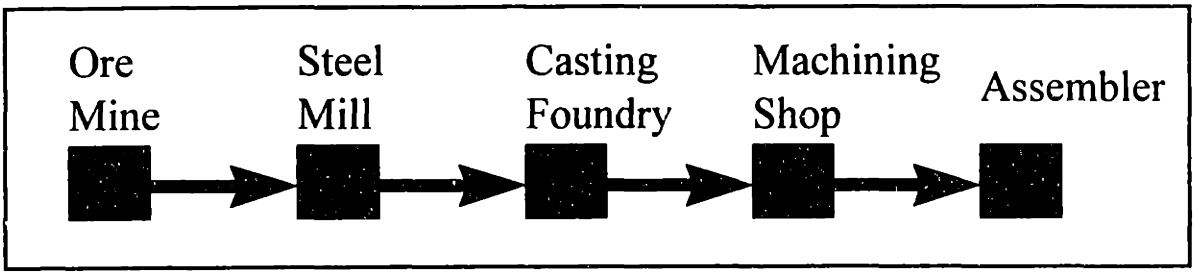


Figure 2.1

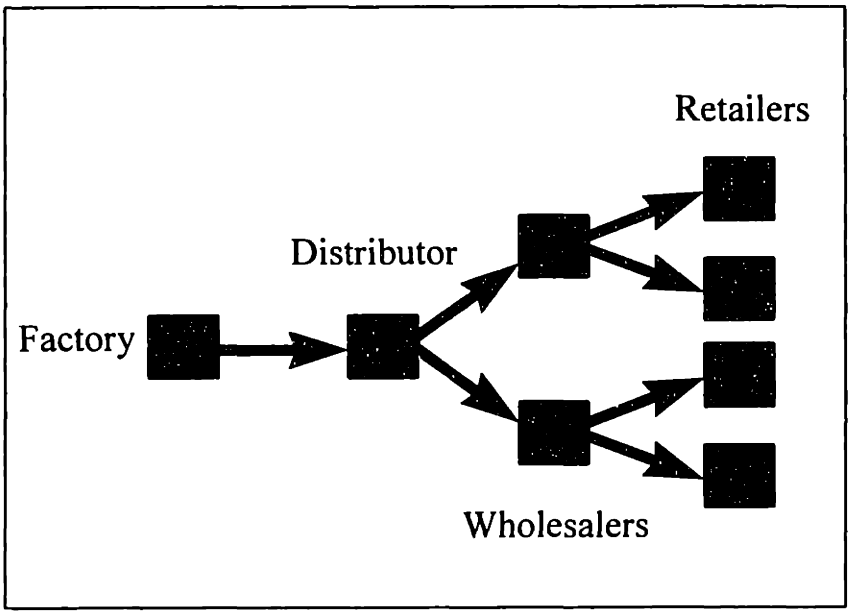


Figure 2.2

Figures 2.1 and 2.2 can be linked by overlaying the assembler with the factory to produce one long LDS. One can think of an LDS as a chain made up of links, each of which is an individual company. Similarly, an individual plant can also be modeled as an LDS. In this case the plant represents the entire production system chain, and each individual machine or processing area represents one link. The level of detail in the LDS model can be specified by the user. However, inputs, outputs and the form of disturbances will be different depending on the level of analysis. In this paper we will begin by looking at the interaction between any two members, or links of the chain.

Then we will show how the effects of these interactions propagate through the system and create serious difficulties. Finally, we will look at how Lean Manufacturing addresses these problems, and compare the responses of several LDS to various demand inputs.

2.2. Managing Inventory in an LDS

The goal (or functional requirements) of the manager of each link of an LDS is to meet the demand of the downstream customer while minimizing total cost (one component of total cost is the cost of inventory, which will be discussed in detail in chapter 3). A common claim from managers in LDS is that meeting the variation in customer demand requires carrying large inventories of raw materials, finished goods and work in process inventory (WIP). However, these large inventories are seldom successful in eliminating backlogs (being unable to meet the customer demand). We will explore why the inherent nature of linear distribution systems makes it difficult to respond to fluctuations in demand.

There are two phenomena that drive the difficulties that LDS managers face. (1) A time delay between the time an order is placed to the supplier and the time the shipment is received to fill that order. This is known as the Order Lead Time (OLT), which starts when the downstream customer orders a quantity of goods, includes all the time spent manufacturing the products (the manufacturing lead time (MLT)) by the upstream supplier, and ends when the goods have been received by the downstream customer. (2) The lack of (accurate) information flowing between any two organizations that do not have a supplier-customer relationship (and even many that do). In most LDS, two

organizations that do not have a supplier-customer relationship (are not adjacent to one another in Figures 1 and 2) will have no communication whatsoever.

The time delay creates a cyclical behavior that is inherent in LDS which will be explored next. The result of the second factor is that in order to meet periodic fluctuations in demand at the downstream end of the chain, suppliers often forecast what quantity of goods their downstream customer will order. Since it is impossible to accurately forecast what the customer's next order will be (even with years of experience), production based on forecast results in large inventories and frequent backlogs.

2.3. A Control System Model

In order to understand the dynamics of LDS, we will model the production system chain using control theory. To begin we will model one link of Figure 2.1 as is shown in Figure 2.3.

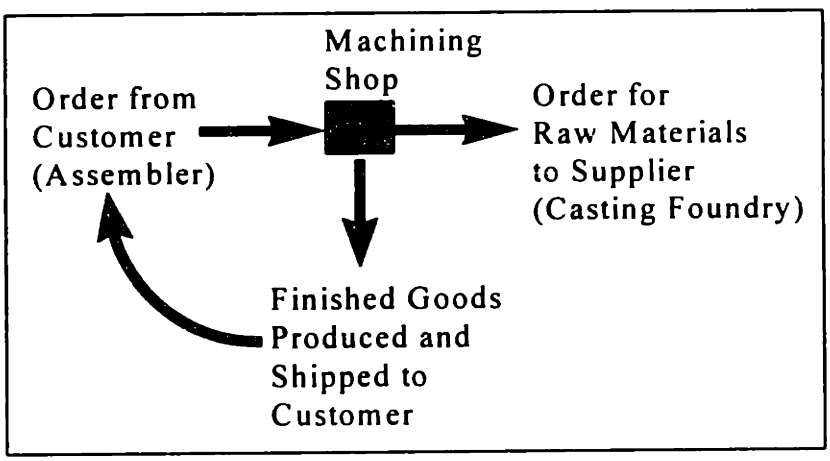


Figure 2.3

Figure 2.4 shows the control system model of the link in Figure 2.3.

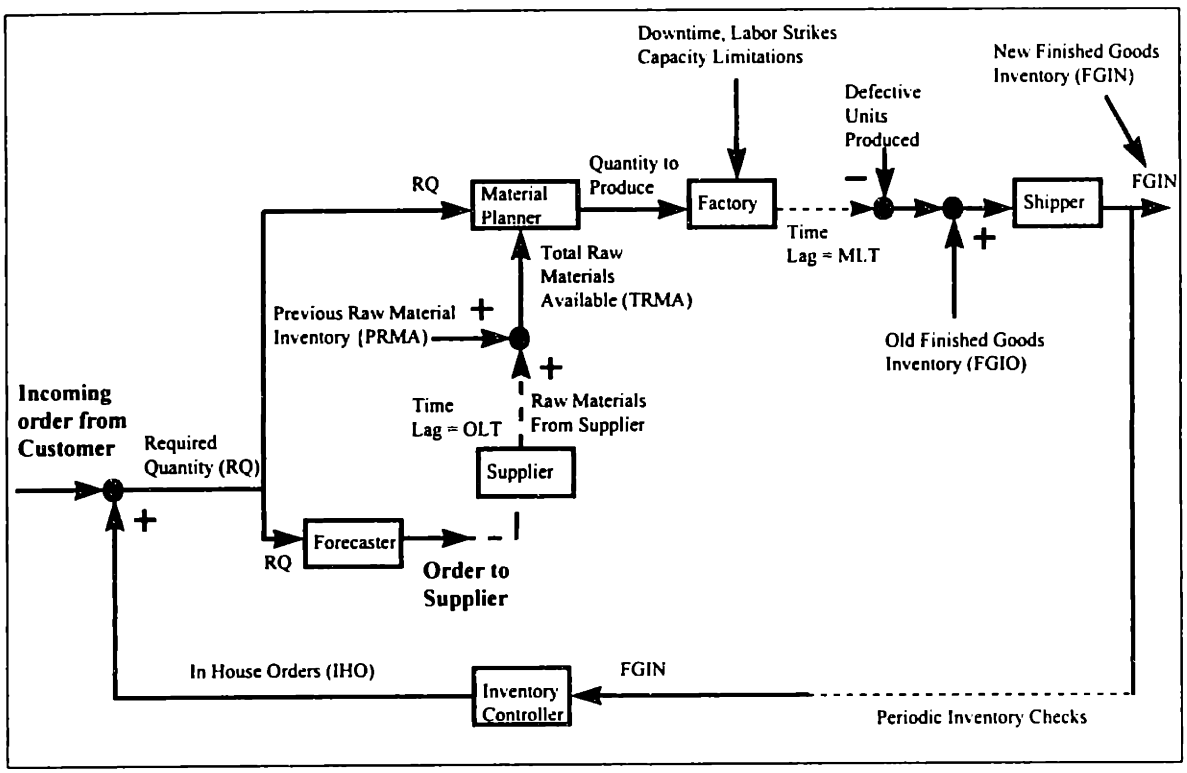


Figure 2.4

There are five control blocks in Figure 2.4: the forecaster, the material planner, the factory, the shipper and the inventory controller. On the upper information path, the forecaster receives information concerning the need for raw materials (the Required Quantity (RQ) that the factory needs to produce to maintain the inventory level of finished goods) and places an order for raw materials at time t. This order is received by the material planner at time t + OLT. On the other path the material planner receives two pieces of information, the Required Quantity of goods to be produced, and the Total Raw Materials Available (TRMA). The material planner takes the minimum of these two quantities and passes this information to the factory in the form of the Quantity of goods to Produce (QP). The factory attempts to fill this order, subject to disturbances in the

form of downtime, strikes, capacity limitations, etc., producing a quantity of units at time t + MLT (manufacturing lead time). The shipper receives the sum of the defect free goods (Good Units Produced or GUP) which were ordered to be produced at t - MLT, plus the previous (Old) Finished Goods Inventories (FGIO), and ships according to the Incoming Order (IO) received from the customer (if demand can be met; if FGIO + GUP > IO). The inventory controller receives periodic information in the form of the number of defect-free units left after shipping. This quantity is designated the new finished goods inventory level (FGIN).

If we perform a Laplace Transformation (Van de Vegt [1994]) on the governing equation for this simple feedback loop, we will be able to analyze the system response to various inputs. We can write the quantity total raw materials available (TRMA) in the S-domain as:

$$TRMA(s) = PRMA(s) + (RQ(s))FO(s)e^{-OLTs}$$

where FO(s) is the transform of the forecaster, TRMA(s) is the transform of the Total Raw Material Available, PRMA(s) is the transform of the Previous Inventory of Raw Materials, and RQ(s) is the transform of the Required Quantity. We will assume the forecaster orders based on the Required Quantity plus the fluctuation in customer demand (required quantity at time t, less the required quantity at time t-1) as shown below:

$$FO(s) = RQ(s) + \eta[RQ(s) - RQ(s)e^{-s}]$$

where η is the Forecaster factor, and reflects the influence of the rising or falling of customer demand on the Forecaster's order for raw materials. The level TRMA will only affect the outer loop if it is less than the required quantity.

The new level of finished goods inventory in the S-domain is given by:

$$FGIN(s) = \left[(RQ(s) \bullet MP(s) \bullet F(s)) - DU(s) + FGIO(s)\right]S(s)$$
$$RQ(s) = IO(s) + IC(s) \bullet FGIN(s)$$

where FGIN(s) is the transform of New Finished Goods Inventory, RQ(s) is the Required Quantity of units, as before, MP(s) is the transform of the Material Planner, F(s) is the transform of the Factory, DU(s) is the transform of Defective Units produced, FGIO(s) is, again, the transform of the Old Finished Goods Inventory level, S(s) is the transform of the Shipper, IO(s) is the Incoming Orders and IC(s) is the transform of the Inventory Controller. Since this system has three inputs, it has multiple transfer functions given by:

$$\frac{FGIN(s)}{IO(s)} = \frac{MP(s) \bullet F(s) \bullet S(s)}{1 - MP(s) \bullet F(s) \bullet IC(s) \bullet S(s)}$$
$$\frac{FGIN(s)}{DU(s)} = -\frac{1}{1 - MP(s) \bullet F(s) \bullet IC(s)}$$
$$\frac{FGIN(s)}{FGIO(s)} = \frac{1}{1 - MP(s) \bullet F(s) \bullet IC(s)}$$

The system is thus dependent on the four functions MP, F, S and IC as well as the three inputs IO, DU, and FGIO. If the Material Planner simply outputs the minimum of

the Total Raw Materials Available (TRMA) and the Required Quantity, we can model the MP as a simple amplifier with a gain of 1 if RQ < TRMA, and a gain of TRMA/RQ if RQ>TRMA. The Factory will also be a simple amplifier with a gain of 1 if there are no disturbances and less than 1 if there are, plus a time delay element, giving a transform of

$$F(s) = K_d e^{-MLTs}$$

where K_d is a function of the disturbances (machine downtime, labor strikes, capacity limitations, etc). For large batches, the changeover time will be small compared with the length of the run, and the capacity lost due to changeover time will be small. However, if the factory needed to changeover to produce several different parts in one day, the length of changeover time would reduce the capacity of the factory greatly.

The Shipper will ship the quantity IO unless GUP + FGIO < IO, in which case he will ship GUP + FGIO. FGIN, which is the information received at periodic intervals by the Inventory Controller, will be set to GUP + FGIO - IO or 0, whichever is larger. We will assume the inventory controller has accurate information about the inventory in the factory (FGIN) and outputs the difference between the Standard Inventory (SI), the present finished goods inventory (FGIN), as In House Orders (IHO):

$$IHO(s) = SI(s) - FGIN(s)$$

The plant will act as a simple closed loop system (with one time delay, the manufacturing lead time) attempting to maintain a constant inventory of finished goods. The feedback given by the inventory controller is crucial element of this loop, so we will explore alternate equations for IHO(s) and their effect on system performance.

2.4. Traditional Linear Distribution Systems (LDS)

In traditional LDS chains we see two main difficulties: (1) two time delays (i) between the time the material planner sends the quantity to produce to the factory and the time product is available to the shipper; (ii) between the time the forecaster places an order for raw materials and the moment they are received by the material planner. (2) Inaccurate information transferred throughout the plant. We will explore the effects of these difficulties separately. First we will assume that all information flows as enumerated in section 2.3. In section 2.10 we will add the effects of imperfect information flow.

We now return to the model of Figure 2.4, from which we can write a deterministic equation to measure the level of finished goods inventory and look at the performance of the system for several patterns of customer demand. We need only specify eight values for the system. These are: (1) the Standard finished goods Inventory level (SI), (2) the Manufacturing Lead Time (MLT), (3) the Order Lead Time (OLT), (4) the initial raw materials inventory, (5) the factory capacity, (6) the quality of the factory (% defects), (7) disturbances (magnitude and schedule), and (8) the gain of the Forecaster control block (which will amplify or dampen the effects of variation in customer demand on the order for raw materials to the supplier).

2.5. Step Increase in Demand

Consider a factory with a stable customer demand of 400 units per week, a standard finished goods inventory level (SI) of 1200, a manufacturing lead time of 4 weeks, an order lead time of 5 weeks, beginning raw materials inventory of 3000 units, and a capacity of 2000 units per week. For the present we will assume that there are no disturbances to the factory, the yield is 100% (no defective products are made), and the forecaster's gain is 0. With the inventory controller output given by SI-FGIN, as shown above, the system response to a 10% step increase in sales (440 units) is shown in Figure 2.5a.

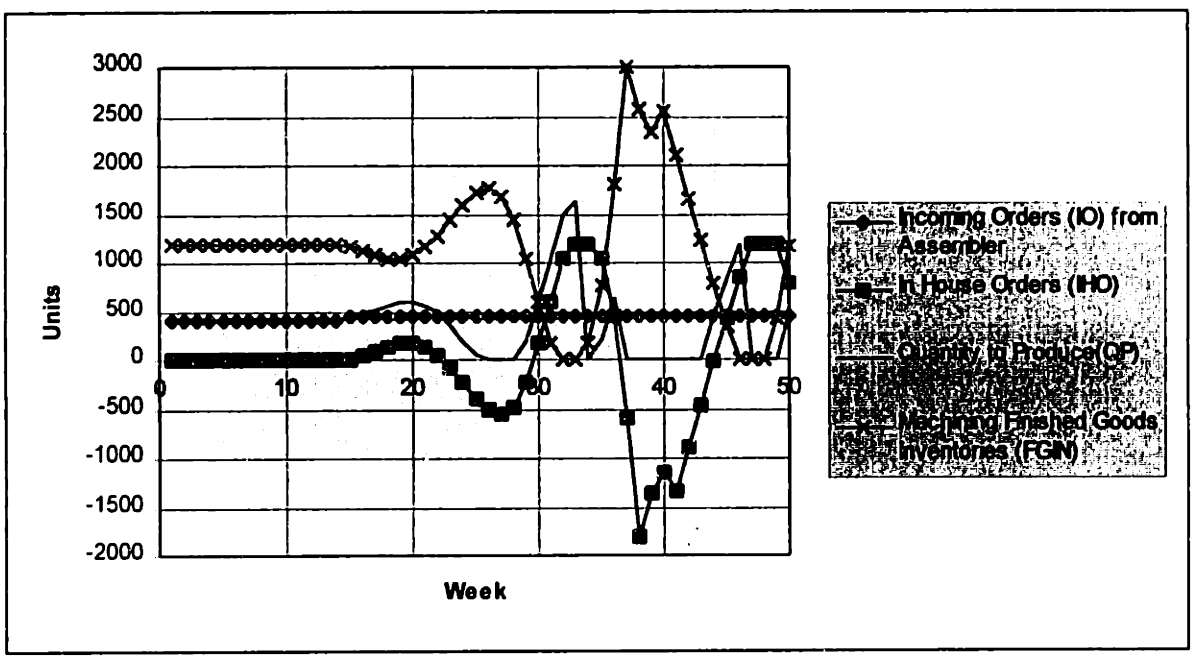


Figure 2.5a Response of Machining to a 10% Step Increase

In Figure 2.5a, the quantities are calculated from the following relations:

IO = 400 (t=0 to 14 weeks), IO = 440 (t ≥15 weeks) IHO = SI - FGIN (t-1)

QP = IO + IHO FGIN = FGIN (t-1) + GUP (t-MLT) - IO

In general a step input to a system will cause the system to oscillate at its natural frequency. For a simple lag system like the one developed here, T, the period of oscillation, is a function of the manufacturing lead time (MLT). The amplitude of the oscillations is a function of the standard inventory level (SI), and the MLT. The oscillation arises from the fact that in-house orders are based on the difference between the desired steady state inventory and the actual inventory of finished goods. When demand rises to 440 units, the material planner orders the factory to produce 440 units. However, this order is not received for MLT in the weeks (the length of manufacturing lead time). As a result, the IC will continue to generate in house orders for goods that are already in production. MLT weeks later the FGIN will begin to rise back the SI level and then surpass it, causing the in-house orders to become negative.

Figure 2.5a shows how an increase in incoming orders from 400 during the first ten weeks, to 440 in the fifteenth week causes the system to oscillate as IHO oscillates. The cyclical behavior is independent of SI. The effect of variations in MLT and SI are shown in Table 2.1.

Table 2.1

	MLT	T	Amplitude	Mean	Backlog pattern
RM = 3000	1	5	40	1200	-
Capacity = 2000	2	11	1280	1600	-
SI = 1200	3	17	2360	2240	120
	4	23	3280	2820	80,440,40
	5	29	4120	3780	320,440,240
RM = 3000	1	5	40	1200	-
Capacity = 2000	2	11	1240	680	-
SI = 2000	3	17	2420	2080	-
	4	23	3760	4120	-
	5	29	4840	4680	160,40

The quantities in Table 2.1 reflect the characteristics of the FGIN as shown in Figure 2.5b.

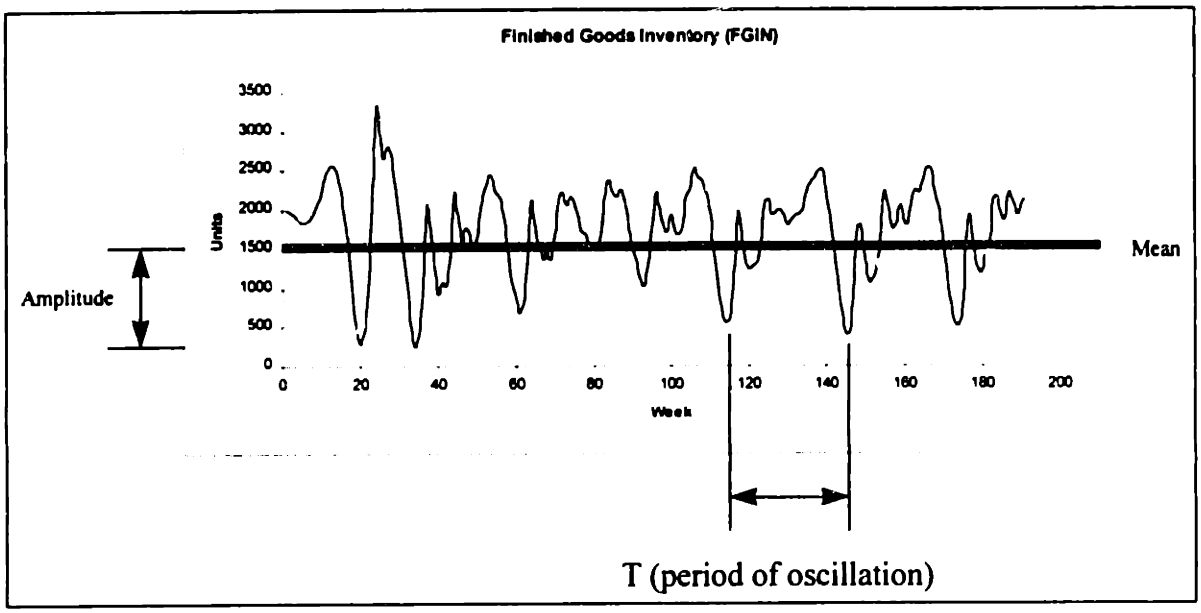


Figure 2.5b Oscillatory Behavior of FGIN for a 10% Step Increase

Increases in the MLT lead to increases in T the period of oscillation, as well as increases in the amplitude of oscillation and resulting backlogs. Increasing the standard inventory from 1200 to 2000 does eliminate some of the backlogs (for MLT = 3, 4) but lengthens the period of oscillation.

Since the inventory controller generates the cyclical nature of the in-house orders, why not simply eliminate the in-house orders and set the required quantity (RQ) equal to the incoming orders (IO)? In essence this will make the system open loop, since the in house orders represent the feedback of information from the output of the system. The obvious answer is that in a perfect world, with no defects, no disturbances, no capacity limits, and no raw material shortages, the inventory controller is unnecessary. However, in a real world, defects and disturbances which prevent the factory from producing the required quantity for the incoming order will reduce the finished goods inventory level to

zero leading to a string of backlogs. We must find a way to produce the right quantity of defect free parts at the right time, even in the face of defects and disturbances.

2.6. Redefining Finished Goods Inventory and Safety Stock

What we can do is control the quantity of inventory. It is clear from Figure 2.5 that simply setting the in-house orders (IHO) equal to the standard inventory (SI) minus the current inventory failed to maintain the level of FGIN at the SI level. Instead, we will establish a new expression for Finished Goods Inventory, FGI, including FGIN and all those goods that we have ordered the factory to produce in the past MLT weeks. Simultaneously, we will rewrite the safety stock or standard inventory to include a quantity equal to the current incoming order multiplied by the MLT. This represents the desired safety stock in the plant plus the safety stock in the pipeline (which is based on current order size). Thus, the new formula for IHO is given by:

$$IHO = TSI - FGI$$

where,

$$FGI = FGIN(t-1) + \sum_{i=1}^{MLT} QP(t-i)$$

$$TSI = SI + IO \bullet MLT$$

The output of the factory to the a 10% step increase in sales is given in Figure 2.6.

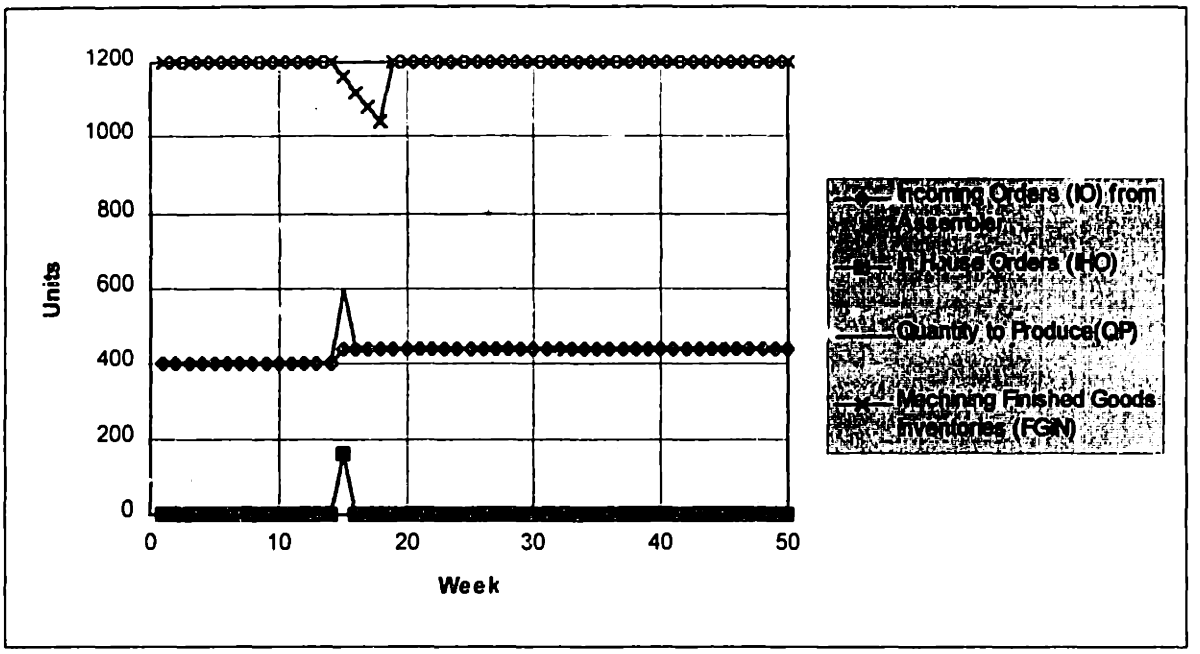


Figure 2.6 Response of Machining to a 10% Step Increase

Clearly, the response is much better. If we then input quality problems, in the form of a 90% yield rate starting in week 15, we see that the system responds as in Figure 2.7.

