TEACHING OPERATIONS MANAGEMENT FROM A SCIENCE OF MANUFACTURING*

MARK L. SPEARMAN AND WALLACE J. HOPP

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-2301, USA
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60208, USA

Recently, there has been much concern over the dumbing down of production and operations management (POM) courses in response to the repudiation of the theory-heavy operations research (OR) approach that became dominant during the 1970s. However, although almost everyone agrees that a new POM framework is needed, there is as yet little agreement on what it should be like. As a result, there is currently a huge variance among POM courses at different universities, ranging all the way from traditional OR courses to almost purely anecdotal case-oriented courses. Although academics have struggled with the search for an appropriate level of methodological rigor in POM courses, our customers (i.e., students and the firms that hire them) have been inundated by a blizzard of management buzzwords. Although many of these undoubtedly contain kernels of truth, the very nature of the buzzword approach is such that it provides little balanced guidance as to what methods work well in a given situation. In recognition of these disparities, POM researchers have begun trying to systematically describe the underlying behavior of production systems. The goal is to provide a framework that will help organize educational approaches and business practices in a consistent fashion. In this paper, we describe our attempt at the needed "science of manufacturing," which we call factory physics, and illustrate how it fits into a new paradigm of POM teaching.

1. Introduction

We are in a state of considerable controversy over the appropriate content of production and operations management (POM) courses in business curricula. The dominant approach of the 1970s of emphasizing the development and solution of detailed models of questionable realism has been rejected soundly by students and faculty alike. But as yet there is no universal agreement on a replacement approach. Some schools have opted for an almost completely case-based empirical approach with virtually no modeling content, whereas others have continued to stress certain models albeit with less emphasis on derivation and solution. The result has been a general trend toward less modeling rigor and a huge amount of variation in the content of POM courses across schools.

As the level of modeling in POM courses has eroded steadily, our customers (practitioners in industry who either hire or seek MBAs) have increasingly come to view the discipline as

* Received July 1996; revision received October 1997; accepted October 1997.
a blizzard of management buzzwords [e.g., materials requirements planning (MRP), manufacturing requirements planning (MRP II), enterprise requirements planning (ERP), just-in-time (JIT), computer-integrated manufacturing (CIM), flexible manufacturing systems (FMS), optimized production technology (OPT), total quality management (TQM), business process re-engineering (BPR), etc., see, e.g., Micklethwait and Woolridge 1996]. Although many of these contain kernels of truth, the very nature of the buzzword approach is to sell a single approach for all situations. Hence, there is little balanced perspective on what works well and when. This often has led to a “management by bandwagon” mentality with unfortunate results. Employees, battered by one “revolution” after another, are settling into a cynical attitude that “this too will pass.” But undaunted, many managers keep to the faith believing that someone, somewhere has a silver bullet that will solve all their operations problems. As a result, buzzword books and consultants prosper, but little real progress is made.

We believe that the failure of POM courses to set a strong and effective tone for manufacturing practice is not caused by an inappropriate use of models but rather to a limited use of models. In general, models can be used to either determine a good (or optimal) policy to design or operate a system or describe how a system behaves. We call models that suggest policies “prescriptive” models and models that describe behavior “descriptive” models.

In most POM curricula, most of the models presented are of the prescriptive variety. These typically are derived from basic mathematical assumptions. This differs from the view found in the sciences such as physics and chemistry where models are descriptive. Although relying heavily on mathematics, these models are not based on mathematical assumptions but, instead, make statements about how natural systems behave. They are not “derived” but are essentially independent conjectures. The overarching goal of physics is to explain the most phenomena with the fewest elementary conjectures. Of course, the models proposed in these fields are being tested constantly (and often rejected).

Practitioners in applied fields such as mechanical and chemical engineering use descriptive models for guidance in designing and controlling complex systems (such as chemical plants). However, these practitioners also have a firm grounding in corresponding descriptive scientific fields from which to approach their prescriptive engineering fields.

An example from one of these fields is illustrative. A civil engineer building a bridge is aware of several design strategies. These strategies are prescriptive in nature and are the result of both experience and models. For instance, over a long span, a suspension bridge is often a good option. Suspension bridges are supported by cables made of steel, which can accommodate enormous tensile stresses but are almost worthless when faced with compression stresses. In contrast, a shorter span is often better served with a reinforced concrete bridge, where the supporting members curve upward slightly producing compression stresses in the load bearing members. Concrete can support large compression stresses but does not work well under tension.