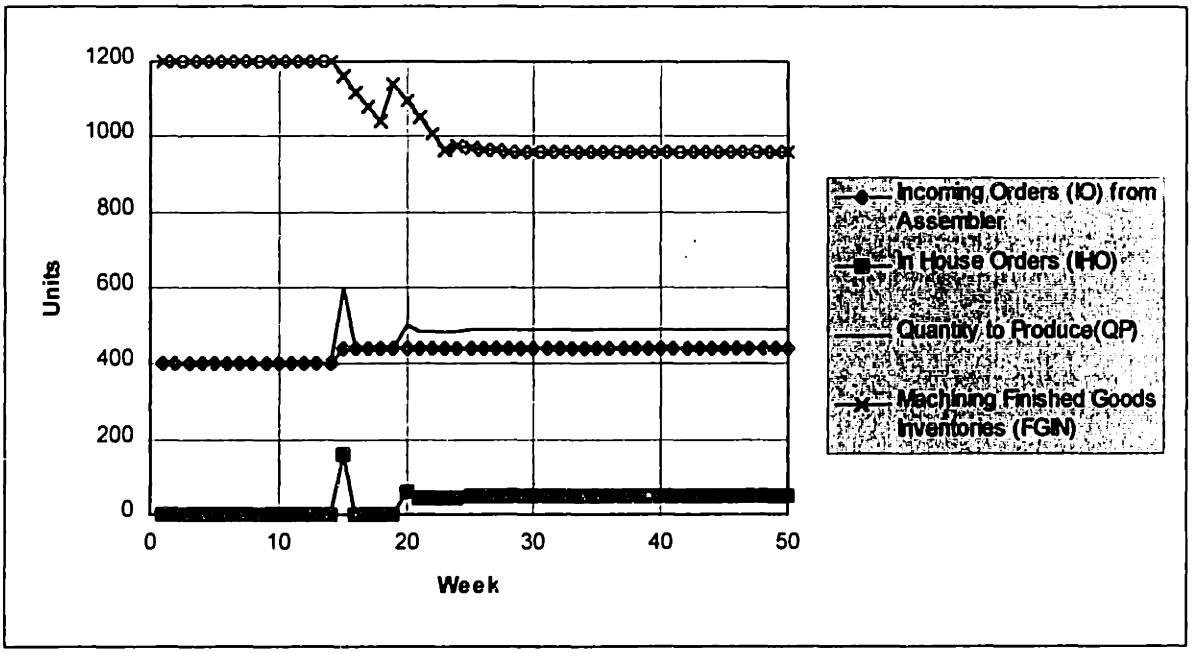


Figure 2.7 Response of Machining with 90% Yield to a 10% Step Increase

The system response is good, but not perfect. The finished goods inventory level never recovers to 1200, but reaches steady state at 955 units. This is a result of the expression for FGI, which is based on QP rather than on the good units that are produced. The system will recover from quality problems, but only if the factory begins producing 100% good units.

One way to address this problem would be to base the expression given for FGI on the good units produced as below :

$$FGI = FGIN(t-1) + \sum_{i=1}^{MLT} GUP(t-i)$$

The system responds as shown in Figure 2.8.

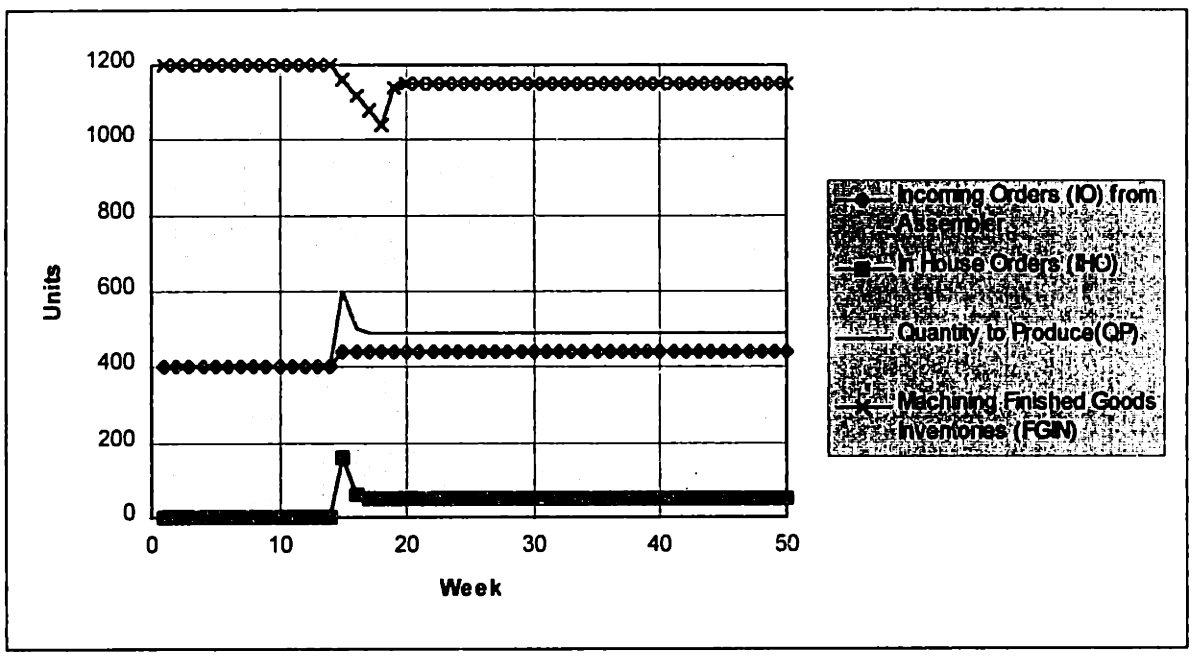


Figure 2.8 Response of Machining with 90% Yield to a 10% Step Increase

There is considerable improvement. In this case the system reaches steady state at 1160 units. However, since the quality of the goods that will be produced at time t is not known at time t (since it is later in the control loop), we can only forecast the number of

good units that will be produced, and thus we are unable to restore the system to the full level of SI. If we try to forecast the number of defective units we will run the risk of overproducing if the quality level is better than we predict.

2.7. Impulse Rise in Demand

We will now input a 20% increase in demand in week 15, and then return to the original demand quantity (400 units) from week 16 on. The result is shown in Figure 2.9.

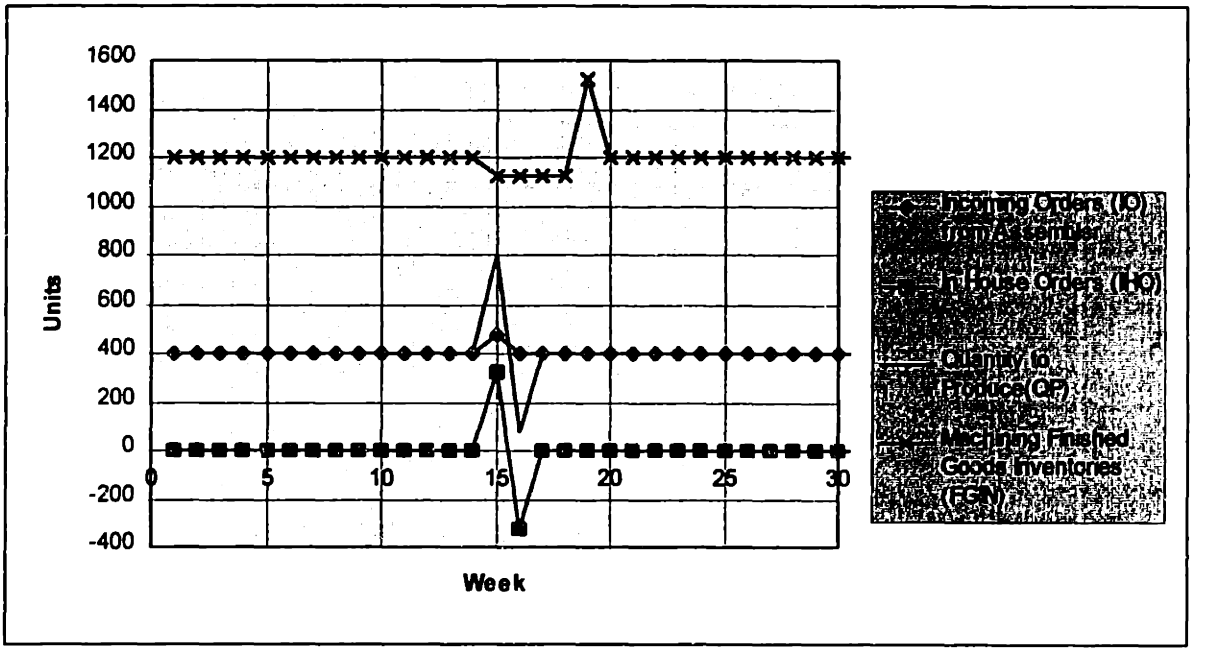


Figure 2.9 Response of Machining to a 20% Impulse Increase

We can see that the factory over produces in the 15th week. This is a result of the effort to adjust the production of the factory to the new level of production, which is 480 units. In essence, the QP in week 15 is attempting to fill the pipeline with enough product for a demand of 480 units. Since demand falls back down to 400 units, there is a surplus of production. This result is similar to that presented in Forester [1965].

If we try to dampen the peak of overproduction by slightly adjusting the expression for TSI to include an average of the customer demand over the length of the cycle time we achieve the following result:

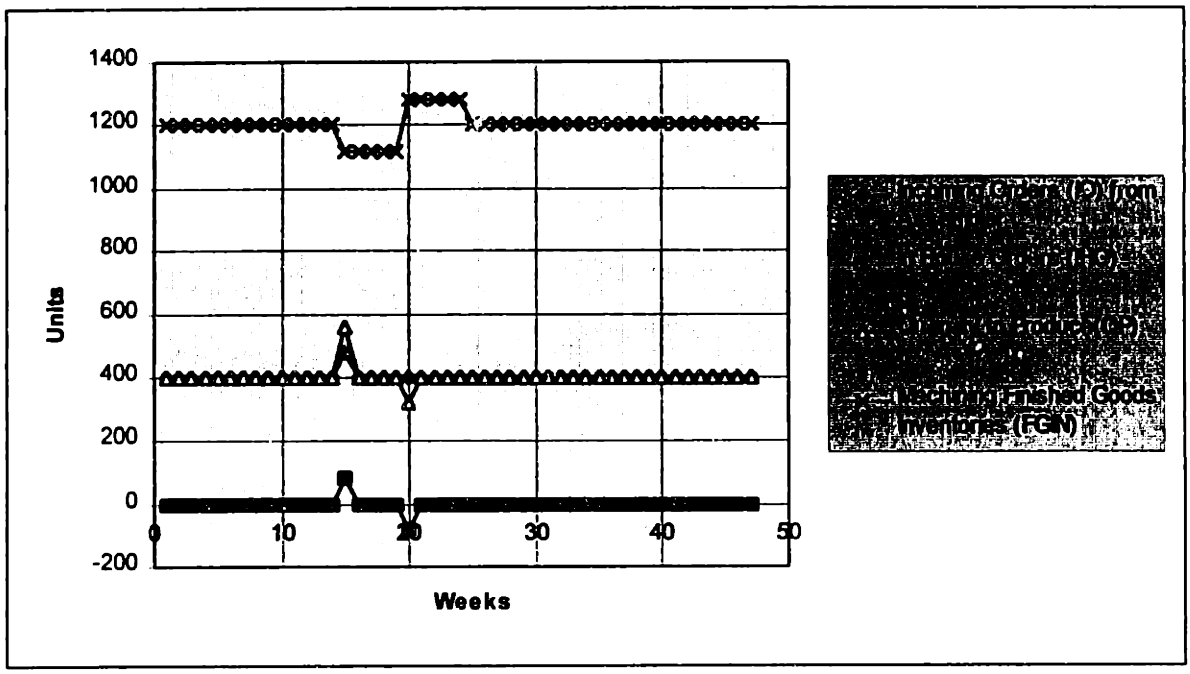


Figure 2.10 Response of Machining to a 20% Impulse Increase (TSI averaged)

This system has a reduced spike in FGI because it is a more conservative ordering approach but it returns to the standard inventory level of 1200 units much slower than in Figure 2.9. This system is slower to respond to fluctuations in the market. Since meeting demand is generally viewed as more important than inventory carrying costs, we will not average the value for TSI over MLT weeks.

2.8. Sinusoidally Varying Demand

The response of the system to an oscillatory demand with an amplitude of 10% (40 units) and a period of 52 weeks is shown in Figure 2.11

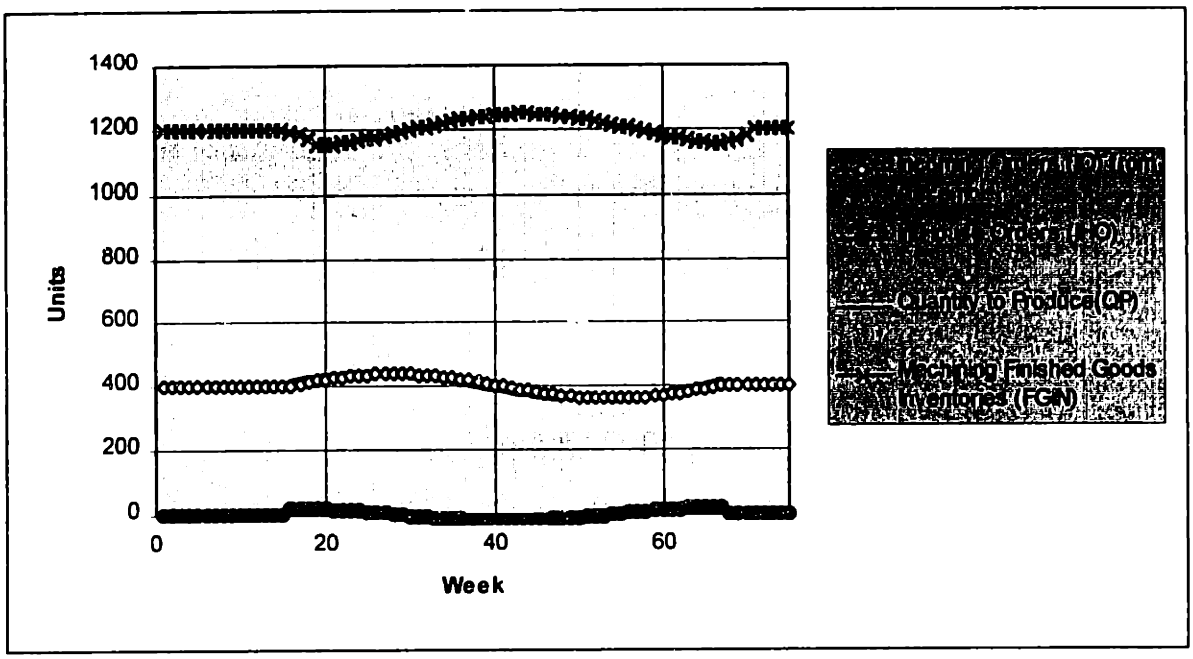


Figure 2.11 Response of Machining to a Sinusoidal Customer Demand with Amplitude of Oscillation 10% of Customer Demand

The response is very good. There is very little overproduction, and the system returns to steady state rather quickly (in a period equal to MLT weeks). In essence, the equation we have derived for the inventory controller simply requires that the orders placed be equal to the customer demand plus the difference between the safety stock and the current inventory, where the current inventory includes FGIN plus all product currently being produced in the factory.

2.9. Capacity Limitations

Having reached a relatively satisfactory system response, we will see how the system responds to capacity limitations. If the capacity of the factory is only 450 units, so that the demand changes from 400 units, or 89% of capacity to 440 units (a 10% step increase) or 98% of capacity, we will see how capacity limitations affect system performance. The response is shown in Figure 2.12.

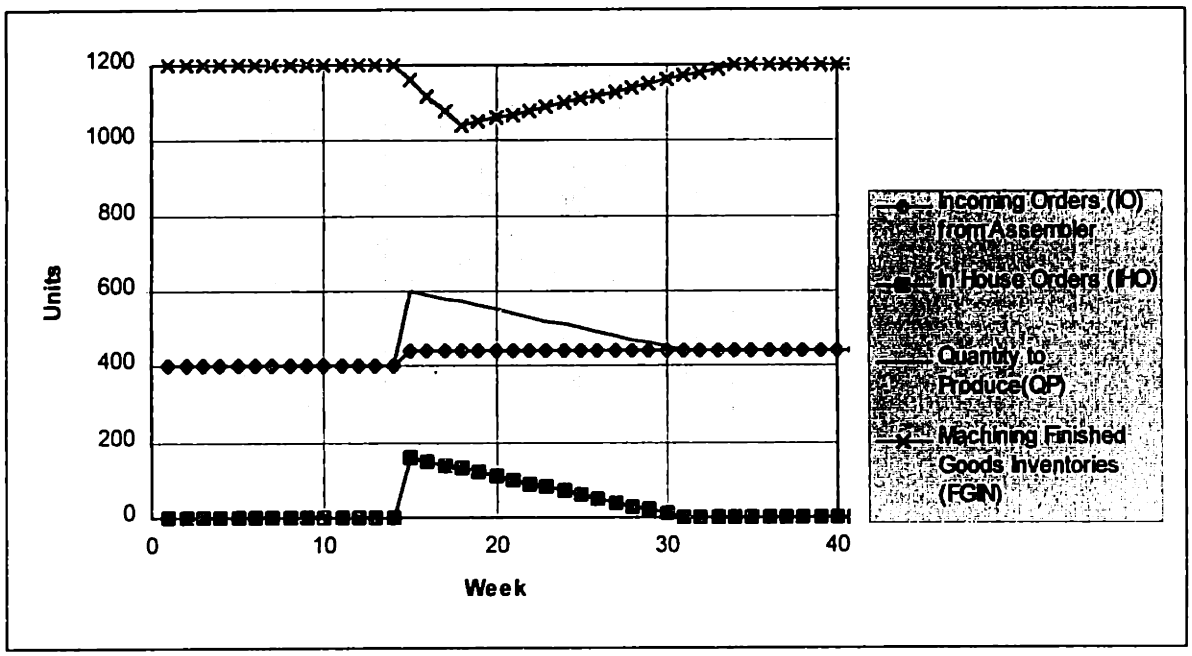


Figure 2.12 Response of Capacity Limited Machining to a 10% Step Increase

Clearly the length of time for the system to recover back to steady state has increased significantly as a result of the capacity decrease. Disturbances serve to limit capacity and so have the same effect.

2.10. Overall LDS Performance

If we now extend our model to include other members of the LDS, we can see how effects are transmitted along the chain. We will add three members to the chain,

essentially modeling the system in Figure 1. The properties of all four members will be the same as the machining plant: SI = 1200, MLT = 4 weeks, OLT = 5 weeks (one week after an order for raw materials is placed, the supplier begins production), RMI(t=0) = 3000, and capacity = 2000. The response of the system to a 10% step increase in customer demand to the Assembler is given in Figures 2.13a, and 2.13b.

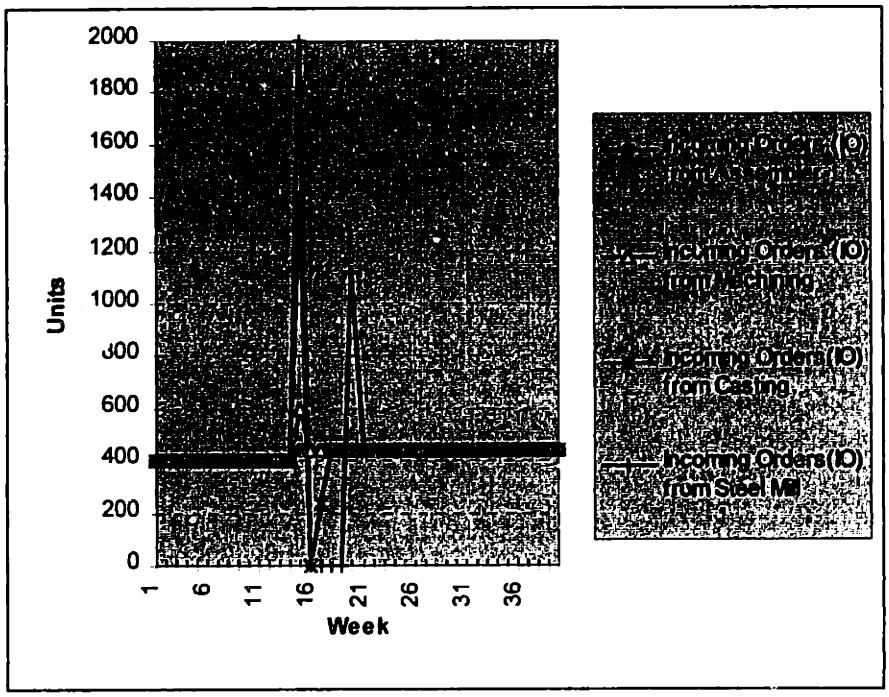


Figure 2.13a System Propagation of Orders in Response to a 10% Step Increase

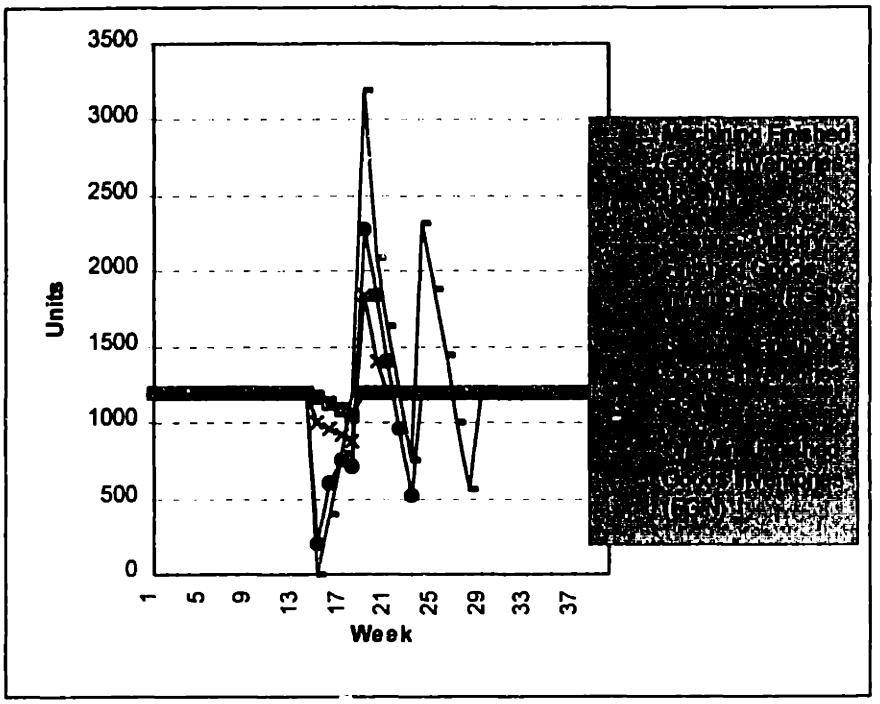


Figure 2.13b System Inventory Response to a 10% Step Increase

One can see the amplification of orders in Figure 2.13a, and the resulting cycling in inventories in Figure 2.13b. The one week time lag for order processing also is apparent in the phase shift of production. This causes the Ore Mine to produce at its peak in week 18 rather than in week 15 when the customer increase occurred.

The response of the system to a 20% impulse rise in demand is similar, as shown in Figure 2.14a and Figure 2.14b.

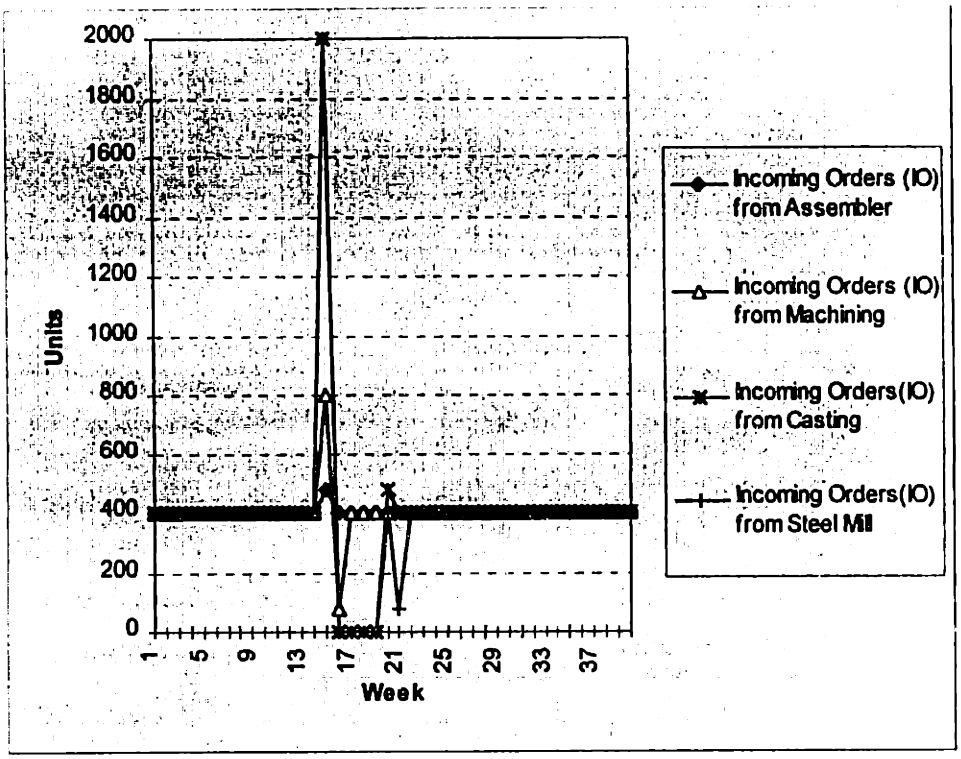


Figure 2.14a System Propagation of Orders in Response to a 20% Impulse Increase

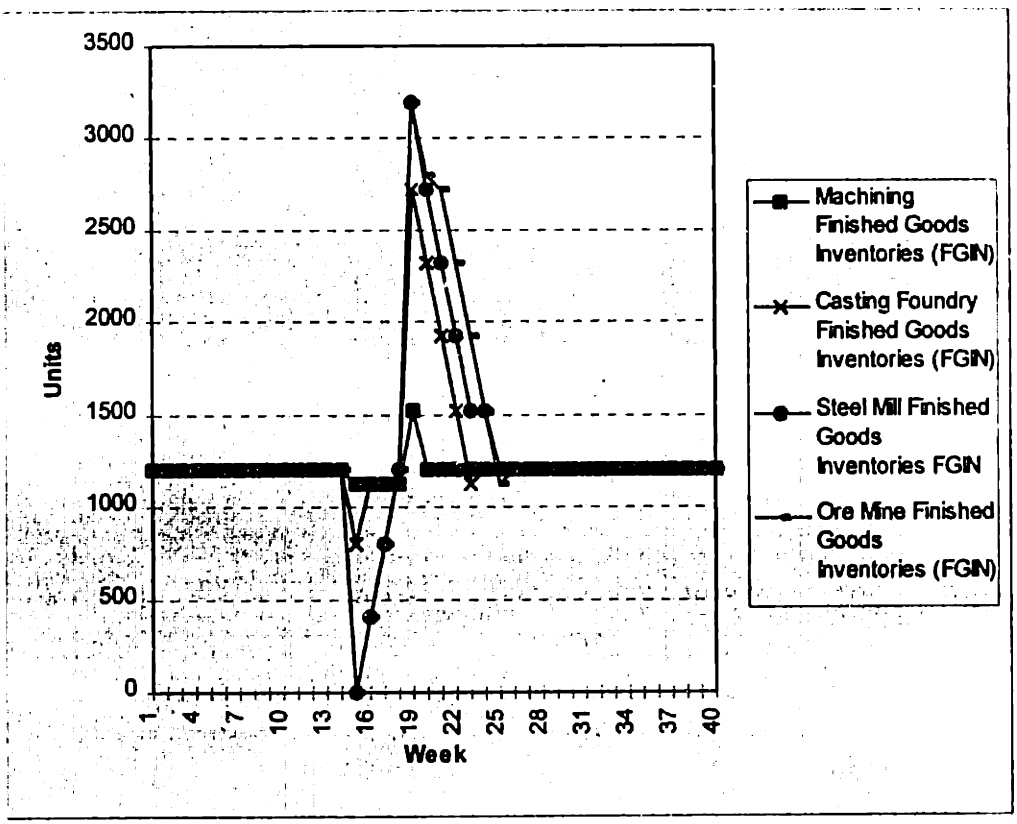


Figure 2.14b System Inventory Response to a 20% Impulse Increase

The key difference is that the 20% increase in customer demand causes an initial spike in both the Steel Mill and the Ore Mine which is larger than the capacity (2000 units), resulting in backlogs for a single week. In addition these two links are slower to return to their standard inventory levels.

Finally, there is the overall system response to a sinusoidal increase in demand (amplitude of 40 units, period of 52 weeks), shown in Figure 2.15a, and 2.15b.

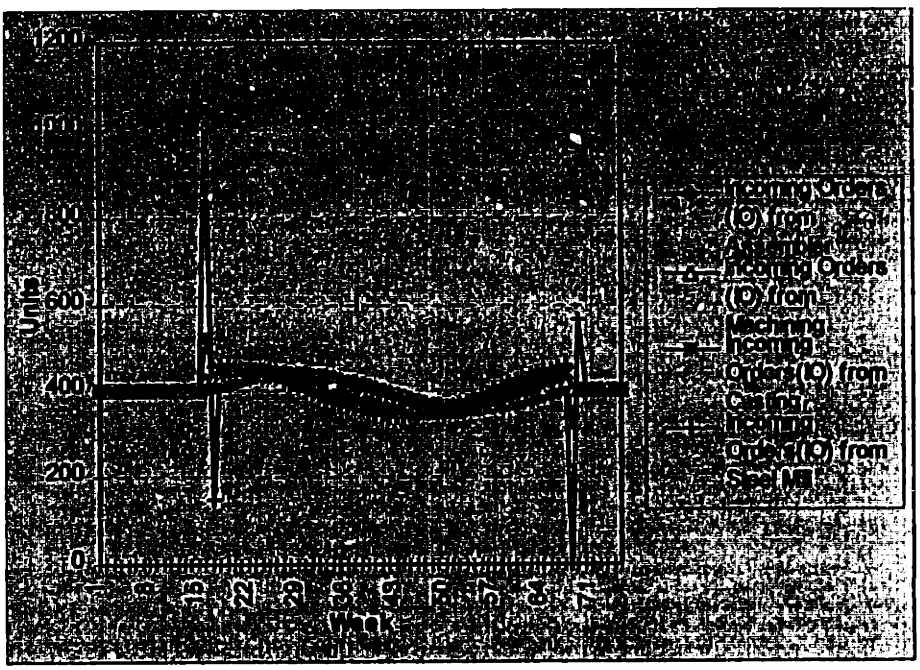


Figure 2.15a System Propagation of Orders in Response to a Sinusoidal Customer Demand Pattern (Amplitude 10%)

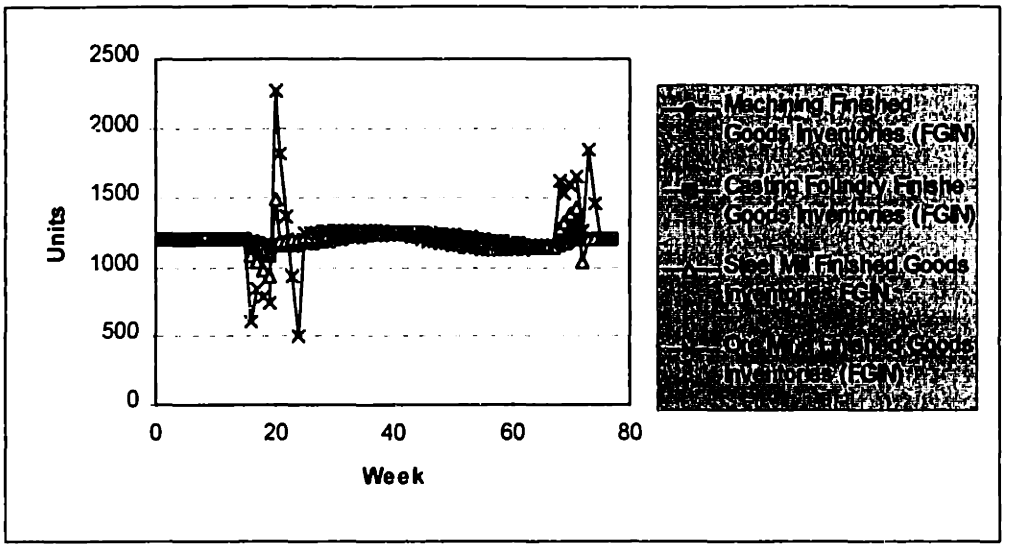


Figure 2.15b System Inventory Response to a Sinusoidal Customer Demand Pattern (Amplitude 10%)

The behavior of the system to a sinusoidal input is quite extraordinary. The initial increment (from week 15 to 16) in the sinusoidal increase is roughly two units. However, the Ore Mine sees an increase of nearly 300 units four weeks later. This is a result of the system trying to fill the factories with product at the new customer demand level.

Increases in the MLT serve to amplify variation in customer demand, as well as lengthen the time it takes each plant to recover to a steady state inventory level. As LDS chains get longer and longer (some in the automobile industry can reach 20 links), the amplification of the initial change in customer demand is compounded, resulting in increasing backlogs as the distance from the customer to the plant increases. The expression for the Incoming Order (IO) for in the week of an increase is given by the following expression:

$$IO = IO'_{x-1} + (IO'_{x-1} - IO'^{-1}_{x-1})\,MLT$$

Where the subscript X refers to position in the chain, and the superscript t refers to the week. So, the IO as a function of the length of the chain can be found by iterating the above expression for X = 1 to n where n is the length of the LDS chain.

Many factories attempt to avoid the spikes caused by variation in customer demand by "forecasting" future demand. If the forecaster sees an increase in downstream demand he will order raw materials for the incoming order plus an additional quantity, knowing that it will take five weeks (the OLT) to receive the order from the supplier, and wanting to have plenty of raw materials should demand continue to rise. Similarly, the forecaster may also order less raw materials from the supplier than the incoming order, following a decrease in customer demand. The effect of this "forecasting" on system response can be seen by setting the forecaster factor equal to 0.1. This is equivalent to a forecaster who orders raw materials equal to the order plus 10% of the difference between this weeks order and last weeks order. The resulting response of the system to a 20% impulse input (compare with Figure 2.14a and Figure 2.14b) is shown in Figures 2.16a and 2.16b.

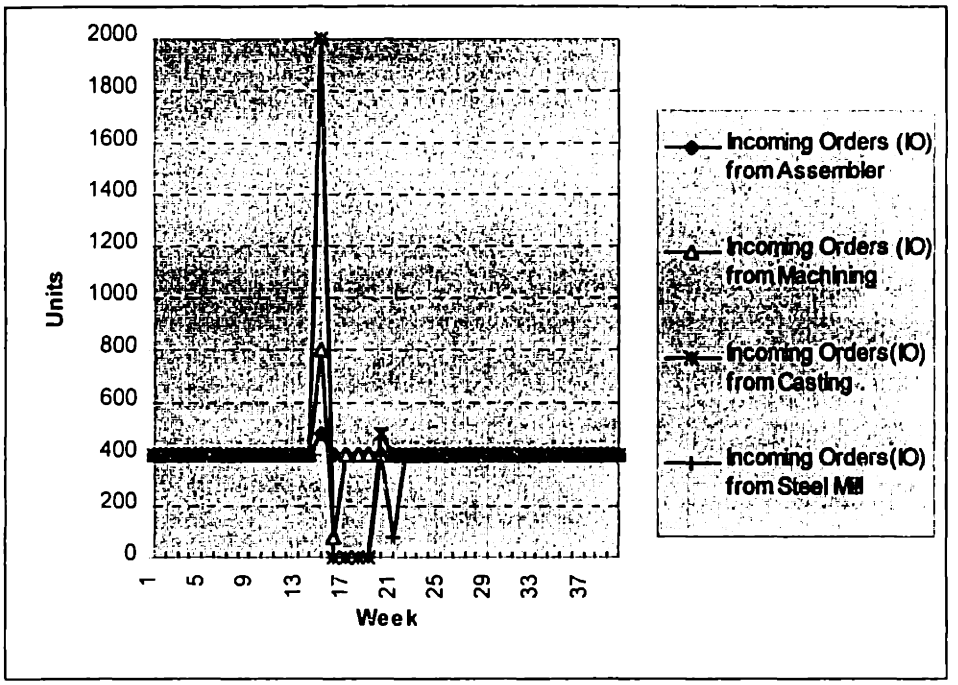


Figure 2.16a System Propagation of Orders in Response to a 20% Impulse Increase with Forecaster factor = 0.1

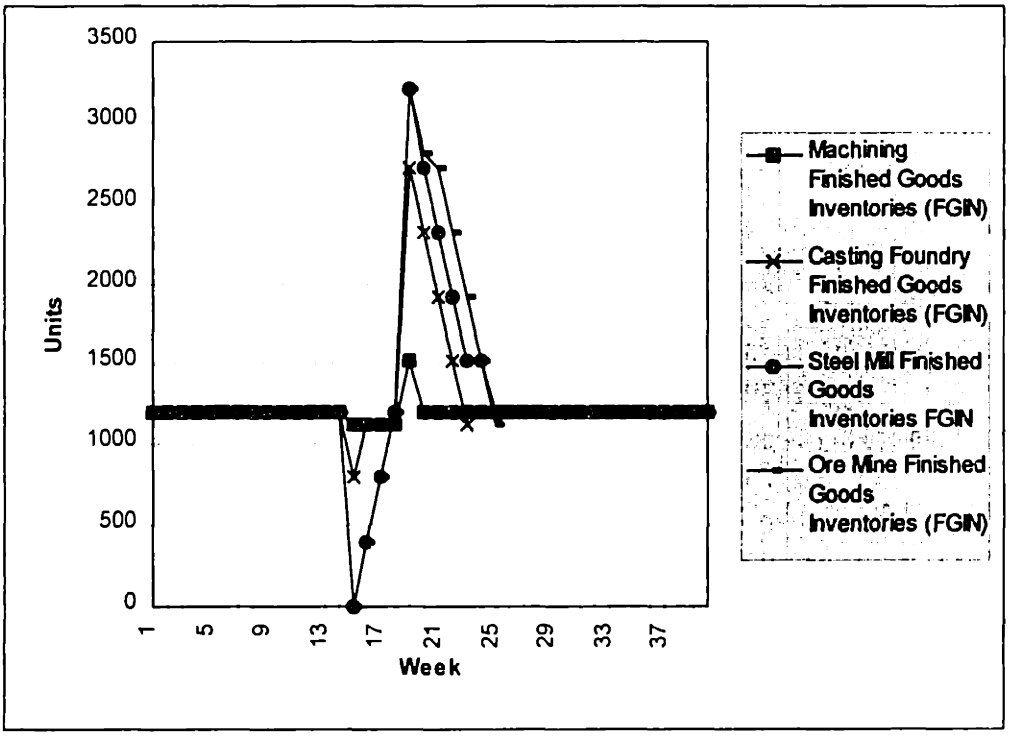


Figure 2.16b System Inventory Response to a 20% Impulse Increase with Forecaster factor = 0.1

The forecasting factor of 0.1 results in an amplification of each order that is passed up the LDS. The end result is that the Ore Mine will see demand raised artificially by a factor of $1.1^3 = 1.33$, resulting in extremely high inventories and a long recovery period.

2.11. Lean Manufacturing Principles Applied

While each plant's response has been greatly improved compared with the original machining plant in Figure 2.5, we can see that the system's tendency to amplify a rise in demand and become cyclic has not been eliminated. To avoid these pitfalls, we must rethink the way production is controlled. This will require us to alter our model of the manufacturing system.

Figure 2.4 must be modified in the following ways: (1) Products must be "pulled" by the shipper which will serve as the signal to begin production rather than "pushing" an order for a given quantity of goods through the factory. The shipper will "pull" or remove product to fill a shipment at a rate which will fill the day's demand from the customer. The length of time between production of successive parts to meet a demand is called the Takt time. (2) We must establish a preset or predefined inventory quantity which will replace the inventory controller by creating a flow of information which flows in the opposite direction of the material flow and sets the total quantity of parts that can be in inventory at any given time. In essence we are specifying a fixed and controlled quantity of inventory, rather than allowing it to fluctuate as was the case in the traditional manufacturing system. Having a fixed quantity of inventory in the factory will eliminate the amplification due to variability in customer demand. (3) We must reduce the MLT of

the factory, and the OLT of the supplier loop. (4) The quality of the factory must be greatly improved and approach 100% yield. (5) There must be a decrease in changeover time which will result in an increase in capacity. (6) The forecaster must be eliminated completely. (7) Lastly, there must be a reduction of disturbance effects.

Most disturbances caused by unplanned machine downtime will be eliminated through implementation of Total Productive Maintenance (see Nakajima [1988]). We will create an accurate flow of information by eliminating the forecaster. Since forecasting is essentially guesswork, the optimal value for the forecasting factor is 0. In this case, the required quantity (RQ) will be passed directly to the supplier. The decrease in changeover time will be accomplished by using the methods of chapter 4.

The manufacturing lead time and order lead time will be reduced by converting the factory from a job shop to cellular manufacturing. This process is detailed in chapter 5. This transition to cellular manufacturing will be accompanied by implementation of improved quality methods, including the transition from final inspection to successive inspections to self checks and finally to installation of Poka-Yoke devices. For details on this transition, see chapter 6.

The new control system, which has been termed "pull" to contrast it with "push" systems where product is produced to fill stock, will result in production control on an individual part basis. Under single piece flow parts will be produced at a rate set by the Takt time. This will create a synchronized flow of product and essentially form a linkage between cells. The production system will be transformed into a set of linked cells. The takt time is the length of time required between successive units of end product, and is determined by the following equation:

$$\text{Takt Time} = \frac{\text{Total time available (= 27600 sec/day)}}{\text{Total number of parts to produce / day}}$$

In short, the factory must produce a part once every X seconds, where X = Takt time. The new control system for the factory is shown in Figures 2.17a and 2.17b.

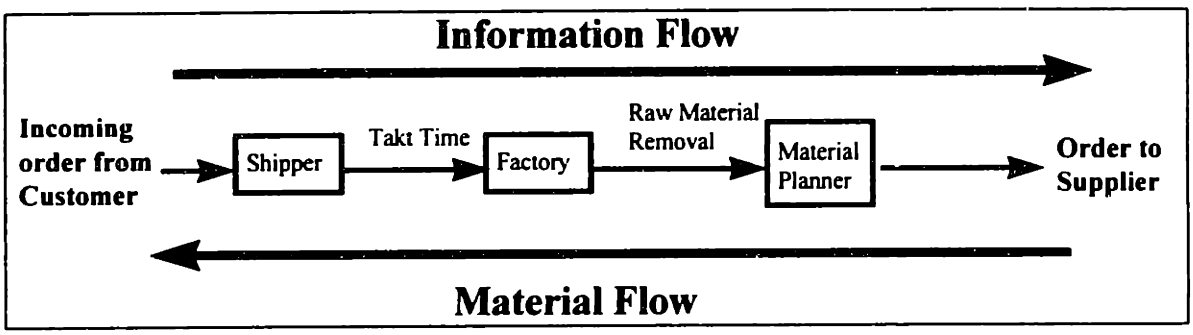


Figure 2.17a

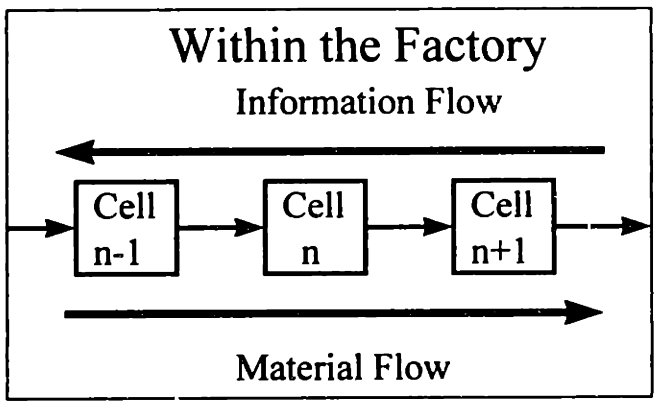


Figure 2.17b

Incoming orders are now received by the shipper who begins removing finished goods from the shipping area at a rate equal to the Takt time. The Takt time is the rate at which the shipper must pull a single part in order to fill the order for the day. The factory in turn pulls raw material from the raw material bins and produces parts at the rate that the shipper is pulling finished goods. Finally, the material planner multiplies the Takt

time by the reorder quantity and determines the frequency that shipments must be received from the supplier.

This system is considerably simpler. Information transfer in the form of the shipper removing product now forms the feedback for the control loop. The factory is not instructed to produce a certain quantity of goods, but rather to produce at a certain rate, as they are pulled to be shipped. By reversing the flow of information, we have transformed a system which "pushes" orders through the factory into a system in which a "pull" of product by the shipper (at a frequency set by the takt time) signals the system to begin production.

2.12. A Kanban Controlled Pull System Producing to Takt Time

In the factory, there is a fixed or controlled level of finished goods from which the shipper removes units to fill the incoming order. One way to fix the total quantity of inventory is through the use of kanban containers. One need only choose the safety stock level and keep a number of containers on hand such that the safety stock is simply the capacity of each container multiplied by the number of containers. If the foreman of the factory requires that all finished goods be placed in the containers, then the inventory can never rise above the safety stock level. Furthermore, as soon as a container is emptied by the shipper the container signals the factory to produce goods until the container has been refilled. The empty container can be a visual signal for the factory to produce, or it may contain a production ordering kanban card which instructs the factory to produce a certain quantity of a certain type of goods. The inventory level is regulated by the number of

containers in the plant, so if one wants to remove inventory one simply removes a empty container from circulation.

The behavior of the LDS under the new production control philosophy will be quite different from the previous case. First, the incoming order is received by the shipper, rather than the material planner. This is the first step in shifting the system from push to pull. Second, the time scale of the control system will be given in seconds rather than in weeks, and the number of units of product will be segmented into the holding capacity of a kanban container. The system response is shown in Figure 2.18. The shipper pulls product in kanban containers throughout the day paced by the Takt time.

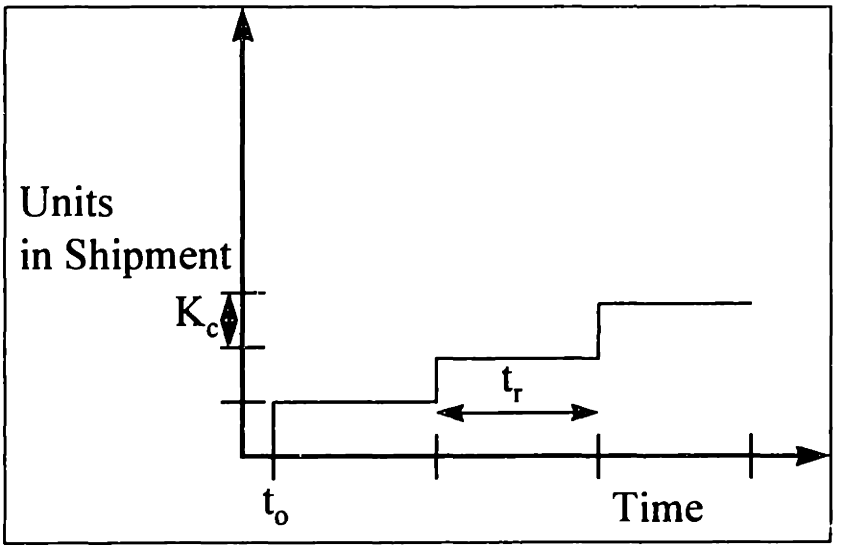


Figure 2.18

where t_r is the time between removal of containers, which is equal to K_c (the kanban container capacity) multiplied by the Takt time. The first container is pulled at time t_o.

The level of safety stock is shown in Figure 2.19.

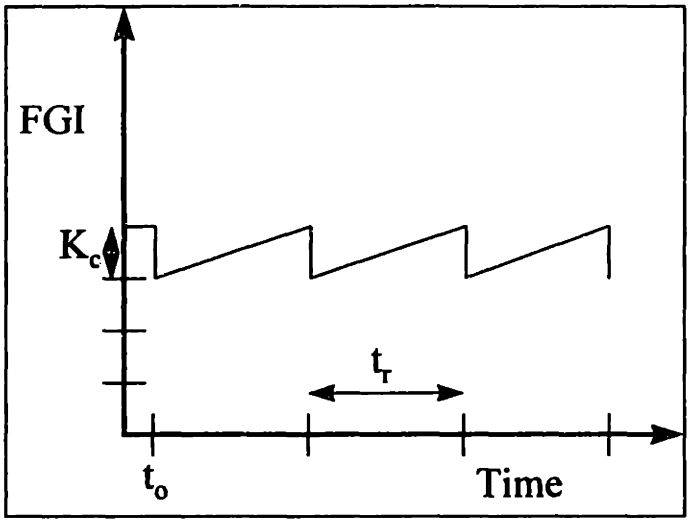


Figure 2.19

This factory has a safety stock of four containers multiplied by the quantity in each container. Similarly, the quantity of raw materials follows the trend in Figure 2.20.

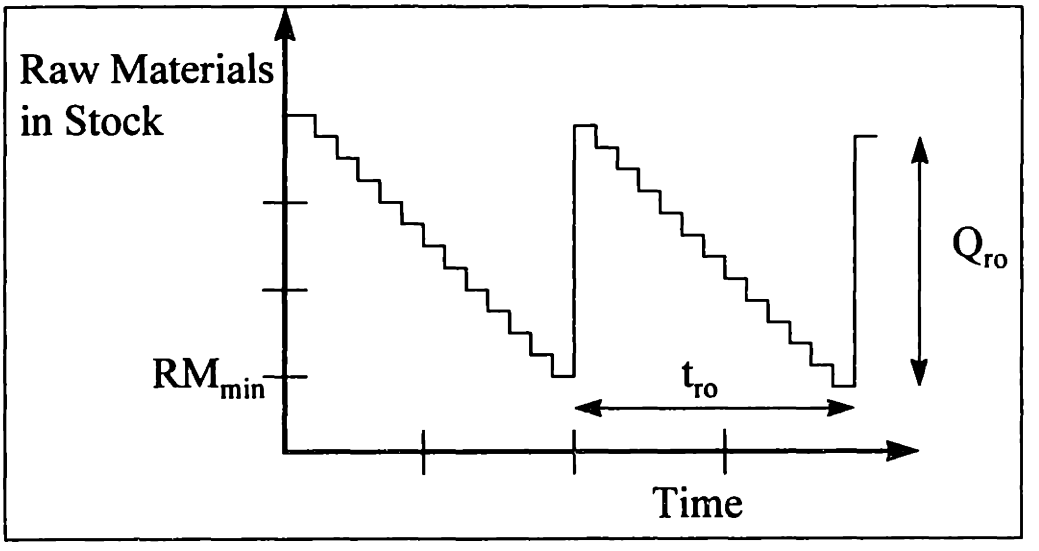


Figure 2.20

where RM_{min} is the minimum level of raw materials, t_{ro} is the reorder time interval, and Q_{ro} is the reorder quantity. Each small step represents a person from the factory pulling raw materials at the rate prescribed by the Takt time.

The fixed quantity of inventory (SI) can be modeled, and depends on the time interval between replenishment of the stores. Material between linked cells is removed and then replenished in a short period of time (the takt time multiplied by the quantity between the cells), whereas the material that links two plants (which is finished goods for one factory and raw materials for the other) is removed and replenished at a rate set by the shipping schedule and is generally much longer than the takt time.

Since the system is made up of four plants all run on a pull system, all four plants behave similarly. While the parameters of each system will be different, they can all be described by the non-dimensional graphs shown in Figures 2.18-2.20. The response of this system to various inputs is rather simple. For a 10% step increase, the Takt time will be divided by 1.1 in each plant which would be accomplished by adding manpower to the cellular manufacturing systems. For a 20% impulse increase the Takt time will be divided by 1.2 for the length of the impulse, and then return to the previous Takt time. This temporary shift in throughput can be accomplished by simply adding manpower to a system of linked cells. There will be no overproduction in a lean manufacturing system, because overproduction is a result of "pushing" an order through the factory rather than pulling the desired quantity from the finished goods stores.

Chapter 3. The Evils of Inventory

3.1. Introduction

The traditional manufacturing manager views inventory as a "necessary evil". Conventional wisdom says that it is necessary to carry a quantity of finished goods to meet customer demand, because customers are not willing to pay for an order and then wait for it to be produced. Chapter 2 showed the problems that are created in a system when a factory produces to fill stock levels. Here we will show how creation of inventory within each factory leads to ever increasing costs and system complexity. We will enumerate why inventory is indeed **evil**. In chapter 5 we will show how to transform a factory into a cellular manufacturing system (chapter 5) and eliminate large inventories and the problems that are enumerated below.

As the levels of inventory increase in a factory, a sequence of events take place which end up costing companies non-incidental quantities of capital, floor space, manpower and dollars. The first event to take place occurs when a part finishes a given process but the next station is not ready to accept it. The part must be stored until the next station becomes available. If there is floor space or table space available the part will probably just rest there. However, if more parts are completed before station 2 is ready, a rack will be needed to store the parts, or the parts will need to be sent to another location to wait for station 2.

Since this will most likely occur between many stations (unless the factory is standardized and balanced), it will not be long before the plant is filled with inventory racks. If these interfere with the working of the machines, the plant will buy or build a

warehouse to store the inventory. Keeping inventory in a warehouse often requires equipment to transport the parts to and from the warehouse (which may or may not be close enough to fill these needs with a forklift). Regardless of the distance to the warehouse, the material planner will need to know what is in the warehouse, so the factory will need some sort of inventory tracking system. Many factories have fully computerized inventory warehouses that tell the material planner exactly what parts are in the warehouse. However, these inventory systems are always fallible, which has made "lost in the warehouse" a common phrase among material planners.

3.2. The difficulties of maintaining a fixed quantity of inventory

It is common in traditional manufacturing environments to carry a specified quantity of inventory at each stage of processing, raw materials, work in process (WIP), and finished goods. However, the quantity of "safety stock" as it is often called, rarely remains constant. What actually occurs is that the material planner will have a level of inventory that he wants to maintain. However, due to the difficulties in keeping track of inventory sitting in a large warehouse near the location of the plant, the material planner often doesn't actually know the level of inventory of each type of product. Since backlogs are extremely costly and may cause the company to lose customers, material planners often carry more inventory rather than less, and are not overly concerned about carrying a little extra "safety stock". A common management response to backlogs is to instruct the material planner to carry more inventory, rather than to discover a better inventory control system.

So, if we now look at the total cost of carrying "just a little extra" inventory, rather than working to balance the manufacturing process (see section 5.11), we will see that the cost is enormous. The factory must purchase or build inventory racks, inventory bins, means to transport inventory between process steps, and to and from the warehouse (usually an armada of forklifts), and the cost of the warehouse. As well, one must include the manpower in the form of material planners and material handlers (forklift drivers), consumed floor space, and perhaps most importantly (but often overlooked) the capital required for all these investments. For companies who are capacity limited, reduction of inventory and inventory systems can add working capital, manpower, and floor space, as well as improve product quality and factory communication.

The costs above can be clearly attributed to the buildup of inventory that arises as a result of large batch sizes, unbalanced factories, and unlinked processing areas. However, a second set of evils is also the direct result of the buildup of inventory, but the connection is not as obvious, especially to upper management. The second set of problems arise as a result of the delay from the completion of one operation until the part reaches the next stage. Often the entire lot will be processed before being moved to the next station. If, for example, the final machine before inspection begins to make defective parts, the entire lot of parts will be defective and need to be scrapped. Compare this to scrapping a single part, if each of the parts were transferred to the inspector after completion of processing.

If engineering changes are made to a product, it is not uncommon to scrap all parts in inventory. All the parts in process, as well as all the parts stored in the warehouse

become worthless scrap. The cost of this is generally attributed to the engineering change, rather than to the true cause, which is the unnecessarily high level of inventory.

In summation, the buildup of inventory is a gradual process, resulting from the imbalances in the factory. An imbalance occurs when one machine or process produces product faster than another. Then, when a part finishes (on the faster machine or process) before the next machine or process is finished, a need for a resting place for the part arises. In Figure 3.1 there are two machining stations. Machine A produces one part

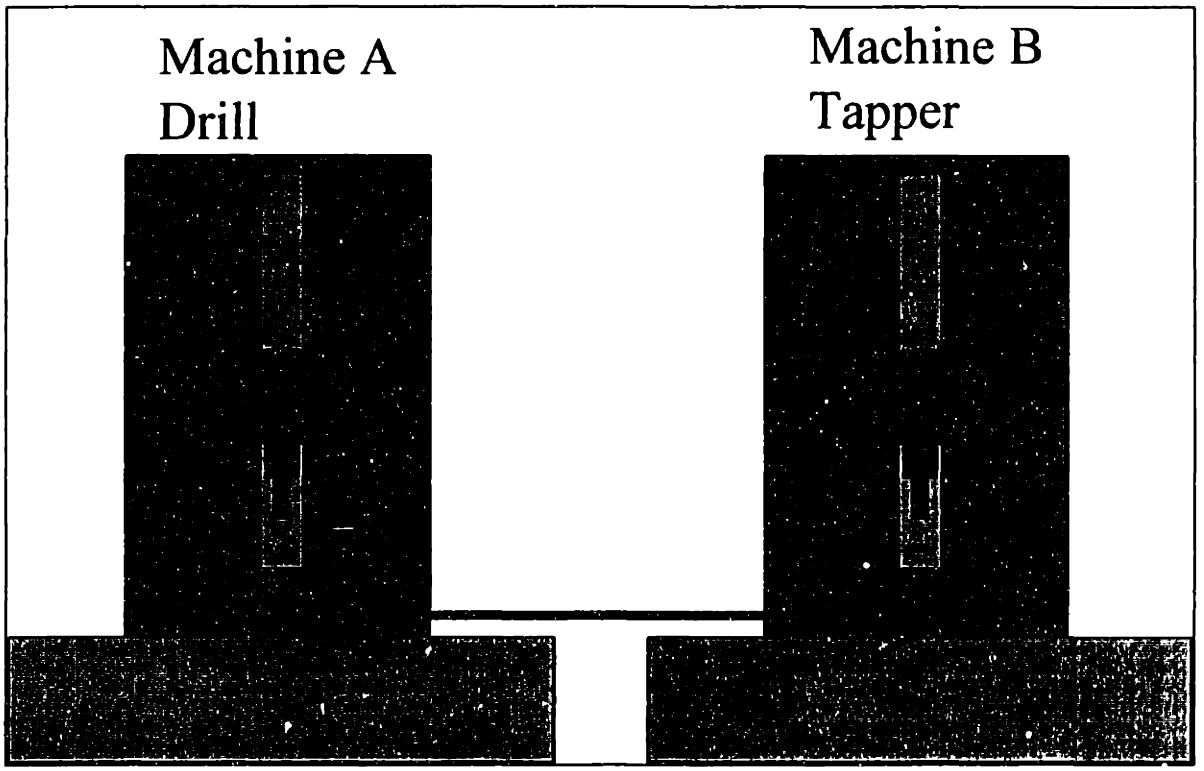


Figure 3.1

every 20 seconds, and machine B one part every 25 seconds. If we plot the production of inventory over time as in Figure 3.2, we see that a slight imbalance of 5 seconds between

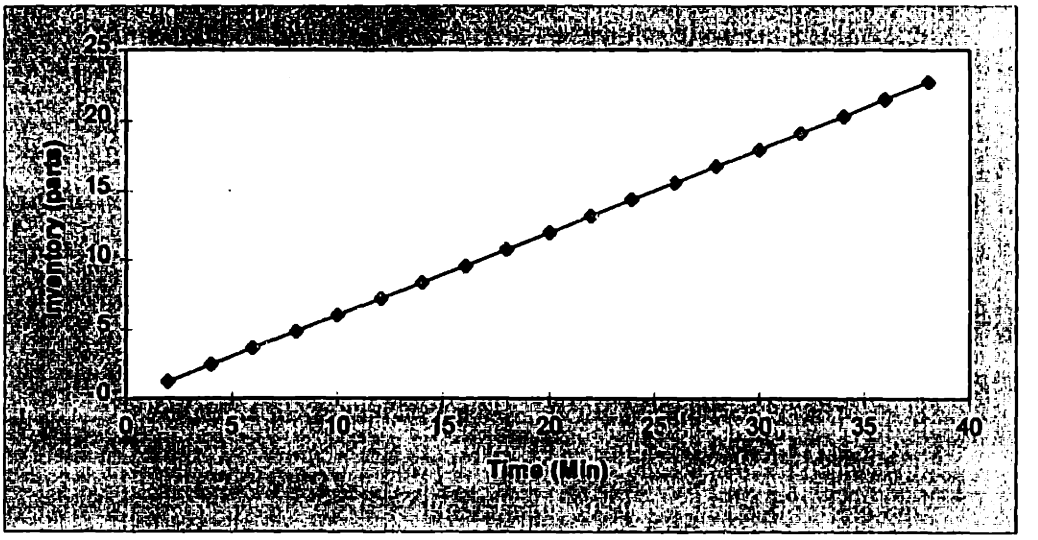


Figure 3.2

two machines can create a serious inventory problem, if both machines are constantly running. The problem escalates with creation of more parts that cannot be immediately processed, eventually calling for a material handling system and a warehouse (in addition to a consumption of floor space). This leads to an increase in the number of parts that are scrapped due to defects and design changes. The costs of the buildup of the "necessary evil" of inventory include capital investment for forklifts, warehouses, occupied floor space, and an inventory control system, and manpower to move the material and people and information systems to keep track of the inventory.

Chapter 4. Reduction of Changeover Time as the Key to Production Leveling

4.1. Introduction

In Ford's day, the era of mass production, the goal of manufacturing was to produce products at the fastest rate possible. Most industries were capacity limited; everything that was made was sold. Thus, the manufacturing philosophy was to set up a machine and then make as many parts as possible before changing over to a new part type. Even if the setup time was inordinately long, one could "spread" this time over a large batch of parts. One produced product mix and variety by carrying large quantities of inventory of each type of part (see Figure 4.1).

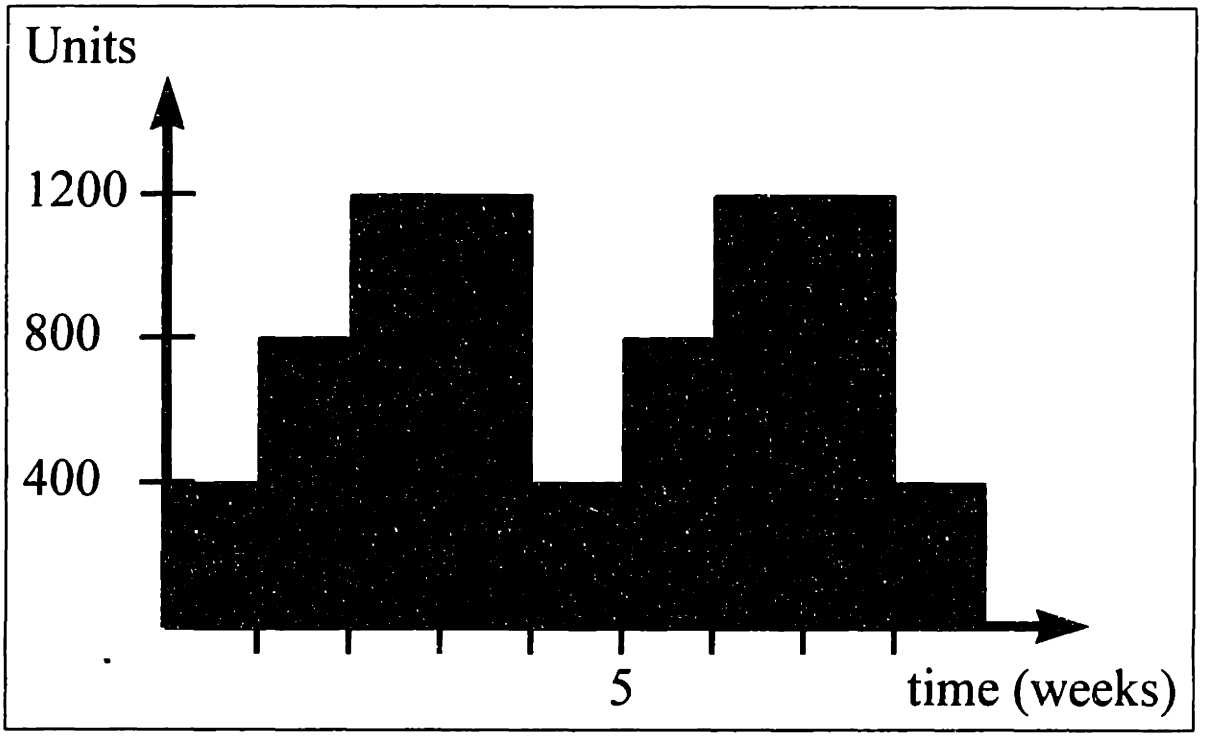


Figure 4.1

However, if one ran out of a type of product, it would often take weeks or months before one could get more of that type of product. For example, if there was an order for just 100 units of product B in week 3, the customer would have to wait more than two weeks before the order would be filled, even though the plant can make 160 units of B in one day.

In today's market, product variety, is demanded by the customer. For example, in the automobile industry, most companies are forced to carry hundreds of different options on each car model. In batch production as shown in Figure 4.1, offering this would require carrying enough stock in each model and options so that one wouldn't sell out before that model and option were produced again. This high level of inventory is called "floating" inventory. Due to the capital costs associated with carrying inventory, it is not cost effective to produce according to the batch production philosophy and satisfy the varying desires of the customer. In order to offer a variety of models and options goods must be produced in small quantities. However, as the batch size (or quantity between changeovers) decreases, the cost of each part will increase, since the changeover time will be spread over fewer parts. This leads to prohibitively high manufacturing costs when changeover times are high. Thus, producing affordable products in small batches (with increased variety) requires reduction in changeover time. After a little reflection we will see that reduction of changeover time is the key to producing products that will meet the varied preferences of customers and also reduces the overall part cost.

We will begin by defining changeover time as the time from production of the last part of type A until the production of the first good unit of part type B, as shown in Figure 4.2.

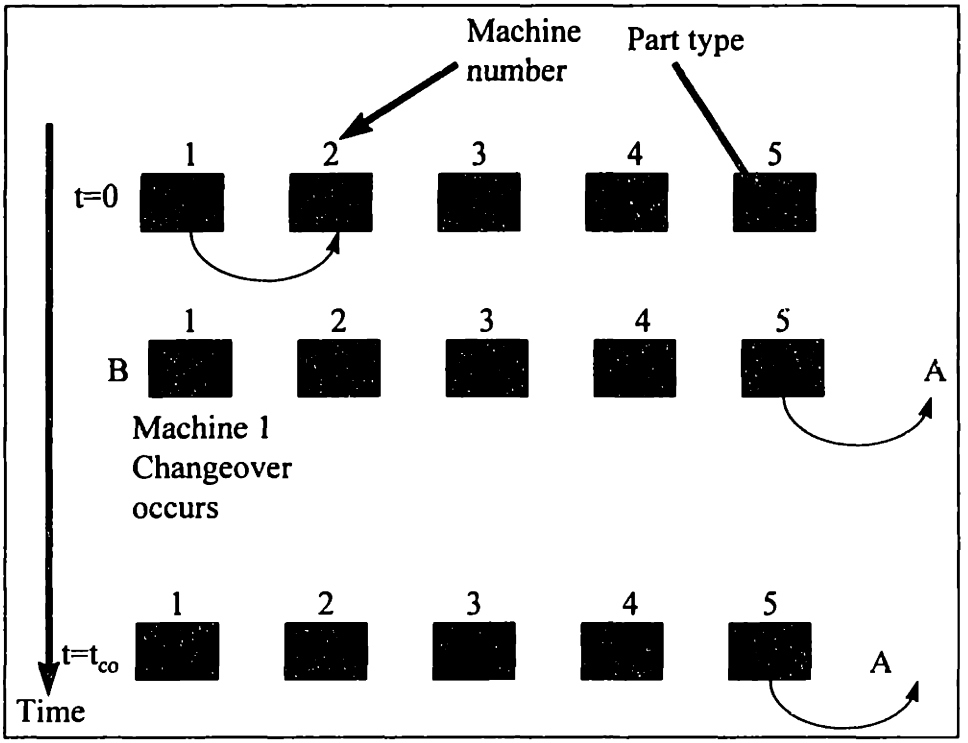


Figure 4.2 Changeover from Part A to B

Thus it includes all the time that the machine is not making parts, as well as the time it takes to run the "test" parts of part type B. The cost of changeover time is two-fold: one part is obvious, the labor that must be paid to the "set up man". The second part is the lost production time while the machine is not making product. These are the direct costs. However, it is the secondary costs which arise as a result of long changeovers and large batch production that cost companies millions of dollars. However, the link between long changeovers and the resulting effects are often overlooked. Moreover, long changeovers are generally accepted as something to work around, rather than something that can be actively reduced. Some managers think only of the direct costs of long setup times, and

think that the reduction in capacity and direct labor for the changeover man are not large enough to warrant investigation into reducing the setup time.

Batch sizes for production are generally set by calculating the economic order quantity (EOQ) given by the following equation:

$$EOQ = \sqrt{\frac{2 \bullet AD \bullet COP}{AICCP \bullet UC}}$$

where AD is the annual demand, COP is the average cost of order preparation(including setup costs), AICCP is the annual inventory carrying cost percentage, and UC is the unit cost. The EOQ is an attempt to minimize the sum of the inventory carrying costs and the changeover costs as shown in Figure 4.3 (Shingo [1985]).

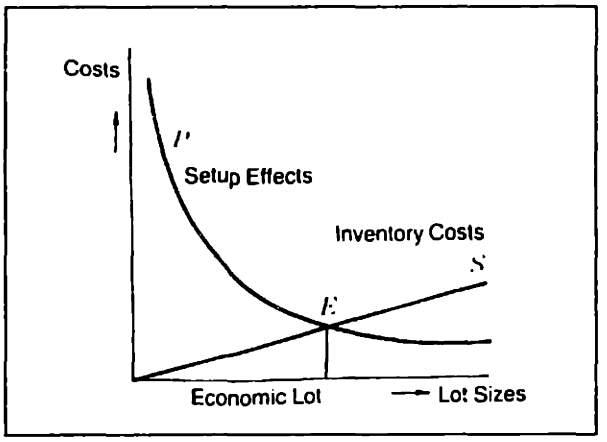


Figure 4.3

If we hold the annual demand fixed, we see that EOQ scales with the square root of the setup time over the product of the inventory carrying cost and the per unit cost. Thus, long changeover times lead to large EOQ's which lead to large inventories (chapter 3) and an inability to produce a variety of parts in a short period of time (product type inflexibility). It is generally assumed that the inventory carrying cost (per part) remains constant independent of the batch size (EOQ) (as is evident from Figure 4.3). Later we

will show that reduction of the batch size actually leads to decreased per unit inventory carrying costs. We can now represent the economic order quantity by the following equation:

$$EOQ = \sqrt{C \bullet COP}$$

where,

$$C = \frac{AD}{AICCP \bullet UC}$$

Next we will derive the cost of order preparation, COP. This will include the costs of generating a work order, plus the costs of setting up the machine(s) (or changing over from the last part type) to produce the lot. We will assume the cost of generating a work order is independent of the size of the order quantity. We are thus left with a relationship between batch size (EOQ) and changeover costs. First, we will consider only direct costs due to changeover. These are (1) the direct labor charges for the time spent by the setup man, and (2) the lost revenue due to machine downtime.

As an example we will show the effects of reducing the setup time from 8 hours (which is common for changing a stamping machine from one set of dies to the next) to 30 minutes on the EOQ. Then we will show how it is possible to reduce changeover times even further using a methodology developed by Shingo [1985]. If we can reduce the setup time from 8 hours to 10 minutes, we can assume that we reduce the cost due to machine downtime by a factor of 48. Concurrently, we will reduce the direct labor of the changeover man from 8 hours to 1 hour (see setup analysis in section 4.2), or a factor of 8. This allows a reduction in the EOQ by a factor between $\sqrt{48} \cong 6.8$, and

$\sqrt{8} = 2\sqrt{2} \cong 2.8$ (depending on the costs attributed to each time). While these appear to be small improvements when compared with overall production costs, we will see next that the effects of this reduction yield savings up to70% of the per unit changeover costs and make the concept of an Economic Order Quantity nearly irrelevant.

4.2. Changeover Analysis

The key to reducing the length of changeovers is asking the questions "What part of the changeover can be done while the machine is running?" and "What is being accomplished by a given (changeover) operation?" In answering the first question, one will find that a very large portion of the previous setup time can be done while the machine is still producing parts, reducing the downtime of the machine due to changeover. In answering the second question, one will find that the direct labor of the setup man can be greatly reduced with very little capital investment.

Shingo (1985) suggests a three step process. He begins by defining operations done while the machine is running (producing parts) as external setup (external to the operation of the machine), and operations which require the machine to be stopped as internal setup. His then suggests the following three step process:

(1) Analyze the setup and distinguish internal from external setup operations.

(2) Convert internal setup operations to external.

(3) Streamline all aspects of the setup operation.

Steps (1) and (2) refer to the first question above, while step (3) addresses the second question. Thus, according to Shingo, to reduce changeover time we must analyze

the changeover operation-by-operation and ask the following questions: (1) "Is this operation usually done while the machine is stopped, or while it is operating?" (distinguishing internal from external operations). (2) For all operation done while the machine is stopped, which are internal setup, ask "is there a way that we could do this while the machine is still operating?" (converting internal setup to external setup). (3) For all operations, internal and external, ask "What is the purpose of this operation? Is it necessary? Could the same functional requirement be served by a simpler, faster operation? If so, what would this require?" (streamlining). However, here we would like to propose a reordering of these steps and add a fourth. We would like to ask the third question first, namely, "Are all these operations necessary?" That way we can eliminate many operations from the beginning, rather than wasting time distinguishing unnecessary internal setup from unnecessary external setup. The fourth question we would like to add is: (4) "For all internal operations, which of these operations can be done in parallel?". Thus, if more than one setup man is available, one can further reduce machine downtime due to changeover by doubling manpower (where possible).

If one analyzes a changeover step by step using this new set of questions one will inevitably find opportunity for great reductions. Shingo (1985) divided changeover time into four categories of operations and estimates their contribution to the total changeover time which is shown in Table 4.1.

Table 4.1

Operation Category	Proportion of setup time
(1) Preparation, after-process adjustment, inspection of raw material, blades, dies, jigs, etc.	30%
(2) Mounting and removing blades, tools, etc.	5%
(3) Centering, dimensioning and setting of operating conditions	15%
(4) Trial runs and adjustment (always internal)	50%

In general, operations in category 1 should be eliminated or converted to external setup, operations in category 3 can be either be eliminated, reduced, or converted to external setup, operations in category 4 (which are all internal setup) must be eliminated, and operations in category 2 should be minimized. From the right hand column we see that elimination of category 4 reduces the internal setup time (which is the most valuable time since it contributes to both direct costs) by 50%. If all operations in category 1 that are currently being done while the machine is stopped (internal setup) are done before stopping the machine, or after the machine has been restarted (converting them to external setups), then the internal setup time can be further reduced by 30% of the original value. This leaves the internal setup time at just 20% of the original length. Further reductions are possible by analyzing operations in categories 2 and 3. However, since the greatest chance for reduction lies in eliminating trial runs and adjustment (category 4), and streamlining operations in category 1, we will concentrate on these two areas, and refer the interested reader to the literature (Smith [1991], Shingo [1985]).

Operations in category 1 include collecting all tools necessary for the setup, checking the raw material for the new part, collecting new blades, cutting tools, jigs, dies, etc., removing any scrap from production of the last part, and cleaning the machine or the

surrounding area. All but the last two operations can be done while the machine is running (and should be). Furthermore, in order to streamline the setup operation itself, it is a good practice to have a "kit" of the necessary tools needed for the setup prior to shutdown (and nothing else). This will save the changeover man time which would be spent looking for a given size of a wrench, or looking for a certain screwdriver. Keeping all changeover tools at each machine is the best practice, because it eliminates the need for a toolbox. However, capital costs for special or expensive changeover tools may prohibit purchasing one set for each machine. In any event, the kit of tools should be brought to the machine while it is still running. Similarly, all new dies, blades, jigs, tooling, and raw material should be brought to the machine prior to the shutdown for changeover. All raw material inspection should take place before it is transported to the manufacturing area (preferably at the supplier).

When changing dies, the new die should be brought to the area and placed on a table just beside the machine. Then when the setup man has all the necessary tools, and the raw materials have been prepared, the machine can be stopped, and the old die removed. The old die should be placed just next to the machine, minimizing die handling while the machine is down. The above operations will require small capital investment in the form of tables for the new and old die. If two forklifts or die carts are available, removing the old die and inserting the new die can be done in parallel, removing the need for tables. Capital investments may also be necessary in the form of a tool box, or tool rack for the kit of tools, as well as a rack for the new set of blades, jigs, tooling, etc. These costs are rather small, and will be insignificant when compared with the savings they provide in terms of direct labor, and lost time due to machine downtime.

Eliminating adjustments and trial runs is a more difficult endeavor, and may require more capital expenditure, as well as substantial time investment and ingenuity. The key concept that must be followed is the following: in making a part, every degree of freedom of the machine must be specified and fixed. In a stamping machine, if one is installing a die, there are two degrees of freedom for locating the die. One can install two stops, one in each degree of freedom, which can be used to locate dies and eliminate adjustments. In machining operations, installation of new cutting tools often requires trial runs and adjustment. This can be eliminated if the location of the tip of the cutting tool (with reference to the tool holder) is known in advance of installation. "Quick Change Tooling" is available now, and is becoming increasingly economical. However, in many cases this will be unnecessary, if the cutting tools (especially new tooling) is standardized. Thus, each time a new cutting blade is installed, the operator can set the machine to a pre-specified location, and begin cutting without trial runs or adjustment. For slightly worn cutting tools, a labeling system can be implemented which establishes the location of the cutting tool, so that it can be reinstalled without trial runs and adjustment. A further source of variability comes from the interaction between the part and the machine. Each part from the preceding process will be slightly different (in structure, dimensions, temperature, etc.), and this will induce additional degrees of freedom that should be taken into account when processing a new part.

In general, trial runs and adjustment are the result of improper (or lack of) gauging, preset points, or general knowledge about the location of various parts of the machine that is being setup. Installation of gauging, frequently used settings (for higher volume parts), and physical stops (for dies, blades, etc.) can eliminate adjustments and

trial runs. These modifications to the machine will cost capital, as well as time to study the machine and design proper gauging. However, the investments will often pale in comparison to the resulting benefits, which include reduced internal setup time, reduced operator labor, and reduced scrap, as well as improved quality.

4.3. The Benefits of Reduced Changeover Time

Using the above methods, companies have been able to reduce four hour setups to 3 minutes (Shingo [1985]). This would allow one to reduce the EOQ (based on internal changeover time) by a factor of $\sqrt{80} \cong 9$. Shingo [et. al] asserts that most changeovers can be done in less than 10 minutes (internal setup time). For machines where this is possible, we can see the opening for an entirely new manufacturing philosophy. The following example will show that one no longer needs to think of an economic order quantity:

If a changeover previously required 8 hours, and the economic order quantity had been calculated to be 500 units, we can see that 480 minutes of setup time has been spread out contributing approximately one minute to each part. Producing a lot of only 100 units would raise this quantity to 5 minutes of setup (non-value added) time per part, which could substantially increase the cost of the part. However, if the setup time can be reduced to 10 minutes, we see that one could produce a batch of 10 parts (instead of 500) and still have only one minute of setup time per part (Monden [1983]).

In the previous section (4.1) we assumed that the carrying cost of inventory was independent of the batch size. However, we can see now that this is not true. If parts are

made in batches of 500, one will most certainly need a bin, or a rack to store the parts. One will also need a forklift to transport them. However, if parts are made in batches of 10 then one can transport them much more easily to the next processing area. Furthermore, one batch will occupy 1/50 the space, and thus will eliminate the need for a separate storage area (as well as the material handling to and from the storage area). Thus, we see that the inventory carrying costs are not independent of the batch size.

We can also see that the cost of the part can also be decreased by decreasing the changeover time. This is accomplished by reductions in storage space, material handling, scrap from trial runs, reduction in overhead costs (setup time is often charged to overhead), and improved quality through elimination of adjustment.

The greatest benefit from reduction in changeover time is not the reduction of direct costs. It is the ability to produce parts in small batches. This enables the factory to produce a variety of parts in each production period, which shortens lead times, and provides greater flexibility to respond to variation in customer demand. This fact is particularly critical to aerospace applications, since production volumes are relatively low. Let us look again at Figure 4.1, but this time with small batch sizes.

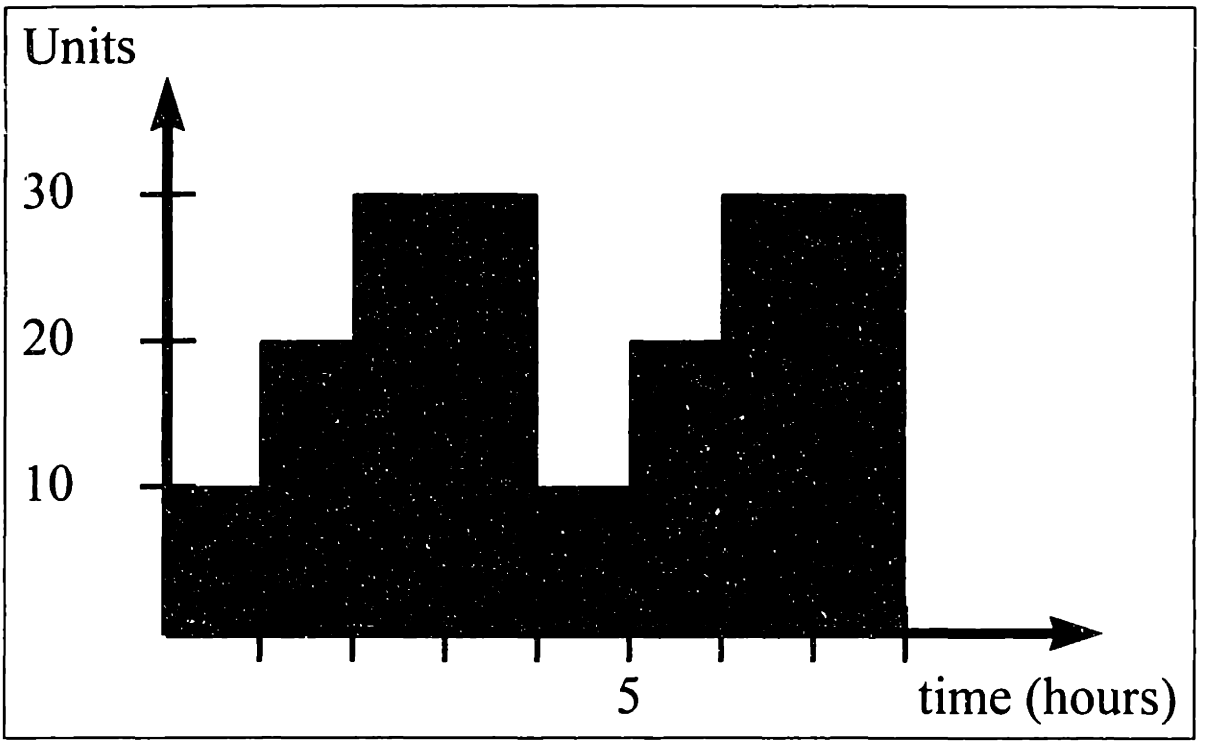


Figure 4.4 After Production Leveling

We see that a customer who wants to order any of the three part types can have their order filled that day, since all parts are made twice a day (note the run time per part is the same for both Figure 4.4 and Figure 4.1). Figure 4.5 shows a comparison of Figure 4.1 and 4.4 for just one 8 hour day of production. This reduction of the batch size is known as production leveling.

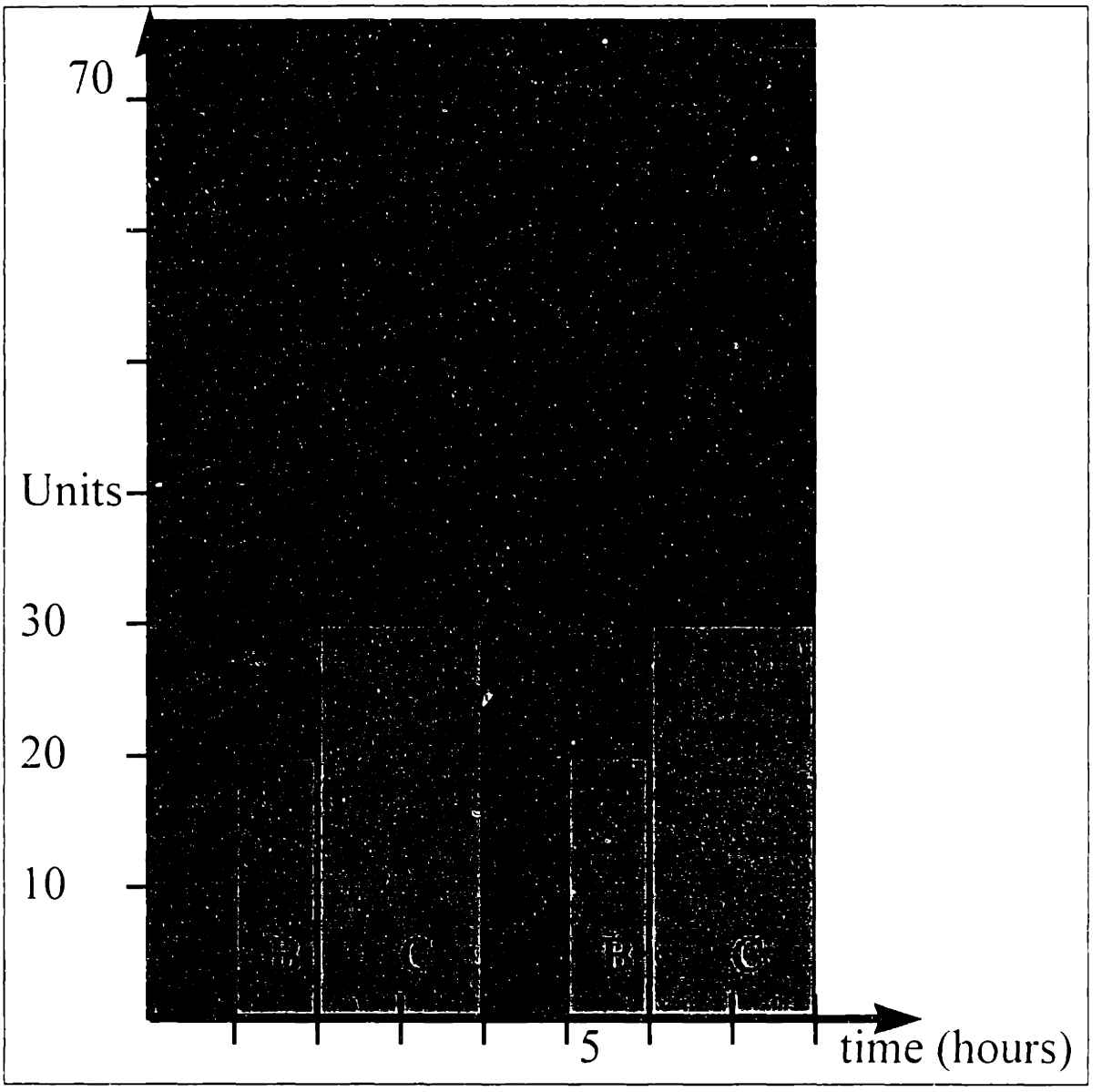


Figure 4.5

If changeover times can be reduced to allow production of parts in small lot sizes, then the factory can respond to changes or spikes in demand by changing the production schedule (while still remaining cost-competitive). If the setup time for the part types A, B, and C as in the figures above to less than ten minutes, we can change Figure 4.4 to level the production, as is shown in Figure 4.6.

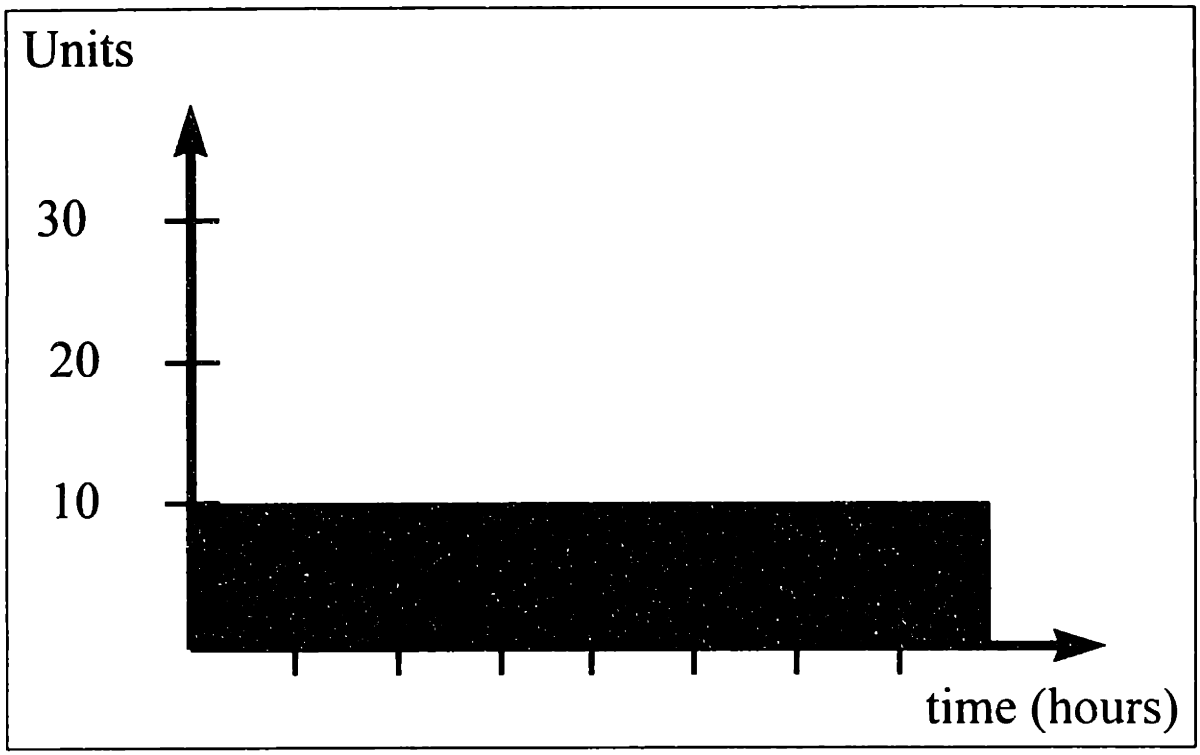


Figure 4.6

In summation, the large batch sizes of mass production (500 units and higher), were a direct result of three factors: (1) a producers market; whatever was made would be sold, so product variety was not a priority, (2) long changeover times, which required spreading the direct costs over a large number of parts and (3) the desire to optimize machine utilization. The second factor brought about the concept of an economic order quantity (EOQ) which sought to balance the costs incurred in changeover (which are independent of lot size) with the costs due to carrying extra inventory (which were assumed to be directly proportional to the batch size), as shown in Figure 4.7 (Figure 4.3 shown again).

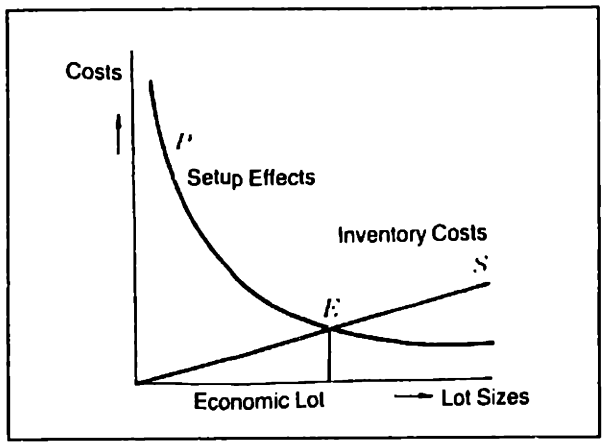


Figure 4.7

EOQ is an attempt to find the point E, which minimizes the sum of the two costs. An analysis of the variables in the derivation of the EOQ was given which yielded additional information not captured explicitly by the formula and showed that inventory carrying costs are not directly proportional to lot size as Figure 4.6 claims. Following that, section 4.2 detailed a standard procedure to reduce setup times which will yield roughly an 80% reduction in internal setup time. The benefits of this reduction include reduced inventories, improved quality, reduced material handling and storage costs, and, most importantly, the ability to implement cellular manufacturing with multiple part-small lot manufacturing, which is the subject of chapter 5.

Chapter 5. Cellular Manufacturing

5.1. Introduction

The concept of a manufacturing cell has been around for at least 50 years, from early work by Mitrafanov (see Black [1991]) and Burbidge [1962] on Group Technology. The quantity of literature written to date on cellular manufacturing would be enough to fill a small library. However, what is missing from that literature, and what this chapter attempts to create, is a comprehensive design framework of a manufacturing system. This chapter will not attempt to cover all the concepts involved in cellular manufacturing to the level that has been covered in individual journal articles. Instead, it will present a top down look at the system, and show how a broad system approach, when designing the individual cells and individual machines for these cells, can avoid many of the pitfalls that have been encountered by industry in their attempts to replicate the Toyota Production System.

Much work has been done which addresses individual issues in cellular manufacturing (such as scheduling, grouping of parts into "families", automation, set up reduction, etc.), but often these papers (or texts) approach the problem from a descriptive level. They discuss an individual case study or set of case studies, or propose modeling techniques to be used for simulation or part grouping. However, for an individual or team embarking to design (or redesign) a process or an entire plant, it is necessary to take a broader scope of the system. This chapter is intended to be a manual for determining (or designing) the production capacity of the process, choosing the philosophy for part grouping, designing the initial quality control devices, designing or redesigning machines

and deciding machine layout to set up the process flow. Testing the results from the preliminary work using simulations will yield information on scheduling sequences that will reduce set up times. Any observed backtracking of parts will lead to adjustments in machining and processing sequences.

The following chapter will provide a framework for understanding and designing a cellular manufacturing system based on a product and a demand rate. A mini-case study at Briggs and Stratton is presented in Appendix 2 to illustrate the use of the framework, and serves to add more detail and insight to the results of the process.

5.2. Cellular Manufacturing vs. Job Shops

Cellular manufacturing (CM) as a system for production refers not only to the layout of the machines or stations, but also to the flow of product. To transform a factory which produces with a traditional (or job shop) layout (see Figure 5.1) and push flow into

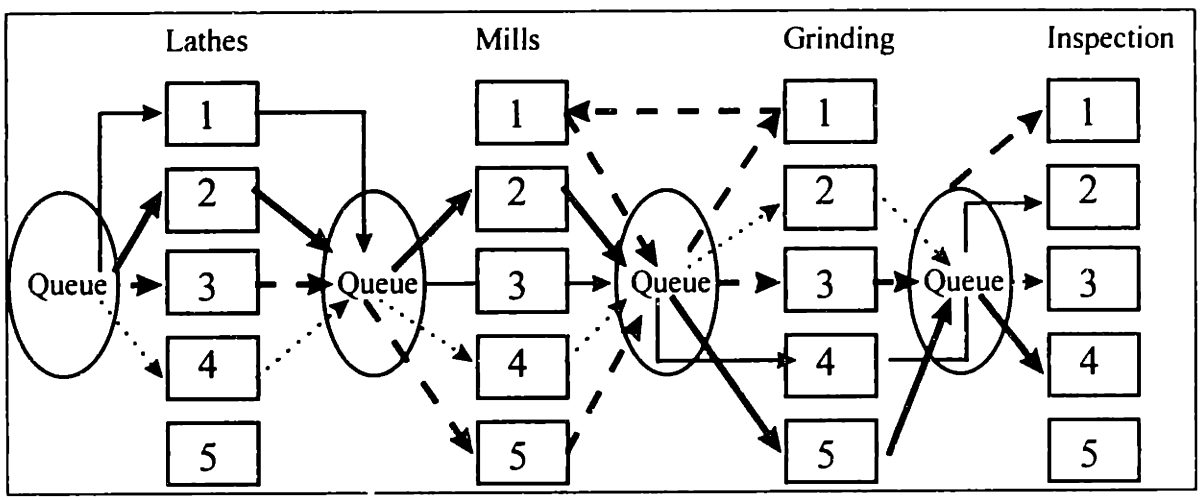


Figure 5.1. Job Shop

a cellular manufacturing system with pull flow (see Figure 5.2) requires a completely new philosophy of production. It will be informative in understanding cellular manufacturing to frequently compare CM to a job shop manufacturing system.

A job shop uses a functional layout or departmentalized layout of machine tools where all machines of the same type (i.e. mills) are placed in the same location. This layout results in routing complexity, transportation wastes, and increased manufacturing lead time as can be seen from the part flow paths in Figure 5.1.

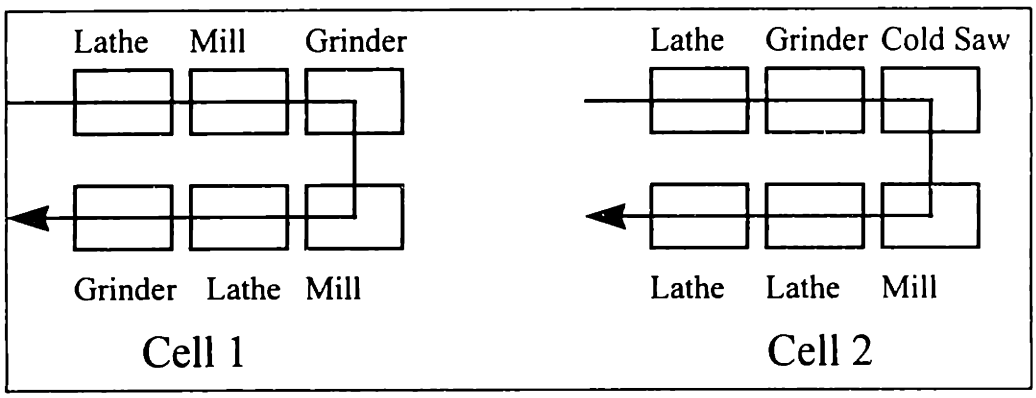


Figure 5.2. Cellular Manufacturing

In cellular manufacturing groups of parts, called part families are processed completely in clusters of machines called cells (see Figure 5.2). The flow of the product through a cell is unidirectional. A part being processed in a cell can skip a machine(Figure 5.3a, Part 2) but cannot backtrack (Figure 5.3b).

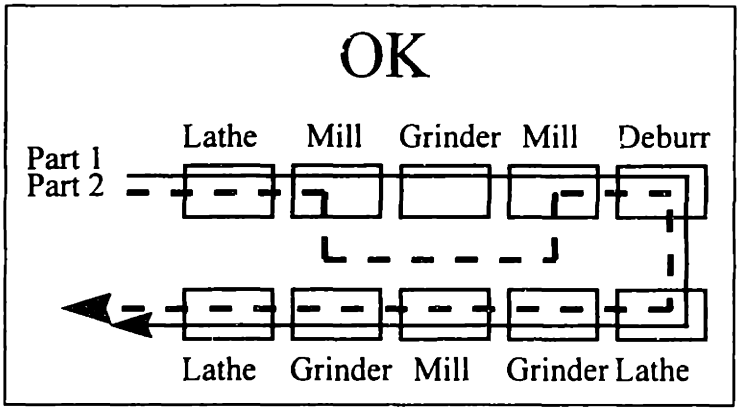


Figure 5.3a

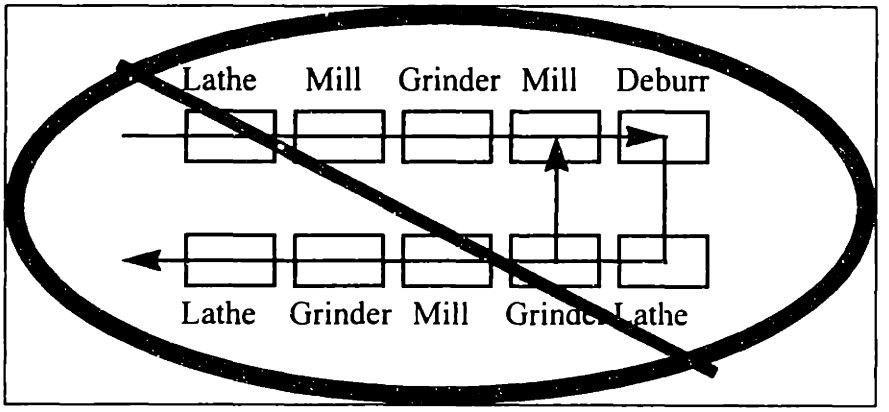


Figure 5.3b

In addition, each part, belongs to exactly one cell and it is easy to determine which machines it will be processed on. In contrast, in a job shop environment a part passes from one process area to the next, being processed on whichever machine is available as shown in Figure 5.1.

A cellular manufacturing system can be viewed as a chain of machines or assembly stations. A cell is a logical unit within this system which can be viewed as a single link as in Figure 5.4.

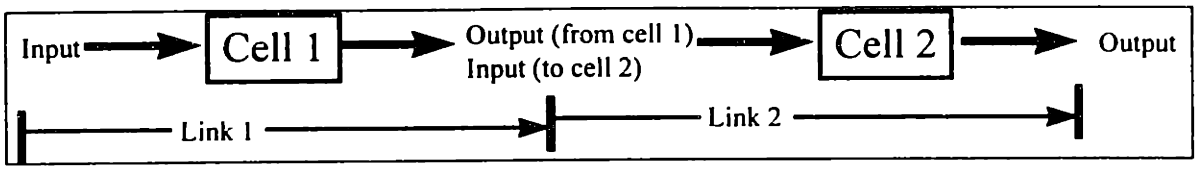


Figure 5.4 Production system with two links

5.3. Manufacturing System Goals

The goal of a manufacturing system is to create defect free products as efficiently (i.e. lowest cost) as possible. Job shop performance is often evaluated on the utilization of the machines. Each operator attempts to keep his machine running at all times. This philosophy results in creation of a large quantity of inventory, since some machines produce parts at faster rates, and some slower. The result is that some machines will be starved (and their utilization will be less than 100%), while others will have a pile of inventory in their queue. This inventory, which Shingo refers to as the waste of stock on hand (Shingo 1992), leads to the waste of transportation, and the waste of storage (see chapter 3).

In a cell, the goal is to have parts flowing from one machine to the next at a rate determined by the takt time. Even when cells are running at capacity, utilization is important only for the "bottleneck" machine. Creation of inventory between machines or between linked cells is viewed as a sign of an unbalanced system which must be balanced. The result is that the inventory in cells is often orders of magnitude less than in job shops.

5.4. Producing to Takt Time

In most plants, the given quantity of an item required in a day's production is known. If it is known exactly how long it will take to produce a given part then the production rate can be set to meet the "takt time". The takt time is given by the following relation:

$$\text{Takt time} = \frac{\text{total time available for production per shift (in sec)}}{\text{required number of parts per shift}}$$

The takt time provides a marker or goal for the cell operators. The cell's goal is to produce parts at a rate equal to the takt time. If cells are linked, then they should all produce to the same takt time. If two cells, A and B are feeding final assembly which uses two parts from Cell A and one part from Cell B in each assembly, then the takt time of Cell A should be twice that of Cell B which should have the same takt time as final assembly. If a cell is producing parts faster than the takt time, there will be a build up of excess inventory. Thus a factory must strive to balance the entire factory to the production rate of final assembly, which should match customer demand rate (see section 5.11 and 5.14). The most effective control mechanism to limit the flow of production to that of final assembly (or the furthest downstream operation in the factory) is the "pull system".

In the pull system, the signal to begin production is given by the downstream machine rather than by a production planner. The shipping area pulls product from the final assembler who in turn pulls parts from the subassemblies who pull parts from the machining area who pull parts from casting, and so on. Upstream machines or cells do

not begin production until they receive a signal from the downstream link, in the form of a Kanban card or simply the removal of material from the stations finished goods area. If every area follows the rules for this system, the production rate of each station, and cell will be set by the shipper (who pulls parts at a rate equal to the takt time).

For example, in the defense industry, the rate of production often varies throughout the year, even on a fixed contract, in order to maintain the skills of the workers in low volume production periods. Implementing a pull system in this industry would keep the entire manufacturing system, from raw materials to final assembly, producing at the same rate. Each process area would adjust their production rate to meet that of final assembly.

5.5. Creating Part Families

In order to begin the transformation from a job shop to cellular manufacturing (or in startup situations) one must decide which parts will be processed in which cells. Groups of parts which will be assigned to a given cell, or linkage of cells are termed "part families". The aim is to assign a certain family of parts to a cell such that each cell will be able to take advantage of reduced changeover times, while having sufficient volume to make good usage of the machines and avoid capacity limitations in other cells. Part families are generally based on Group Technology, or formed around a product line.

5.5.1. Group Technology

Group Technology (GT) is a method used to create part families which has been recently expanded to include codification and classification schemes (see Bibliography).

Part families are formed based either on similar part geometries or similar process routings. If one chooses to group parts based on geometric features, e.g. small shafts, then one should group parts together such that the common geometrical features will allow for common fixturing and common processing sequences. Common fixturing will reduce changeover times, which will facilitate lower batch sizes (section 4.3), and common sequencing will allow for machine grouping which avoids under-utilization of equipment. Grouping parts based purely on size should be avoided, since it does not reduce fixturing costs and does not facilitate good utilization of resources.

Part families based on process routings are created by looking at the processing sequence of all parts and grouping those parts together that require the same machinery and can be processed in the same sequence. The advantage to this method over geometrical groupings (which generally have smaller changeover times), is that the information necessary to create part families (the process routing sheets) is readily available whereas effective geometric coding of all parts can be a time consuming process and sometimes fail to produce satisfactory results, in terms of machine utilization or part processing sequences.

5.5.2. Product Line Cells

Another way to determine the family of parts to be processed in individual cells is to select a product line and manufacture all the parts in that product line in one cell or one group of cells linked to final assembly. This family grouping allows for the array of cells producing a given product (or line of products) to be "a factory within a factory" (Drucker [1992]) having its own cost accounting structure, maintenance staff, and other support

labor. The cost accountability makes it is easy to determine the direct and indirect labor costs for the family of parts.

A key benefit of product line grouping is that one can create a manufacturing system which produces one "set" of parts in each cycle that can be assembled to form one final product. For example, for a car, one would create an array of cells that would produce individual parts at a rate which meets the takt time for the finished product (a completed car). Thus the rim cell would produce rims four times as fast as the chassis cell , and twice as fast as the door cell (if the final product was a two door car). Inventory in this type of system can reach the ideal limit, which is one part per machine or assembly station, with no parts in between process steps.

5.6. Production Environment: Machine Flexibility, Material Handling

Parts can be produced in a multitude of different production environments, some of which are listed in Table 5.1. When choosing from any of the systems in Table 5.1 one should be careful and understand the benefits and drawbacks of each system before investing valuable capital. Depending on how the work will be performed and the type of material handling, one may choose from systems that are very flexible within their part family (Manned Automatic Cells, FMS) but which have difficulty adjusting to design changes or addition of new parts.

Table 5.1 Types of Systems referred to as Cells

"Cell" Type (processing)	Material Handling	Work Method	Control System	Inspection	Machine Flex.	M.H. Flex.
Manual Work Cells (serial)	Manual	Manual (assembly or manual machines)	Manual	Manual or Poka-yoke	High	High
Manned Automatic Cells (serial)	Manual	Single cycle automatics, Dedicated machines or CNC	Manual	Manual or Poka-yoke	High for CNC, low for others	High
Transfer Lines (serial &/or parallel)	Auto	Manual or Auto	Auto	Manual or Auto	High for manual work, low for auto	Low
FMS (parallel)	Auto	CNC	Auto	Auto	Low	Low

Auto- Automated, FMS- Flexible Manufacturing System

Others (Manual Work Cells, Manned Automatic Cells) are very flexible to volume changes, since their production rates are generally limited by the number of operators rather than the machines' cycle times. In addition, these two also process parts serially which is a prerequisite for tracking the processing of parts (see Chapter 6). Table 5.1 is laid out in order of capital investment, from lowest at the top to highest at the bottom, suggesting that Flexible Manufacturing Systems and Transfer Lines should be reserved for high volume, long product life parts which will not undergo significant design changes that might make pallets or fixturing obsolete.

In all manned cells, operators are responsible for multiple machines or assembly operations. In contrast, in a job shop there is generally a one to one correlation between operators and machines (this is sometimes called the "one man, one machine" paradigm).

In addition to higher labor costs, a lack of multi-skilled workers makes production extremely sensitive to absenteeism.

5.7. Batch Size

The batch size refers to the number of parts that are processed on each machine before moving the "batch" to the next operation. Since one part is generally processed at a time, the other parts in the batch must wait, either before the machine, waiting to be processed, or after the machine, waiting for the rest of the batch to be processed before moving to the next operation. When changeover times are large, the job shop philosophy is to process parts in very large lots, to "spread" the changeover costs across the batch of parts (Chapter 4). In contrast, in cellular manufacturing, changeover times are reduced by grouping families of parts to promote common fixturing and thus do not require lengthy time-consuming changes in the configuration of the machine. However, since frequent changeovers occur within a family of parts, reduction in changeover times is a prerequisite for cellular manufacturing families of parts. This active approach to reduction in changeover time eliminates the need to "spread" changeover costs over a large batch of parts.

5.8. Reduction in Lead Time by Eliminating Lot Delay

Large reductions in lead time can be gained from the flow philosophy of cellular manufacturing. Since there is no queue between machines in cells (as compared with job shops- see Figures 5.1 and 5.2), there is no need to wait for the entire batch to be processed before sending a part on to the next machine . Thus, parts can flow one by one

from one machine to the next in what is termed "single piece flow". The lead time (LT) to produce X parts using single piece flow(SPF) is given by:

$$LT_{SPF} = \sum_{i=1}^{n} MCT_i + (X-1) \bullet MCT_{BN}$$

where n is the number of machines, MCT is the machine cycle time for each machine, and BN denotes the bottleneck machine (the machine with the longest MCT). Producing parts in batches(as is done in job shops) yields the following expression for batch production lead time (LT_{BP}):

$$LT_{BP} = LS \bullet \sum_{i=1}^{n} MCT_i$$

where LS is the lot size. If we rewrite the equation for single piece flow as follows

$$LT_{SPF} = X \bullet \sum_{i=1}^{n} MCT_i - (X-1) \bullet \sum_{i=1}^{n} MCT_i + (X-1) MCT_{BN}$$

$$LT_{SPF} = X \bullet \sum_{i=1}^{n} MCT_i - \left[(X-1) \left(\sum_{i=1}^{n} MCT_i - MCT_{BN} \right) \right]$$

(and note that X=LS), we can subtract LT_{SPF} from LT_{BP} to obtain

$$LT_{BP} - LT_{SPF} = \left[(LS-1) \left(\sum_{i=1}^{n} MCT_i - MCT_{BN} \right) \right]$$

The factor in brackets, is simply the time for one less than the total quantity of parts to be produced, multiplied by the sum of the machine cycle times not including the bottleneck machine. As the batch size increases and the length of cycle time not including the bottleneck machine increases, the benefits of single piece flow increase greatly.

It should be noted that in general cells are designed so that the operator's motions are the "bottleneck", and machine cycle times are only important for maximum capacity

calculations. In this case, the number of operators will have a large affect on the length of the "bottleneck" activity. As a result, the difference in lead times between single piece flow and batch production will be dependent on the staffing of the cell.

5.9. Scheduling

One of the major difficulties in managing a plant is the scheduling of parts. The scheduler must assign priorities to all parts, make sure all parts are completed by their due dates, and (in a job shop) keep machine utilization as high as possible. The design of a job shop system makes the scheduler's job a nightmare. It is not uncommon in a job shop to have half the parts "red-tagged" (a priority tagging which gives the part a priority in each of the queues in Figure 5.1), because they are overdue. Red tagging parts is no longer effective in setting part priorities if two red-tagged parts arrive at a queue simultaneously. As the number of red-tagged parts grows, there arises a need for a priority list among "red-tagged" parts to decide the priority of parts should two red-tagged parts meet in a queue. Thus, the priority of red tagged parts problem compounds itself leading to a part by part priority list and an army of expediters.

As a result of these scheduling difficulties, the lead time of each lot of parts depends on the quantity of red-tagged parts in the shop, and whether it is "bumped" from its position in a queue by expedited parts. The lead time of a lot of parts at any given time is unpredictable and many orders fail to be completed by their due date.

In cellular manufacturing each batch of parts must be scheduled only once. Each part type belongs to one family which belongs to one cell (unless cells are duplicated) or linkage of cells where the part is processed from start to finish. The scheduler's main task

is to schedule parts so that setup time is minimized, without making parts (which require a large quantity of changeover time) wait an inordinately large amount of time. Thus, scheduling in a cellular manufacturing environment is greatly simplified resulting in elimination of overdue parts, greatly reduced lead times, reduced number of schedulers, and lower production costs. For algorithms to schedule cells which process a large number of part types, see the literature review in the Bibliography.

5.10. The Effects of Lead Time Unpredictability

One effort to reduce the quantity of overdue parts in a job shop is to include a buffer in the planned lead time for a given lot of parts. This pushes the release of the work order for the lot back several months. The release of the order to the floor signals the material handler to buy the raw materials and begin processing the part. If the part does not encounter disturbances (being bumped by red-tagged parts) then it may spend months in the shop either as WIP or as finished goods. The longer a part spends in the factory, the greater likelihood that it will be damaged, and it may even be scrapped due to a design change or an order cancellations. Furthermore, the capital that was used to buy the raw materials and then to pay for processing the goods is tied up until the goods are sold to the customer.

In cellular manufacturing the lead times of parts are predictable because there is no queue for parts to wait in, and the cycle times of parts belonging to a given cell are known through standard work instructions. Thus, it is easy to calculate the time required for a given lot to be produced in the cell which manufactures that part type. Capacity calculations in cellular manufacturing are much simpler than in job shops, due to the fact

that capacity is calculated for each cell, rather than for an amalgamation of machines, which does not take sequence of processing into account. By knowing the capacity and lead time for each part type the scheduler is able to release the work order for a part much closer to the due date (than in a job shop) and still be confident that it will be completed on time. This allows for design changes to be built into the part and reduces the chance that the part will be damaged (from transportation or storage) or scrapped due to an order cancellation or design change.

5.11. Balancing the System

The maximum throughput of any process is limited by the machining, assembly, or inspection step with the longest cycle time (the time needed to complete the given operation) which is designated the "bottleneck". However, the chief goal of the process engineer is to balance the entire system, so that all linked cells and processes are producing parts at the same rate and that they are producing them at the same rate they will be used in the final assembly. Thus, every cell that processes a given part should be producing to the same Takt time.

In optimizing a cell for maximum capacity the process engineer works with each part type to balance, or equilibrate the cycle times of all machines. This is done through optimizing the tool paths or assembly operations or by shifting workload from one machine to another or between assembly station (see section 5.14). The eventual goal is to balance all the machining times (or assembly station times) so that when the cell is running at capacity, no station will be idle. When this is achieved, storage of parts and

waiting by operators are completely eliminated; an operator will reach each machine just as it finishes its machining cycle.

However, once capacity has been established for a process, the process engineer must be concerned with setting up operator movements so that the Takt time can be met. This requires dividing manual tasks so that all operators have an equal workload. This is a challenging task, since operators often have different physical skill. The result of an imbalance due to operator differences will be a creation of idle time, or Shingo's [1989] "waste of waiting".

Achieving true flow is by no means a simple process. It requires dividing up the processing operations among the machines, as well as deciding on a sequence for the operations for each part type. It requires visibility between linked cells so that imbalances in production rates can be discovered and eliminated.

5.12. Tracking Defects in Manufacturing

In a job shop parts are processed in large lots which move as a unit from one station to the next. It is not uncommon for an entire lot of parts to be scrapped due to a machine malfunction since the problem will not be detected until after the entire lot has been made and transported to the inspection station. Furthermore, it is generally difficult to determine which machine processed a given part, and operators will sometimes "pass the buck" and claim they didn't make the defective part, in order to defer blame for the costs of repair. In job shops, quality problems are very difficult to solve.

One of the main benefits of a cellular manufacturing system is that the path of each part can be easily tracked through the plant. Thus, when a defect is found, the

problem can be traced back to the individual machine which processed it, and the cause of the defect can be eliminated. This "traceability" of defects leads to greatly improved quality as well as reducing the need for inspections. Often poka-yoke ("mistake - proofing") devices will be built into machines which prevent defective parts being made by halting the machines (see section 6.4, Shingo [1986], Hirano [1988]). These inexpensive methods ensure that all operations are completed properly and check that all necessary parts at that step have been added. In addition, go-no go inspection devices will be placed in between machines which prevent defective devices from being passed on to the next station (see source inspection, sections 6.2 and 6.3).

5.13. Designing Cells

When designing cells, one must be certain that the prerequisites for cellular manufacturing are in place before attempting to shift to cells (or in designing new cells). The prerequisites include reliable machines, short (<10 minutes) changeover times (for cells which manufacture multiple part types), and an able work force. One can quickly see that cellular manufacturing cannot be successfully accomplished without taking into account both the system and the machine (see Figure 5.5).

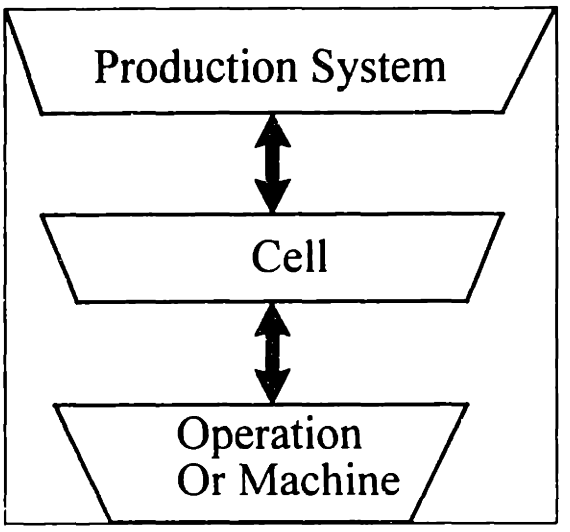


Figure 5.5

Each level's operation interacts and therefore places constraints or functional requirements on the adjacent level. One must keep in mind both the big picture (the production system) as well as the smaller picture (machine or operation design) when designing a cell or an array of cells. At the system level one must take into account customer demand rate (which determines the Takt time), length of product life, and skill level of operators. At the process or machine level one must be careful to design machines which are ergonomic, easy and fast to load and unload (proper fixturing), have minimal changeover times, and machine footprints (the rectangular size of the machine on the floor) which reduce operator walking distances. Only if all these issues are considered in designing cells will the full benefits of cellular manufacturing be achieved.

5.14. Cell Design Methodology

When moving to cellular manufacturing, one must remember that cellular manufacturing requires quick (<10 minutes) changeovers, reliable machines and a willing, able (cross-trained) workforce.

I. Begin with a finished product that will be sold as is to customers. For the final product: translate demand from customers into a Takt time for each individual part in the final product. For takt time use 7 hours 40 minutes = 27,600 available seconds per shift (or adjust depending on the labor contract). Then assess how long the product will be in operation, and the likelihood of design changes, and their impact on the process, in terms of fixturing, machining, etc.

II. Break out parts according to size and weight. Those parts that are too large (require two hands), too heavy (>10lbs) or too small (parts which can be grabbed between two fingers may present handling difficulties) are candidates for automated material handling systems, i.e. transfer lines.

III. For remaining parts, obtain estimates of machining/assembly times for each operation, including machine time and manual time. Machining times may come from past experience, or a material removal data handbook . Assembly times may come from a Boothroyd and Dewhurst estimate, or from timed samples. 2 ft / sec should be used for walking speed to estimate operator travel times between machines.

IV. Survey existing equipment and assess capacity by comparing the required processing time for each operation with the takt time. If designing or buying new equipment, buy machines with enough capacity such that predicted customer demand is 85% of capacity (the cycle time of every machine should be less than 85% of the Takt time) using bottleneck or theory of constraints analysis. Thus, the cell will not be running

at 100% of capacity (see Black [1991] on designing cells to run at less than 100% capacity.) to stay with customer demand, and will be able to increase production should customer demand increase.

V. If capacity is greater than demand translated into Takt time, (i.e. if total of machining cycle time plus manual time for each machine is less than Takt time) then go to step VII.

VI. If there is insufficient capacity, then the following adjustments can be made[1]:

a) Reduce the number or length of operations currently done on the bottleneck machine by shifting them to lightly loaded machines.

b) Reduce manual operations through better fixturing, or kaizen of machine loading (or installation of devices which automatic the unloading of parts).

c) Consult design engineers to reduce the time required for feature generation, i.e. reduce depth of drilled hole, increasing dimensional tolerances to allow for a more robust or faster process[2].

d) Investigate the possibility of using an alternate processing method. For example, one may be able use a cold saw to remove a large quantity of material faster than a mill. If designs is sufficiently robust finish passes can be eliminated which will greatly reduce machine cycle times.

---

[1] *There is a common mistake made which is to duplicate the bottleneck machine. Proponents of this argue that having two machines will aid in situations where one machine goes down. However, duplicating machines creates the variability inherent in having two different machines doing same processing step. It also eliminates the trackability of parts which is necessary for accurate defect resolution. In general, it is not good to sacrifice process repeatability for partial production in machine downtime.*

[2] *This also can have large effects on quality. See section 6.1 designing quality into parts.*

e) For machining applications, consult engineers in the upstream process to see if they can form the part using a near net shape method which will reduce the quantity of material to remove.

f) Reorder sequence of operations. This may reduce fixturing time, or allow removal of material on a machine which can do so more rapidly, i.e. doing large face milling before drilling operations.

g) Remember that the strength of manned cells is the volume flexibility. A quick way to add volume (if possible) is to add another shift, starting with one operator for the entire cell (or system of linked cells) and then adding more manpower incrementally as is necessary. This will increase overhead costs, but in plants which are not currently operating on three shifts, it will save the company the capital investment in new equipment by taking advantage of the volume flexibility of cells.

g) Investigate new technologies (high speed machining, automated operation).

h) As a last resort, consider duplicating the bottleneck machine (but see footnote 1 on page above, and avoid this measure if at all possible).

VII. Layout

a) The layout of the cell is very important. Two general layouts have been the most successful. These are the U shaped cell as shown in Figure 5.6,

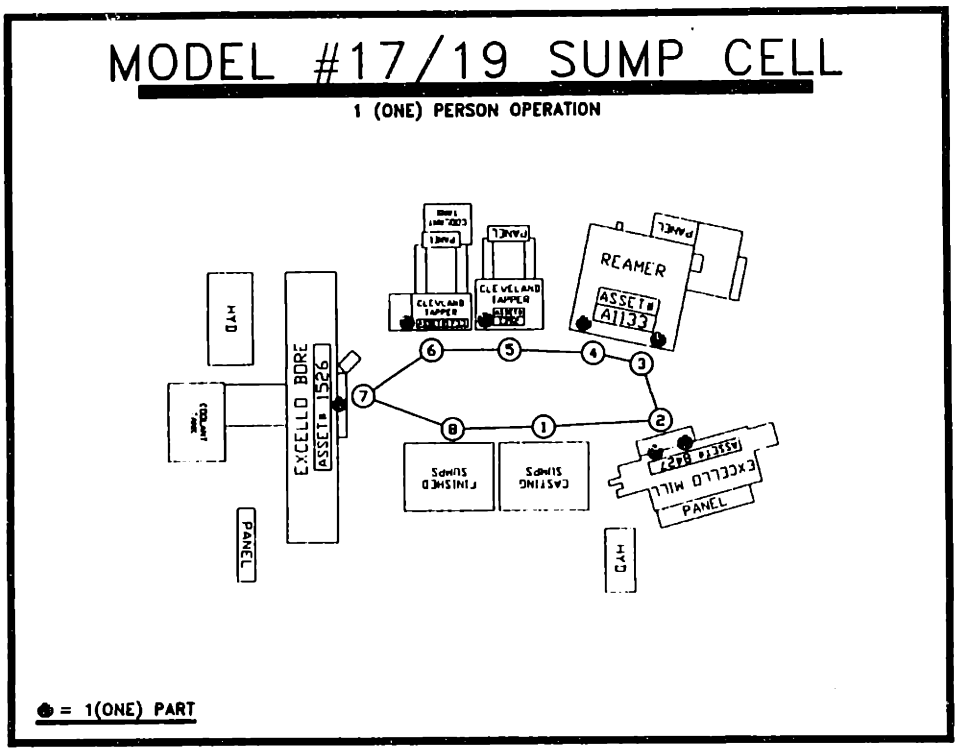


Figure 5.6

and two parallel rows, as in Figure 5.7.

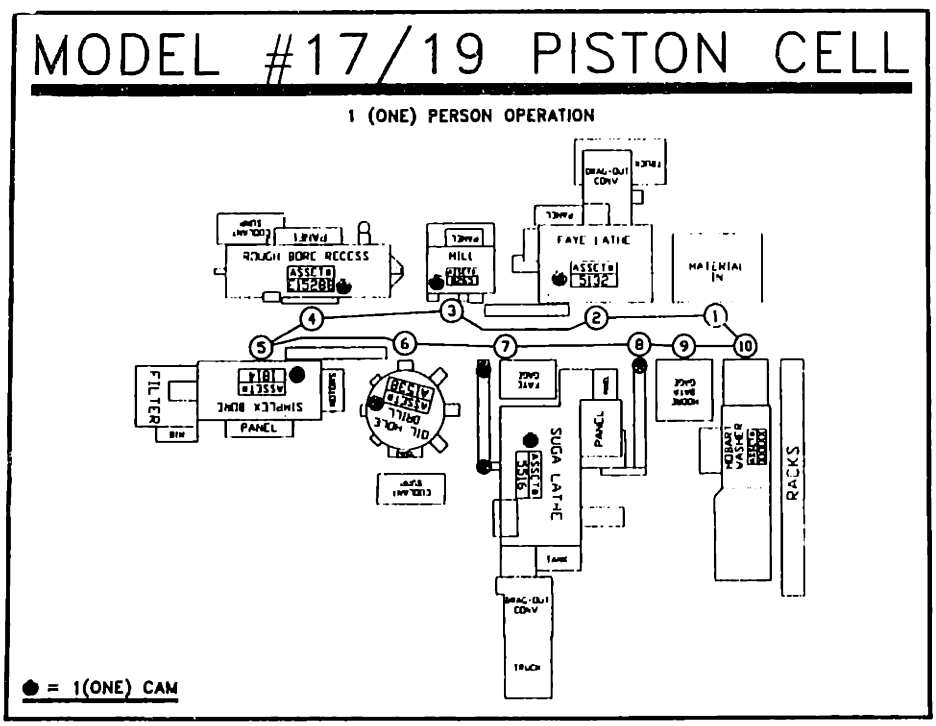


Figure 5.7

The aim is to have an operator who is running the entire cell start and end one loop in the same place to minimize walking time and distance. A layout using two parallel rows of machines facilitates access to the cell for changeover, machine repair.

b) For cells where there will be several different part types which will use different combinations of machines, the cell should be laid out to minimize travel time and distance for the operator. For a cell with two part types, which each require one machine that the other part type does not use, the best layout (if processing sequence allows it) is given in Figure 5.8.

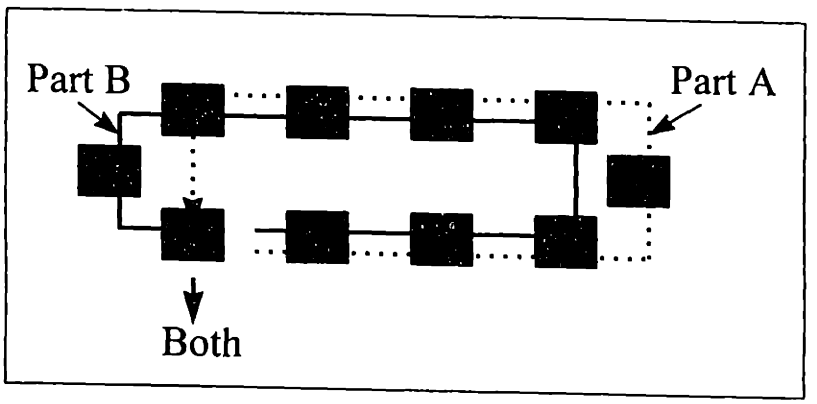


Figure 5.8

For cells where there are several part types, all of which share a few machines, but some of which require additional machines, the layout should be (if processing sequence allows it) as is Figure 5.9, where the order A-B-C represents the ranking in terms of volume of parts produced in a given time period.

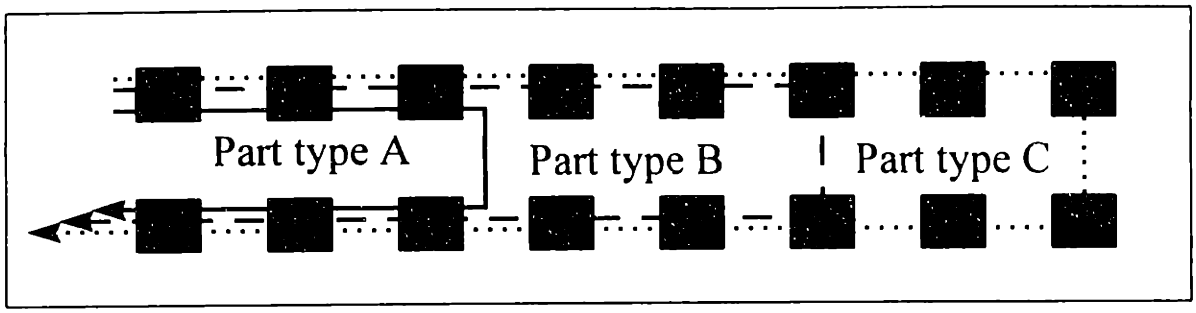


Figure 5.9

VIII. Balancing the cell.

a) Once cells have been created, and sufficient capacity is available, the next task is to balance operations. There are three methods of balancing a cell:

(1) Combine operations into one machine or one assembly station to increase the loading on these operations.

(2) Divide operations on a machine that is heavily loaded. These operations can be done on a completely new machine, or added to a lightly loaded machine (or assembly station).

(3) Analyze each operation to see if there is a way to improve the process which results in reduced processing time, either manual or machine time.

IX. Designing for multiple operators. Analyze cell operation for a single operator. Add a second operator by dividing the cell roughly in half, and attempt to create a division such that the operators each share 50% of the manual and walking time. Repeat this process for 3 operators, 4 operators, and so on. When the manual time for each operator is equal to the bottleneck machine cycle time, the cell will be operating at capacity. For each number of operators, go back to step VIII and balance the cell. With

more than one operator, one can balance manual times by sharing loading and unloading duties on machines that form the junction between two operator's loops as in Figure 5.10 at the Mill.

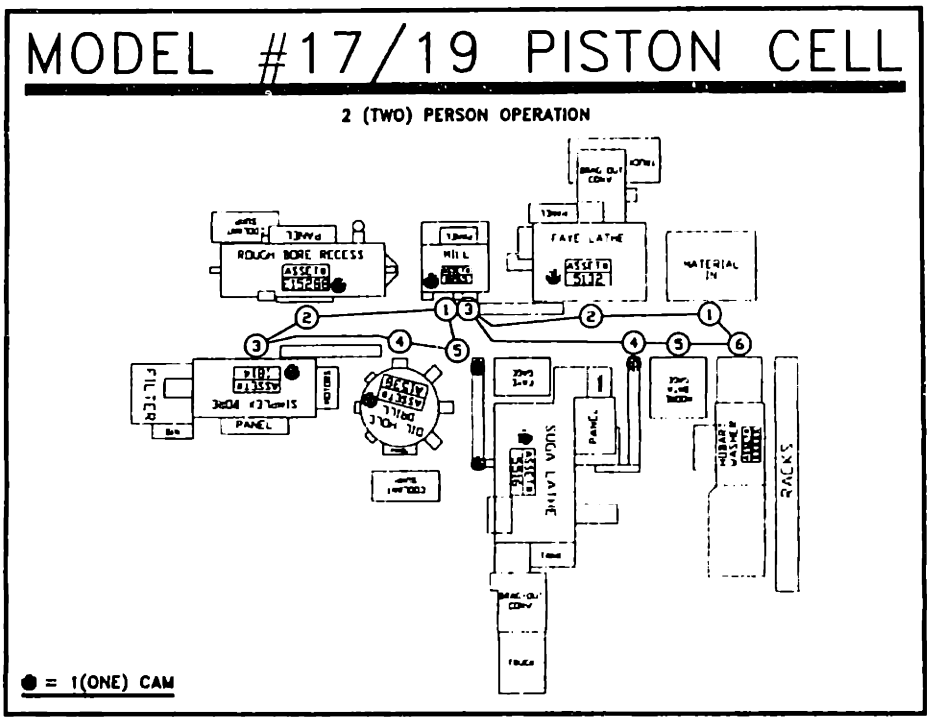


Figure 5.10

IX. Installation of quality control devices. Ideally, one should work with design engineers to create part designs which are robust enough to handle the variation inherent in the process that will be used to manufacture the product. However, in cases where this is not economically feasible, one must build quality control into the machines. This can be achieved through installation of poka-yoke devices, and simple go-no go gauges to check part characteristics.

a) The first step in deciding which machines to mistake proof and which features to inspect is to perform a priority analysis.

(1) Assess the likelihood of each defect (from past data, or from statistical analysis of test runs).

(2) Assess the likelihood of detecting each defect which depends on the inspection device, method and frequency.

(3) Assess the magnitude or effects of each defect. This can be measured in dollars, or rework time, or other measures.

(4) Calculate the priority value (PV) for each defect from the following relationship:

$$PV(defect_i) = \frac{p(defect_i) \bullet C(defect_i)}{p(\text{det}ection(defect_i))}$$

where p() denotes the probability of occurrence, and C(defect) is the cost of the defect.

(5) From this list, perform a cost benefit analysis for devices or methods that will eliminate or reduce each defect. Then implement as many as possible. For more details on poka-yoke devices see section 6.4 and Shingo [1986], Hirano [1988].

X. The following tips will aid in creating and optimizing cells:

- Operator movement should be counter clockwise when viewed from above the cell. This facilitates loading for right handed people (especially on lathes and other non-front loading machines).
- Installation and work should be done at a height of 36-38"
- Always use critical features as a reference when fixturing parts. This has substantial effects on quality,.
- Minimize changeovers by using a single fixture for multiple parts

- Minimize fixturing costs and machine purchases in drilling and tapping operations by combining these processes on one machine.
- All switches should be walking switches which reduces manual time by 3-5 seconds depending on fixturing.
- Carpal tunnel syndrome can be minimized by replacing push button switches with light sensor switches.
- Make a to-scale floor plan and cutouts to-scale of machine and play with different layouts, analyzing the walking distance of the operator, accessibility of the machines to maintenance, available floor space, and links to upstream and downstream processes.
- After initial design is complete, present to operators for feedback. Experience has shown that operator input prior to creation of a preliminary design reduces the efficiency of the design process. However, operator input is invaluable to the optimization process, and should not be neglected. Only by adopting a philosophy of continuous improvement with the operators as agents will it be possible to achieve the full benefits of cellular manufacturing.

Chapter 6. Quality in Manufacturing

6.1. Designing Quality into Parts

The definition of a defect (see Taxonomy Chapter 1) is a part which fails to conform or perform to a desired requirement or purpose. The purpose of any product is to serve a function, and thus each part, and each design can be thought of as satisfying functional requirements (see Suh [1990]). Since it is often difficult to test whether a part satisfies the functional requirements, we specify characteristic dimensions and limits which the part must fall within. However, the inherent problem with this method is that the translation of functional requirements into dimensional requirements often results in either labeling a part defective which is in fact functional, or accepting a part which is not functional.

Thus the ideal inspection is one that tests whether a part meets functional requirements. The functional requirements are met by design parameters; thus if we want to create defect free parts, we must design parts so that they meet the functional requirements despite the variation inherent in the processing technique. Retrospectively, parts can be redesigned so that process variation does not impair performance. In short, the best way to improve quality, or to build quality parts from the beginning is to find a design that is so robust that parts meet the functional requirements despite variation in the processes.

6.2. Improving quality in processes

Improving the quality of a process is no small endeavor. Many company wide practices such as Total Quality Management, Total Preventative Maintenance, and Quality Circles dedicate large quantities of resources and look at quality as something that relates to all aspects of a factory. In this section, only the quality of a given process will be discussed, and only methods to improve quality through reducing the number of defective parts that are made before the defect causing problem is detected and fixed. In essence, this section will illustrate how to transform the present practice of inspection, where parts are inspected at the end of a process (fabrication or assembly). Final inspection, as it is called, results in long defect-problem solving time lags, during which more defective parts are often made. The transformation from final inspection to successive check and then (when possible) to self-check, will essentially shorten the response time to fix defect causing problems.

One goal of a manufacturing system should be **to produce** 100% defect free products. Since this goal is often viewed as unattainable, many factories aim at a smaller goal: only **ship** defect free products. This goal has led to the practice of inspecting parts just before shipping, after all processing and assembly steps have been completed. Inspections at various intermediate stages are done in cases where defects may be undetectable later (but cause the product to malfunction in the hands of the customer). However, if all features of a part or product can be inspected after the final operation, many plants choose to have only final inspection. However, in many cases, defects cause parts to be reworked or scrapped, and often all the work that was completed after the defect occurred will be wasted (if the part is scrapped), or need to be redone (as in an

assembly when the product must be disassembled and then reassembled). Thus, intermediate inspections can save the company the wasted labor and processing that goes into these defective parts. However, the key point that must be remembered is that inspections only detect defects, they do nothing to prevent them. So the only way a factory can prevent defective products from reaching the customer is by implementing 100% inspections on all functional aspects of all parts before they leave the factory. However, inspections cost the company money and fixing all the defective parts costs even more. Clearly it is best to catch problems as soon as possible after they occur, or prevent them from occurring in the first place.

One way to reduce the time lag between the creation of a defect and detection of the problem (during which more defects are often made) is by implementing more inspections. The cost of inspections including labor in the form of an inspector's wages, capital for inspection tools, and lost revenues as a result of increased cycle time (adding inspections adds another step to the process) makes it economically unfeasible to have a quality inspector inspect all parts after each operation. Some methods to reduce the inspection costs have taken advantage of the characteristics of some defects (defects due to variability) and implemented SPC (statistical process control) which allows them to inspect a small portion of the parts and infer the quality of the rest. The difficulty with this is obvious; it is possible for a defective part to slip through, all the way to the customer. Clearly if one wants to have only defect free products leave the factory, 100% of parts must be inspected.

Instead of reducing the number of inspections by relying on statistics, one should reduce the time required for each inspection, as well as increase the number of

inspections that can be done by the operators themselves. In this way, the cost of inspecting 100% of the parts can be reduced to a feasible level. An added benefit is that if one can detect problems immediately after they occur, the company can save money on rework (or even warranty work), so that these inspections will actually pay for themselves as well as prevent defective products from reaching the market.

The key two step process to implementing 100% inspection while eliminating costly final inspection by a quality engineer is the following:

(1) First implement successive checks at each handoff, and devise ways to inspect the work that has just been created in a quick and simple way. A successive check is an inspection by the next operator to touch the part, and is done before any additional processing is done.

(2) Convert successive checks to self-checks where possible. Successive checks are used so that an unbiased operator can check the work, preventing the operator from passing on his own defective parts down the line. The key to progressing from successive checks to self checks is making the inspections entirely objective. For machined parts, the part can be slipped into or over a mold that can check to make sure the part has the correct dimensions. This is exceptionally simple for checking the location and depth of drilled holes, checking the outside contour of machined parts, and checking that all parts have been added to an assembly (perhaps though a simple weight check). These devices can be used in self checks, since the go-no go decision is objective, as long as the operator does not try to pass a part that fails the test. Moving to this type of inspection will require honesty, ingenuity and creativity. One must instill the idea of responsibility for quality in both operators and supervisors so

that they do not try to pass on defective parts, and so that they can devise devices to quickly check their work.

6.3. Reducing the Defect-Detection Time Gap

The most important improvement that is made by a move from final inspections to successive checks to self checks is in the reduction of the time gap between creation of a defect and its detection. Figures 6.1 - 6.3 show how this time gap shrinks as one progresses towards self inspection.

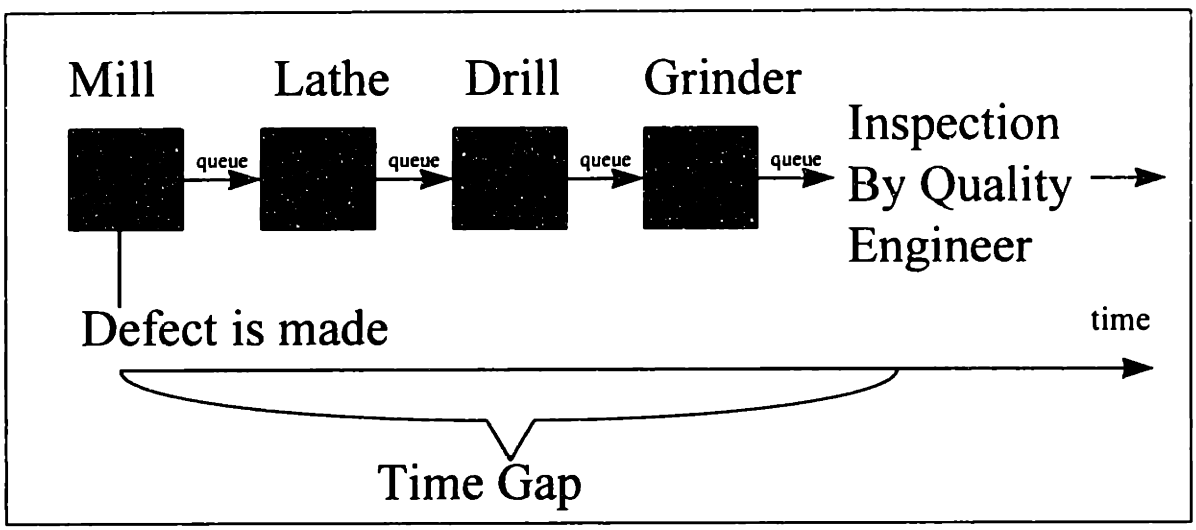


Figure 6.1

In Figure 6.1, the time lag includes all operations that happen to the part after the defect has been made and before the defect will be detected (more defective parts can be made during this time if the defect is due to a broken machine tool, improper machining method, or other problems that do not create simply one isolated defect).

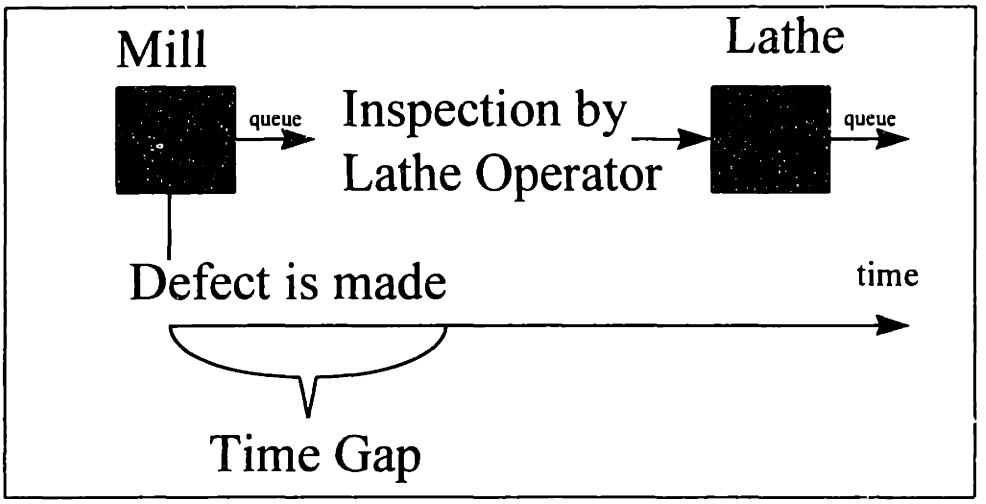


Figure 6.2

In Figure 6.2 the time gap shrinks to the length of time before the operator of the next machine handles the part. In a job shop, this may be a significant quantity of time and if parts are produced in batches, often the entire batch may have the same defect. However, in cellular manufacturing this time lag is small, since the queue is only one unit.

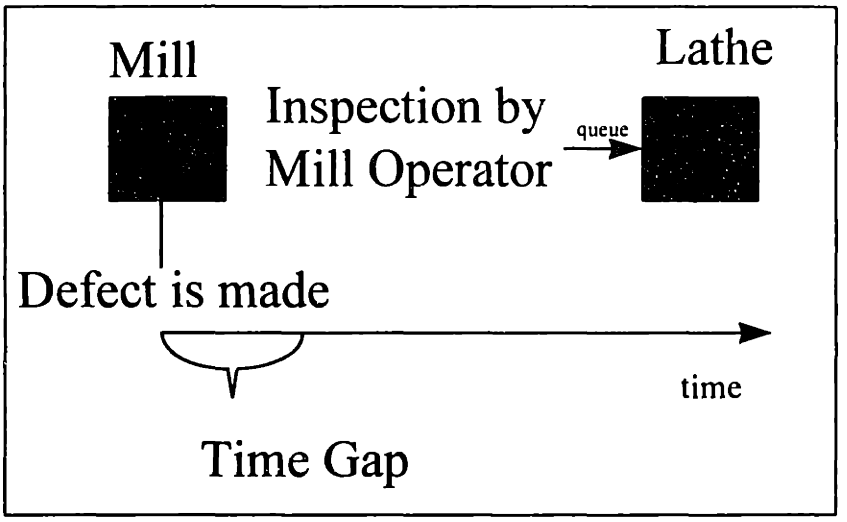


Figure 6.3

In Figure 6.3 the time lag has shrunk down to the amount of time that the operator spends on the given operation before he or she checks the part. Self-inspection produces visibility of the problem after the first defective part is made (if it is detectable).

Catching defective parts prevents adding more value to parts that will be scrapped or reworked later. Clearly this reduction in time lag can lead to: quicker and easier detection of what the problem is that is causing the defect, reduction in wasted time in the form of value added to scrapped parts, and wasted time spent assembling a part that will have to be disassembled and then reassembled. Overall, quicker elimination of defect causing problems will result in a reduction of the number and cost of bad quality parts.

6.4. Poka-Yoke Devices

Since the goal of cellular manufacturing systems (CMS) is to produce the desired (by the customer) quantity of goods at the exact time they are needed with minimal inventory (and zero overproduction), it is essential that the quality of these parts be perfect. Many quality engineers would say "Impossible. Manufacturing processes are by nature variable, and defects are a fact of nature. The best we can do is to keep them to a minimum." This kind of thinking has led to statistical process control, frequent inspections, and material review boards. However, if one thinks with this mindset, one will never achieve perfect quality. The best one could hope for will be to catch all defects before they reach the customer.

However, we can take a (perhaps) idealistic approach. If we ask "What must we do to insure perfect quality in all the parts we make?" we will find the answer is not to inspect the parts before they are shipped, and not even to inspect parts after each operation. It must be understood that inspections only discover defects, but never prevent them. Granted, some statistical methods can follow the variation in a feature's dimension and alert the operator that the machine is about to make a defect. However, these

methods rely on probabilities, and are never 100% certain. What we are aiming for, what we require, is that no defects are made.

There are two ways to achieve this. One was given in section 6.1, designing defect free parts. The only other way to achieve zero defects is to understand exactly what causes defects, and eliminate the source of the defects, rather than eliminating the defective parts once they have been produced. A brainstorm of causes of defects might yield the following for a machining operation:

- improper alignment of work piece
- improper tool path
- broken tool
- material abnormalities
- operator errors

For an assembly operation, the following defects are common:

- missing parts
- incorrect parts
- over-torqueing/under-torqueing of fastening devices
- improper alignment of assembled parts

The most common categorization of defects is done by a fishbone or Ishikawa diagram with four categories: man, machine, material, and method, as in Figure 6.4.

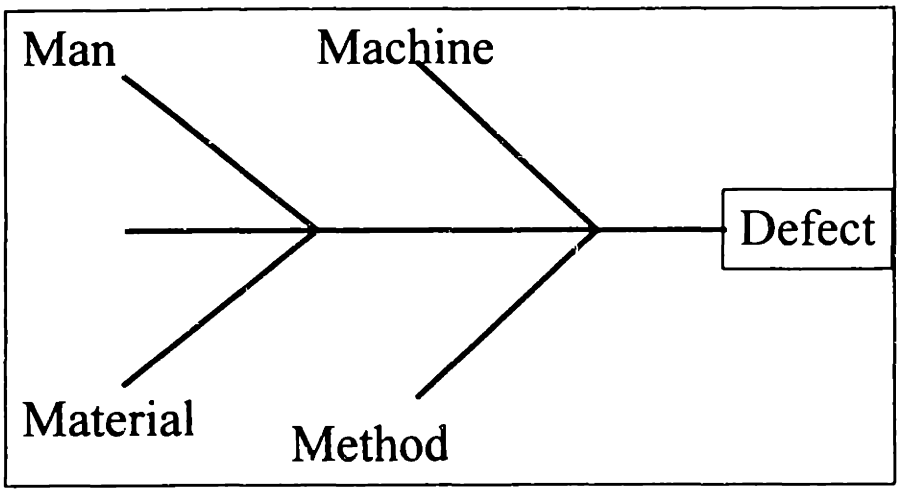


Figure 6.4 Fishbone diagram illustrating possible defect causes

In each of these categories there are literally hundreds of factors that can lead to defective products, making the task of eliminating them all seem insurmountable. However, if we look instead at how the defect occurred rather than its cause, then we can prevent it from occurring in the first place. For example, if we try to stop an operator from making mistakes by chiding them for their mistakes, or if we only give them better work instructions, we are bound to be disappointed. Human beings are fallible and mistakes are bound to be made. However, if we make it impossible to make a defect, then we open up the possibility for perfect quality. This is not to minimize the benefits of good work instructions or worker training, but merely to illustrate that this alone will not prevent defects from occurring. Good work instruct n cause significant reductions in the number and frequency of defects.

Let us now begin to look at the way defects occur. When a machine with a broken tool contacts a part, defects are bound to occur. One answer to this problem might be preventative maintenance, checking and replacing the tool after certain intervals. However, accidents can occur, and tools that appear sound or functional may actually be

defective, or precariously close to breaking. How can we solve this problem without running extensive tests on a tool before each part is made? These are the kinds of questions that must be answered if we are to achieve zero defects or perfect quality.

Let us start with some simple concepts. For assembled parts, clearly we must find a way to make certain that all parts that should be added to an assembly are done so properly. The best way to do this is set aside the parts that are required for a given assembly, and before seals are closed, or the assembly moves on, check that no parts are remaining. This is a rather simple, common sense notion, but one that is rarely followed in industry. Often some parts will be "kitted" for a given assembly, but other parts that are used in large quantities (sometimes called pan stock) are taken as needed from a bin. For example in the aerospace industry, most nuts and bolts are not included in kits for most assemblies. It is no surprise that one of the most common assembly defects is a deficient quantity of fasteners or other pan stock items. Since it may not be efficient to include these in the kit for an assembly (due to material handling costs, or quantity of use), one can, in essence, kit the pan stock items at the assembly station. For example, if an assembly requires 12 bolts, 8 washers, 6 lock-washers, and 8 nuts, these parts can be collected before beginning the assembly and then placed in a small bin to use for the assembly. This eliminates a common problem, which is that assemblers often forget that they have not installed a given bolt, or seal, or washer, but have no way of checking. Sometimes the defect is hidden by a cover, so the assembly continues to be processed until it is tested (or perhaps by the customer). On hand-kitting with pan stock items, combined with kitting of larger parts can eliminate nearly all "missing parts" defects for assembly operations.

Another frequent defect that occurs in assembly is misalignment of parts, or incorrect sequencing of part layers. These can be solved in two ways: (1) To **reduce** defects, the part can be visually inspected and compared with a drawing of a correctly assembled part (but this will not remove the defects that cannot be seen by a short visual inspection). (2) To **eliminate** defects, we must find a way so that parts cannot be installed incorrectly. This may require redesign of the parts (adding symmetry, making asymmetry more extreme, adding alignment pins or tabs, etc.) but can sometimes be achieved through extensive and clear labeling, detailed work instructions, and a simple inspection device which checks alignment of the assembly.

In machining, it is often more difficult to prevent defects from occurring and easier to detect defects after they have been made. This has led to the design of many devices which check important dimensions following each machining step, and prevent the part from moving on if it is defective. While this saves the plant the value that will be added to the part (which will be later scrapped) in the next set of operations, and alerts the operator to the problem, it does not eliminate the cost of scrapping (or sometimes reworking) the part. Defects in part machining are often caused by misalignment in loading, broken tools, improper machining, or incorrect tool paths/part programs (on a CNC machine). One capital intensive method to eliminate defects caused by incorrect machining is to purchase (or retrofit, if possible) a machine with a sensor that tells the controller where the cutting tool is (in three dimensional space, referenced to the part) at all times. This is often very costly, but can prevent defects due to environmental factors (thermal expansions, propagated vibrations, etc.) if the sensor is equipped to detect for these kinds of variations in operating conditions.

A less expensive way to eliminate (most) defects is to study both the part and the tool path and see where defects are caused. If it is a drilling operation, for example, and the most common defect is an incorrectly located hole, a limit switch can be installed that prevents the machine from drilling (prevents motion in the axial direction) unless the drill tip is in a certain location referenced to the part (perhaps using four light sensors to designate a location through which the drill must pass).

To eliminate misalignment during loading of parts, the most common solution is to install contact sensors that prevent the machine from starting up, unless the part makes contact with the fixture device in a predetermined number of locations (these locations should be determined by kinematic analysis). If the points are selected correctly, contact with these points will assure correct gripping by the fixture device. These devices have solved many quality problems where the source of the defect was just a slight misalignment of the part.

At this point we can see the systematic framework that we must use to eliminate defects. Rather than make assumptions about a machinist's skills or an assembler's memory, we must prevent defects from occurring even when the operator fails or the assembler forgets.

Another type of defect that occurs frequently is a damaged tool defect. Either the tool itself may break and thus damage the part, or the machine may malfunction and damage the part. The easiest way to avoid the great majority of these difficulties is through a preventative maintenance program. However, this will reduce the number of occurrences of this problem, but will not ensure that it **never** happens. The way to prevent this from occurring is to install a device to check the soundness of the tool and

machine before each machining cycle. This can be accomplished by means of a touch off inspection (to check for a broken tool), or a system self-check by the machine. This will reduce the number of machine malfunctions and broken tool defects to zero if it is accompanied by routine maintenance.

6.5. Rough Framework for Designing Poka-Yoke Devices

There are two types of processes that one must distinguish: (I) already existing processes that have been run for a significant quantity of time and (II) new processes that are being developed. The main difference in implementing poka-yoke devices in already existing processes is that the type of defects that occur (and occur most frequently) are already known. In new processes the process designer must guess which defects are most likely to occur based on critical dimensions of the part, key parts in assembly, etc., and design poka-yoke devices to prevent these defects. The simple design framework is as follows:

I. For existing processes,

(1) Collect data on past defects,

(2) Calculate (for each defect) the product of: a) the probability of occurrence, b) the probability of detection, and c) the cost of the defect. Calculate the priority value of each defect from (reprinted from section 5.14):

$$PV(defect_i) = \frac{p(defect_i) \bullet C(defect_i)}{p(\text{det}ection(defect_i))}$$

(3) Assess which defects can be prevented by Poka-Yoke devices, and which defects will require capital investment (redesign of the part, or the fixture, capital intensive machine sensors, etc.).

From the previous section we can see that the defects which are the easiest to prevent using Poka-Yoke devices are the following:

- improper fixturing of parts
- defects caused by excessive travel of machine head (i.e. over drilling)
- in assembly, defects of forgetting to add certain parts (either nuts, washers, seals, bolts, or other small items, including pan stock parts)
- improper alignment of a part during assembly

II. For new processes,

(1) Analyze each part and note critical dimensions and characteristics.

(2) Analyze the process, including all machines and operations that will be performed to make the given part, and brainstorm the possible defects that might occur.

(3) Assess a likelihood of each defect occurring as well as the extent of the damage that it would cause (i.e. scrapping of the part, cost to rework the part, etc.).

(4) Assess, as in I above, which defects could be prevented by the cheapest methods. The key in this process is that there is no previous data, so it is unclear what defects will occur in a given process. The best that one can do is make educated guesses, implement the cheapest solutions to prevent the greatest number of defects and then implement more devices after the process begins making parts, if defects occur.

It is important that only low-capital poka-yoke devices are installed when a new process is being designed, since some defects may not occur and investing capital to prevent defects that would never have occurred is just wasting capital.

Chapter 7. Conclusions

We now define T_{all} as the length of time the customer is willing to wait for product after having paid for the order. We will call T_{all} the allowable lead time for post-payment production. For some products this may be close to zero. However, for many products, such as prepared food, cars, and all items that are sold through catalogues, T_{all} ranges from a few weeks to two years. For most products, the order lead time (OLT) using traditional manufacturing methods is much longer than T_{all}. As long as the OLT is longer than T_{all}, it is necessary for factories and sales rooms to carry an inventory of finished goods which the customer can obtain product within T_{all}. The factory must then produce to fill this inventory.

Chapter 2 illustrated the results of producing to fill inventory levels on production systems. Slight variations in demand from the customer created cycles of demand which propagated along the chain of suppliers. The orders to each link in the LDS were amplified higher and higher as one moved further from the customer up the chain, as the result of the MLT time lag of each link of the chain. As a result, the inventories of both raw materials and finished goods also escalated and then fluctuated wildly.

Chapter 3 showed how increases in inventory lead to ever increasing costs in the form of invested capital, damaged finished goods, scrapped product due to order cancellation or design changes, and a tremendously costly inventory control system of forklifts, warehouses, conveyors and computer systems. The vision of inventory as a "necessary evil" was transformed into a vision of inventory as merely evil.

If one can lower the manufacturing lead time and the order processing time to reduce the OLT below T_{all}, then one can produce to fill an order for product rather than to fill an inventory level. Chapter 2 showed that the only way to escape the cyclical nature of LDS, which most industries are composed of, was to reverse the flow of information and produce to a Takt time prescribed by customer demand. The goal of manufacturing system optimization is thus to reduce OLT beneath T_{all} and eliminate all defects.

There are several contributions to OLT: order processing time, part type production intervals, manufacturing lead time, and length of shipping intervals. The order processing time is consumed mostly by information processing and paperwork, and can be streamlined by improving information flow from sales to manufacturing. The part type production interval's contribution to OLT was illustrated in Chapter 4. A system which requires a long period of time to changeover between part types is often unable to produce product of a given type cost effectively in small batches and thus must finish the batch of the part type it is currently manufacturing before beginning the new part type. As a result, there is often a substantial delay between the time an order arrives for one part type and the time production on that part type begins.

The manufacturing lead time is made up of four components, processing, transportation, storage, and inspection. A Time Division Analysis (see Chapter 1) created from a process map (see Appendix 2) will show which activities are value added to the process and which must be eliminated. Processing is the only value-adding component of lead time, and even processing is filled with a large quantity of non-value added time and energy. The theoretical minimum manufacturing lead time is the sum of the value added processing steps. Chapter 5 showed that implementing cellular manufacturing eliminates

a large quantity of storage time (which makes up 95% of the lead time in aerospace manufacturing), and allows one to approach this theoretical minimum. Storage of parts between the first and last operations is a sign of a system imbalance which must be eliminated. Transportation of parts is also reduced or eliminated by adjusting layouts to minimize the distance that parts must travel to be processed.

One must remember, however, that it does little good to produce an entire order of parts in a very short period of time if the parts are defective. Chapter 6 showed not only how to improve the quality of products through implementation of Poka-Yoke devices, but also how to reduce or eliminate inspection altogether. Since a defective part cannot be used to fill an order, eliminating all defects is crucial to achieving the goal of OLT < T_{all}.

The final component of lead time is the shipping interval time, which includes the shipping time, plus the time the finished goods are stored before being shipped. This can be reduced by moving production to the location of the customer (as compared with moving manufacturing facilities to regions where the labor is cheap), and making shipping more frequent.

If we define inventory as raw materials or product that has been purchased by the factory but not paid for by the customer reducing the OLT to T_{all} will result in a complete elimination of inventory. The elimination of inventory not only avoids the difficulties discussed in Chapter 3, but also frees the factory from the cyclical behavior of linear distribution systems (chapter 2). One can see that a transition to Lean Manufacturing is a process whereby the OLT is constantly reduced through continuous improvement which eliminates the non-value added activities, and yields a manufacturing system which only

produces what the customer needs, when the customer needs it and in the quantity the customer demands, all in a period of time less than T_{all}, the allowable customer lead time.

Directions for Future Research

Although a plethora of literature exists on the principles of Lean Manufacturing, there is generally a lack of focus on the system level issues, as well as the integration of the three levels of production, system, cell and machine (or operation). At each of these levels, there is a need for more investigation to break long-held beliefs and myths of manufacturing that prevent the implementation of Lean Manufacturing.

In the area of modeling, an approach using control theory to understand the intricacies of Kanban-linked pull systems will serve to illuminate the effects of altering the Kanban container quantity, and the number of containers. In the area of cell design, there is a need to create a system with automated material handling (for large and heavy parts) which retains the volume flexibility of a manned cell. Finally, in the area of quality, an axiomatic design approach to eliminating defects through machine design represents the next step on the road to "zero defects".

Bibliography

Black, J.T., *The Design of the Factory with a Future*, McGraw Hill, 1991

Cochran, D.S., *The Design and Control of Manufacturing Systems*, Ph.D. Dissertation, Auburn, 1994

Burbidge, J.L., *Principles of Production Control*, London, Macdonald and Evans, 1962

Burbidge, J.L., *The introduction of group technology*, New York, Wiley, 1975

Cox, J. C., Blackstone, J.H., Spencer, M.S., editors, *APICS Dictionary 8th edition*, Falls Church, VA, APICS, 1995

Deleersnyder, J.-L., Hodgson, T.J., Muller, H., O'Grady, J., "Kanban controlled pull systems: an analytic approach," *Management Science*, vol. 35, no. 9, pp. 1079-1091, 1989

Drucker, P.F. *Managing the Future*, New York, NY, Dutton Press, 1992

Forrester, J.W., *Industrial Dynamics*, Cambridge, MA, The M.I.T. Press, 1965

Forrester, J.W., *Principles of Systems*, Cambridge, MA, Wright-Allen Press,1968

Helander, M., *A Guide to the Ergonomics of Manufacturing*, London, Taylor and Francis, 1995

Hirano, H. editor, *Poka-yoke Improving Product Quality by Preventing Defects*, Portland OR, Productivity Press, 1988

Monden, Y., *Toyota Production System*, Atlanta, GA, Industrial Engineering and Management Press, 1983

Nakajima, S., *Introduction to TPM Total Productive Maintenance*, Portland OR, Productivity Press, 1988

Ohno, T., *Toyota Production System- Beyond Large Scale Production*, Portland OR, Productivity Press, 1988

Sekine, *Single Piece Flow*, Cambridge, Productivity Press, 1990

Senge, P.M., Sterman, J.D., "Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future," *European Journal of Operational Research*, 59 pp.137-150, 1992

Shingo, S., *The Shingo Production Management System- Improving Process Functions*, Cambridge, MA, Productivity Press, 1992

Shingo, S., *A Study of the Toyota Production System*, Cambridge, MA, Productivity Press, 1989

Shingo, S., *Non-Stock Production*, Cambridge, MA, Productivity Press, 1988

Shingo, S., *Zero Quality Control: Source Inspection and the Poka-yoke System*, Cambridge, MA, Productivity Press, 1986

Shingo, S., *A Revolution in Manufacturing: The SMED System*, Cambridge, MA, Productivity Press, 1985

Smith, D. A., *Quick Die Exchange*, Dearborn, MI, S.M.E., 1991

Suh, N., *The Principles of Design*, New York, Oxford Press, 1990

Van de Vegt, J., *Feedback Control Systems*, Englewood Cliffs, NJ, Prentice Hall, 1994

Womack, J.P., Jones, D.T., Roos, D., *The Machine that Changed the World*, Harper Perennial, New York, NY, 1991

Cellular Manufacturing

Black, J.T., Chen, J.C., "Decoupler-improved Output of an Apparel Assembly Cell," *The Journal of Applied Manufacturing Systems*, Winter, pp. 47-58, 1994

Chan, F.T.S., Jayaprakash, B., "Design of automated cellular manufacturing systems with simulation modeling: a case study," *International Journal of Computer Applications in Technology*, vol. 8, no.1/2, pp.1-11, 1995

Cheng, C.-H., Madan, M.S., Motwanis, J., "Designing cellular manufacturing systems by a truncated tree search," *International Journal of Production Research*, vol. 34, no. 2, pp. 349-361, 1996

Mukhopadhyay, S.D., Sarkar, P., Panda, R.P., "Machine-component grouping in cellular manufacturing by multidimensional scaling," *International Journal of Production Research*, vol. 32, no. 2, pp. 457-477, 1994

Noaker, P.M. editor, "Strategic Cell Design," *Manufacturing Engineering*, March, pp. 81-84, 1993

Offodile, O.F., Mehrez, A., Grznar, J., "Cellular Manufacturing: A Taxonomic Review Framework," *Journal of Manufacturing Systems*, vol. 13/3 pp. 196-220

Pardue, M.D., "Use of Cooperating Expert Systems for Real-Time Fault Detection and Correction in a Manufacturing Cell," *Computers Electrical Engineering*, vol.19, no. 2, pp. 103-112, 1993

Price, W., Gravel, M., Nsadanda, A.L., "A review of optimization models of Kanban-based production systems," *European Journal of Operational Research*, 75, pp. 1-12, 1994

Singh, N., "Design of cellular manufacturing systems: An invited review," *European Journal of Operational Research*, 69, pp. 284-291, 1993

Group Technology

Chen, H., "Operator Scheduling Approaches in Group Technology Cells- Information Request Analysis," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, No. 3, pp. 438-452, 1995

Heragu, S. S., "Group Technology and Cellular Manufacturing," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, No. 2, pp. 203-214, 1994

Min, H., Dooyoung, S., "A group technology classification and coding system for value added purchasing," *Production and Inventory Management Journal*, 1st quarter, pp. 39-42, 1994

Snead, C.S., *Group Technology*-Foundation for Competitive Manufacturing, NY, Van Vostrand Reinhold, 1989

Flexible Manufacturing Systems

Angelva, J., Piltonen, P., "Real time data management in a flexible manufacturing system(FMS)," *Journal of Materials Processing Technology*, 52 pp. 76-82, 1995

Flynn, B.B., "The effects of setup time on output capacity in cellular manufacturing," *International Journal of Production Research*, vol. 25, no. 12, pp. 1761-1772, 1987

Shang, J., Toshiyuki, S., "A unified framework for the selection of a Flexible Manufacturing System," *European Journal of Operational Research*, 85, pp. 297-315, 1995

Vakharia, A.J., "Methods of Cell Formation in Group Technology: A Framework for Evaluation," *Journal of Operations Management*, vol. 6, no. 2, pp. 259-271, 1986
Wemmerlov, U., Hyer, N.L., "Research issues in cellular manufacturing," *International Journal of Production Research*, vol. 25, no. 3, pp. 413-431, 1987
Wemmerlov, U., Hyer, N.L., "Procedures for the Part Family/Machine Group Identification Problem in Cellular Manufacturing," *Journal of Operations Management, vol. 6, no. 2, pp. 125-145, 1986*

Ang., C.L., Willey, P.C.T., "A comparative study of the performance of pure and hybrid group technology manufacturing systems using computer simulation techniques," *International Journal of Production Research*, vol. 22, no. 2, pp. 193-233, 1984
Chen, H.-G., "Operator scheduling approaches in Group Technology cells- information request analysis," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 3, March, pp. 438-452, 1995
Dale, B.G., Russell, D., "Production Control Systems for Small Group Production," *Omega The International Journal of Management Science*, vol. II, no. 2, pp. 175-185, 1983
Huang, P.Y., Rees, L.P., Taylor, B.W., "A simulation analysis of the Japanese Just-In-Time technique (with Kanbans) for a multiline, multistage production system," *Decision Sciences*, vol. 14, pp. 327-343, 1983
Lakhmani, G., "Relationship between Group Technology and Material Requirements Planning," *AIIE Conference Proceedings*, pp. 483-486, 1984
Lee, L.C., "A study of system characteristics in a manufacturing cell," *International Journal of Production Research*, vol. 23, no. 6, pp. 1101-1114, 1985
Min, H., Shin, D., "Group Technology classification and coding system for value-added purchasing," *Production and Inventory Management Journal*, vol. 35, no. 1, 1st quarter, pp. 39-42, 1994
Monden, Y., "Notes and Communications on 'A simulation analysis of the Japanese Just-In-Time technique (with Kanbans) for a multiline, multistage production system'," *Decision Sciences*, vol. 15, pp. 445-447, 1985)
New, C., "M.R.P. & G.T., A new strategy for component production," *Production and Inventory Management*, 3rd quarter, pp. 50-62, 1977
Onyeagoro, E.A., "Group Technology cell design: a case study," *Production Planning and Control*, vol. 6, no. 4, Jul.-Aug., pp. 365-373, 1995
Perrego, T.A., Petersen, H.C., Hahn, W.F., "Perrego algorithm: a flexible machine-component grouping algorithm based on Group Technology techniques," *International Journal of Production Research*, vol. 33, no. 6, pp. 1709-1721, 1995Rabbi, M.F.,
Sato, N., Ignizio, J.P., Ham, I., "Group Technology and Material Requirements Planning: An Integrated Methodology for Production Control," *Annals of the CRP*, vol. 27, no. 1, pp. 471-473, 1978
Sihna, R.K., Hollier, R.H., "A review of production control problems in cellular manufacturing," *International Journal of Production Research*, vol. 22, no. 5, pp. 773-789, 1984

Spencer, M.S., "Scheduling Components for Group Technology lines (A New Application for MRP)," *Production and Inventory Management*, 4th quarter, pp. 42-49, 1980
Stecke, K.E., Solberg, J.J., "Loading and control policies for a flexible manufacturing system," *International Journal of Production Research*, vol. 19, no. 5, pp. 481-490, 1981
Suresh, N.C., "Optimizing intermittent production systems through Group Technology and an MRP system," *Production and Inventory Management*, 4th quarter, pp. 76-84, 1979
Vaithianathan, R., McRoberts, K.L., "On Scheduling in a GT Environment," *Journal of Manufacturing Systems*, vol. 1, no. 2, pp. 149-155, 1982
Wemmerlov, U., Hyer, N.L., "MRP/GT: A framework for production planning and control of cellular manufacturing," *Decision Sciences*, vol. 13, pp. 681-701, 1982

Appendix 1. Mini-Case Study:

Cellular Manufacturing at Briggs & Stratton

with Mike Krawczyk

Briggs and Stratton is the largest producer of engines for portable power. At their Milwaukee, Wisconsin plant, they house the entire process, from die casting, through final assembly. Their casting facility is the 2nd largest user of aluminum in the United States. They have been in business for over half a century, and have become the best known name for engines that power everything from portable generators to lawn mowers. Some of their engine designs have been relatively stable for the past 25 years. It is thus surprising to find that they would have reason to redesign their factory.

It was a drive for profitability that led Vice president Greg Socks to direct resources to implement cellular manufacturing at Briggs. After brief exposure to the benefits of cellular manufacturing, he endorsed a plan to start a pilot project co-headed by Mike Krawczyk to test the feasibility of cellular manufacturing in the machining area at Briggs and Stratton.

The pilot project was enormously successful, and after a few months, Krawczyk was instructed to implement cells for other parts of the engine. Krawczyk, now working under Don Klenk, is in charge of designing most of the new cells for machining at Briggs and Stratton, whether it is in their Milwaukee plant or in Briggs' southern plants . On his wall in his cubicle, among various Japanese manufacturing books was a sign that read, "If

it isn't adding value, it's adding waste". It was this philosophy that guides Mike through his designs and optimizations of manufacturing practices at Briggs.

Generally, Krawczyk begins the cell design process with a product which is to be produced at a rate to match future customer demand which is often determined by market research. In addition, he is given an estimate of the product life, which helps in determining the capital investment that is justified for the given product. For example, a product with short life cycle between design revisions will probably not be a good candidate for dedicated machinery which is inflexible to changes in part characteristics or geometry. From the forecasted customer demand rate Krawczyk calculates the Takt time which sets the production rate for the cell. The Takt time is calculated from:

$$\text{Takt Time} = \frac{\text{Total time available (= 27600 sec/shift)}}{\text{Total number of parts to produce / shift}}$$

Based on this rate, which has units of time (seconds), the next step is to analyze the process which will create the part, carefully dividing up each operation and estimating the time required. If any processing steps are longer than the takt time, then they will present capacity problems. When considering investments in new machinery or new technology, Krawczyk generally had to justify investments based on a 6 month or less pay-back period.

If there is excess capacity with the machines available, one should look at including other parts in the same cell. Each cell should produce products at the same rate as they will be needed for final assembly, so it is not beneficial to design a cell which

produces faster than final assembly can produce. However, in general Krawczyk tries to design cells to run at 85% capacity based on the production rate he receives from marketing. This allows for an increase in production should demand exceed market predictions, or as is often the case in portable generator cells, sales occur in large demand "spikes" after natural disasters. One of the largest benefits from cellular manufacturing is volume flexibility. The throughput of most cells can be increased or decreased simply by adding or removing operators. This makes cells ideal for startup conditions where initial demand for a new product is low and then rises as the product gains popularity and market share. One can design cells with enough capacity to produce at the higher production rate, but keep low operating costs in the startup period by operating them with only one operator, and a low production rate. This will also help to prevent overproduction of goods that cannot be sold.

Following capacity estimates, ergonomic issues must be considered. If the part is heavier than 10 lbs., or is too small to pick up without a special tool, then one must consider automating the material handling. In addition, when considering types of machinery, and floor layout one should consider that in normal cell operation, operators walk 3 or more miles a day. Thus machines with large "footprints" (the rectangular dimension of the machine on the floor that encompasses the machine and determines the distance an operator must walk to move a part to the next machine) should be avoided as much as possible. In addition, machines should have a standard work height of 38"+/-2" so that operators do not have to reach up, or bend over, as this will cause health problems which translate into real dollars in costs to the company in addition to the pain and aggravation which will have adverse affects on operators' motivation, and quality of

work. In addition, carpal syndrome has been recognized as a serious problem. Replacing push button switches with light sensor switches on most one cycle automatics can alleviate these difficulties.

These general issues must be addressed in all cell design. However, actual design of the cells from capacity estimates for each machine and operation, to layout issues require separate analysis. In section 5.14 a step by step design process was given which shows how to start with a final product, a production rate, a set of machines, and a given floor space and create a fully operational cell. Some of the issues addressed include machine and process optimization to increase capacity, balancing a cell, and installation of quality improvement devices.

Improvements in quality by shifting to cellular manufacturing are facilitated by single piece flow processing, and trackability of parts.. Before the change to cellular manufacturing, most machines produced parts which were then placed into bins and then transported to the next machine in large batches. Many parts were damaged due to handling, and others sometimes skipped operations due to the ambiguity in part flow paths. These difficulties combined with the normal problems of worn cutters, miss cast parts, and improper coolant flow during cutting resulted in 5 to 10 percent of all parts failing inspection after being machined. After the shift to cells, roughly .5 percent of parts are rejected, with the great majority of these being defective castings.

The number of consecutive defects that can be produced in cellular manufacturing is limited by the frequency of inspection and can be calculated from the following relation:

$$MCD = \frac{T_{insp}}{t_{takt}}$$

where MCD is the maximum number of consecutive defects, t_{takt} is the takt time, and T_{insp} is the period of time in between inspections. For example, at Briggs and Stratton model 19 engines are being produced at a rate of 879/day, and one part is inspected each hour (in most cells). It is thus possible to produce 114 consecutive defective parts before discovering the problem. In the past, it was not uncommon for entire bins of several thousand parts to be defective. Defective parts may be discovered if parts are inspected before being assembled in the engine, or they may be discovered if the engine fails test. However, as long as inspection is less than 100% of parts, it will be possible to install and sell an engine with defective parts. Briggs has failed to implement 100% inspections for several reasons. First, there is a belief that inspections add too much time to each part. The takt time for model 19 engines (and all parts) is presently 38 seconds, and the inspection of parts, even using "go-no go" gauges, requires 5 of more seconds to test each of the critical dimensions. The cells are currently not staffed to include this 5 second inspection in their operation sequence. In addition, 95% of defects are the result of poorly cast parts, which is causing current investigation in redesign of dies and machining fixtures.

Cellular manufacturing must be addressed from a system level, and detail must carry down to the machine design level. Failure to address system and machine level

issues will result in a manufacturing system that fails to receive the full benefits of cellular manufacturing. At Briggs and Stratton, a large part of the machining area has shifted to cellular manufacturing, and is linked to the production rate of final assembly through Krawczyk's design of cells to meet Takt time. Implementation of cells has resulted in a tremendous decrease in inventory within machining. Before cellular manufacturing was implemented, each machine had a bin of raw materials and a bin of finished parts which would hold hundreds of parts. Presently, the number of pieces of inventory within a cell is generally equal to the number of machines in that cell. However, the number each type of cast part waiting to be machined numbers in the thousands and the number of machined parts waiting to be delivered to the assembly line is often 1000 units (of each part), which is equivalent to more than one day's supply of finished goods. This is a result of failing to modify the links between casting and machining and between machining and assembly. There are many improvements that can be made. However, the shift to cellular manufacturing has enabled Briggs and Stratton to greatly improve the efficiency of production, achieving a Flow Efficiency of 39% in machining and 31% in assembly (Flow Efficiency is the ratio of the length of time value is being added to a part to the length of time it spends in the factory). Their past performance sets the stage for a move to Just In Time (JIT) Manufacturing, and a further decrease in inventory by an order of magnitude.

Appendix 2. Process Mapping

The key to understanding any manufacturing system is to look at the factory at two different levels, the macroscopic or system level, and the microscopic or operation level. The production system can be broken down as in Figure 1.

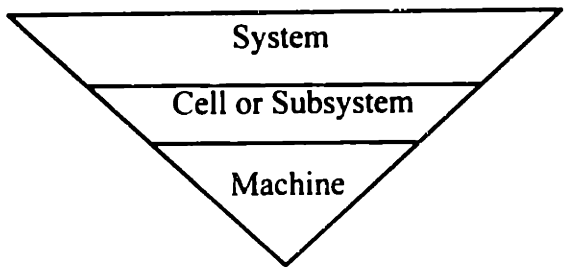
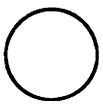


Figure 1. Breakdown of a Production System

Analyzing the plant from the system level will require a Time Division Analysis which divides all activities into four categories: processing, storage, transportation and inspection (see Table I). Of these, only processing is value added. Furthermore, not all of the processing time is value added. A microscopic study of the processing operations, will involve dissecting them into individual motions, whether they be motions of the machine, part, or human. Thus, in order to truly understand a company and understand what is not lean, we must take a macroscopic look at the plant, and eliminate all unnecessary storage, transportation and inspection, and then take a microscopic look at the processing of the material into finished products and remove any non-value added motions.

Table I. The Four Activities of Production

Activity	Symbol	Description
Processing		A physical change in the material or its quality (assembly or fabrication).
Inspection		Comparison of the material or part with an established standard.
Transportation		The movement of material or parts; a change in position.
Lot Delay		In batch production, while one piece is being processed (or inspected), the rest of the batch waits, either to be processed or to move onto the next machine
Storage Delay		The entire lot waits in storage while other lots are being processed.

Measurables

Key measurables may be defined to focus the researcher's energy towards gleaning the most important information. There are two types of measurables. The first are "absolute" measurables, such as defects, and setup time that can be compared to an absolute standard. Thus, we may say a setup should not require 8 hours, or a 50% first pass yield rate is not acceptable. We cannot allow relative measures to be used, because then a company may be satisfied with a 90% reduction in set-up time from 10 days to 1 day, when in reality, the set-up could be done in 1 hour. By introducing measurables as "absolute" we open the door to truly revolutionary ideas for improvements. The second

type are "guideline" measurables such as lead time, and floor space. These measurables are only important if they are referenced to a particular part or process. One cannot say that 10,000 square feet of floor space is too much or too little without knowing what manufacturing system that space houses. This division of measurables (see the summary for a listing of all the measurables) will facilitate finding ways to improve the system to reach the relative optimum levels in the guideline measurables, and the absolute optimum levels in the "absolute" measurables.

The measurables can also be categorized into the three levels of a production system (from Figure 2). Certain measurables are tied to the system level, some to the subsystem level, and others to the machine level. On a macroscopic system level, the measurables include cost, quality(usually some number of defects or defective parts per million, and rework time), lead time and flexibility. There are two different categories of defects that cost a company dollars and time. The first are defects made in the plant which can be corrected in the plant before the product reaches the consumer. They will be termed Internal Defect Costs (IDC). The second set of defects are those that are discovered by the customer. All warranty rework is included in this category which will be termed Post Processing Defect Costs (PPDC). There are also two different lead times which are important, namely Order Lead Time(OLT) and Manufacturing Lead Time(MLT). OLT includes product design and thus measures the time from first concept all the way through to finished product. MLT measures the time from production of the work order or product request through final delivery.

Other measurables that will be collected are: Floor space (for each product line, or group of products), setup time (which will be divided into external and internal setup

times), and Work In Process (in both units and dollars (approximate)). These measurables will help to explain the state of affairs in both quality and lead time. If one is looking for non-leanness or wastes, it will usually be found in the form of long lead times, or defective products. In a lean company, good products are produced when needed and in the form and quantity needed.

The microscopic issues of importance include ergonomics, tool paths, fixturing and changeovers (which are linked), and process variation. The microscopic and macroscopic levels are tightly integrated, and should not be approached separately. Often the system or plant layout will be built around individual operations, and process characteristics. If one attempts to separate the system from the individual operations there will inevitably be a loss of information.

Summary of Measurables

Absolute

- Internal Defect Costs- Costs incurred due to defective parts that are either scrapped or must be reworked.
- External Defect Costs- Costs incurred on parts that are returned by the customer and must be reworked or replaced. Includes all warranty rework.
- Escapes (internal and external)- number of defective units that reach the customer. If the customer is another plant of the same company, it is an internal escape. If the customer is outside the company it is an external escape.
- Backlog of ordered units- Quantity of units that have been ordered but not delivered.

- Changeover time- The quantity of time between production of two different types of parts. The quantity of time that a process is not producing parts because it is being reconfigured for the next product.
- On time delivery percentage- The percentage of all delivered parts that are delivered on or before the due date.

Guideline

- Direct and indirect labor hours per part
- Order Lead Time
- Manufacturing Lead Time
- Floor Space- The area, in square feet, that a given machine, process, cell, or product line covers.
- Number of changeovers (for a given process, per week or per day)- The frequency that a given machine or assembly line or station must stop production to change over to a new product.
- Units produced (per day, per month)- The number of units that are shipped per day or per month
- Work-In-Process (in units and dollars)- The quantity of units or dollar equivalent of all parts in the factory, including all raw materials, and all parts partially or completely processed, but not shipped.
- Inspection Stations- Number of stations where the part is inspected by an employee dedicated to inspections

- Manpower- The distribution of manpower at various stations of the process, including direct and indirect labor.

Analysis Approach

Upon entering a factory, the best way to understand how product is made from raw materials to shipping is to develop a process flow map for the product of interest. The process flow map is made by classifying all plant activities into the four groups listed above (processing, storage, transport, and inspection). Each of the four activities is given a symbol. The result is a process map with the Process flow on the Y-axis and the Operation flow on the X-axis. The distinction between a process and an operation is not one of time scale; the two have different subjects of study. A process is a flow of **product** from raw materials to finished parts. Operations are the actions of **man**, or **machine**, and what they do to the product. In our analysis, we will label the Y-axis System, rather than Process, which will allow us to include areas of the factory that are not a part of processing, such as marketing and design.

After completing this macroscopic process flow map, one must break the processing down further into individual motions, as shown in Figure 2. For example, analysis of a machining step would be accomplished by a time study of the operation at that process step, including examination of the tool path, set-ups, and fixturing. Dissecting individual operations also includes noting how many degrees of freedom the machine or part has in the operation, and how the parts are handled between machines or assembly steps. This reduction of the processing into basic kinematics allows one to see exactly what **is**, and what **is not** required to make the part (what adds value from the eyes

of the customer). It will also allow one to see how defects are made, and thus how they can be prevented.

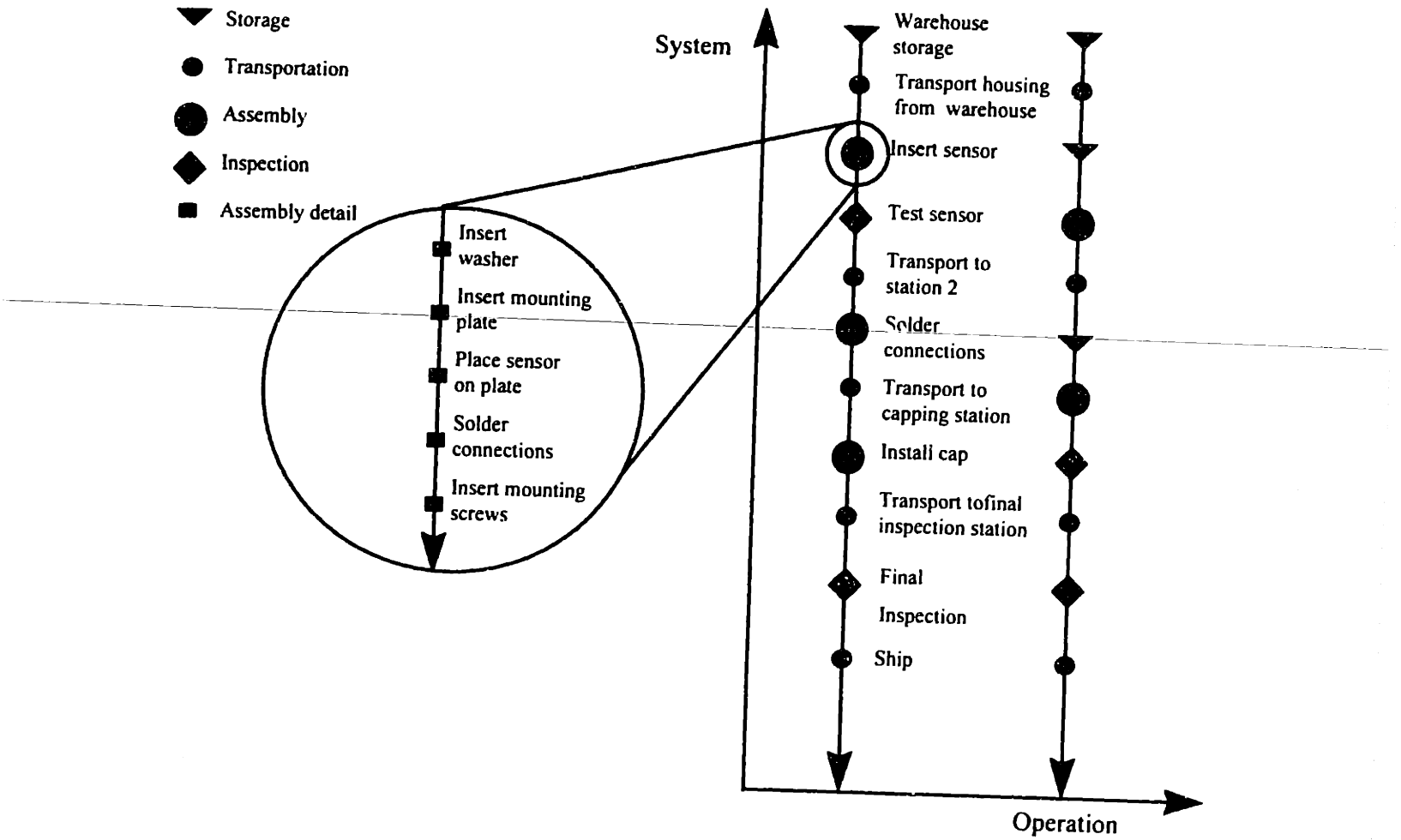


Figure 2. Process Map of an assembly process