How do civil engineers know these things? Early in their education, before taking a course on building large structures, they take a set of courses called “engineering science.” One of these, “statics and dynamics,” covers compression and tension forces. Here one learns how an arch transmits load from its top to its base. Another early course describes the “strength of materials” such as steel and concrete. In our parlance, these are descriptive courses. Only after these basic concepts are understood, does the prospective engineer begin to take “design” or prescriptive courses.

One could argue that the models taught in POM courses serve a similar function to those taught in engineering science courses. Both sets of models are elementary and are used as building blocks for more complex systems. However, there is a fundamental difference. As Little (1992) points out, most of the mathematical models used in POM and industrial engineering (IE) are tautologies. Given a particular set of assumptions, the system behaves as follows. The emphasis is on proper derivation from the assumptions to the conclusions.
and not on whether the model is a realistic representation of an actual system. In essence, the truth of the model is self-contained. Little even demonstrates that Little's law is one such tautology and that, given this fact, there is no point in checking it with empirical data.

Many of the relationships taught in engineering science courses are fundamentally different in that they do make conjectures about the outside world. They invite the student to check particular statements against empirical evidence (and students do exactly this in laboratory sections). \( F = ma \) is one such conjecture. This "law" certainly is not a mathematical tautology; indeed it isn't even strictly true (it is only correct for "slow" speeds as compared with the speed of light). Nonetheless, it is useful and is at the heart of many complex engineering models. Also, \( F = ma \) and the several other Newtonian laws are remarkable for their simplicity. However, as any sophomore engineering student can attest, the field of statics and dynamics is anything but simple even though it is based solely on a small set of extremely simple statements about nature.

There does not appear to be an analogous basic "science" on which one builds the prescriptive field of operations management (or industrial engineering, for that matter). Fortunately, a number of researchers and teachers are beginning to address this gap. At the highest level, Schwarz (1996) offers a new and useful paradigm for teaching operations management that compares the practice of operations management with that of a manager of a financial portfolio. Three components, information, control, and buffers, must be managed together in order to have an effective operation. If one of these three is lacking (e.g., imperfect information regarding demand), it must be compensated by one or both of the other two (e.g., more control by assigning due dates to customers or more buffers by carrying safety stock).

At more of the operational level, Little (1992) asks the fundamental question, "Can we find 'laws' of manufacturing?" and conjectures a number of possible candidate laws. Askin and Standridge (1993) also put forth a set of general laws for manufacturing systems. Similarly, Buzzacott and Shanthikumar (1993) have put together a comprehensive set of queueing and other stochastic models of manufacturing systems, many of which use sophisticated approximations.

In more narrowly focused work, using both theoretical models and empirical studies, Karmarkar and others (Karmarkar 1987; Karmarkar, Kekre, Kekre, and Freeman 1985) have clarified the way we think about lot sizing. Suri and De Treville (1986) also used simple models to gain insights into how just-in-time systems work. Much of the work of Albin and Whitt (e.g., Albin 1984, Whitt 1983, 1984, 1993) has been devoted to better approximations of queueing models of manufacturing and service systems. Some of these approximations are based on queueing theory whereas others are used in a way akin to many approximations found in mechanical engineering—they work.

Paralleling the increased interest in descriptive modeling has been an increase in the number of case studies dealing with fundamental relations. For instance, Bourland and Suri (1993) provide a comprehensive case study that highlights basic relationships between performance measures in a manufacturing case setting. Their teaching notes illustrate a means of combining lectures to cover relationships, a case study to provide an example, and the use of software for empirical studies.

Unfortunately, the above examples are still the exception rather than the rule. It remains difficult to publish a paper with "insights from simple models." Teaching using software is more difficult than simply assigning problems or questions. "Laboratory" sections are still rare in POM classes. Finally, prescriptive models continue to outnumber descriptive ones in most POM courses.

The root cause for the lack of uniformity in POM classes is that there is no agreed on "science of manufacturing" or "science of logistics." But, though we still are far short of a true science, we believe that some important concepts are known. The above-cited works represent important pieces in this emerging field. We call our own attempt at structuring a
science of manufacturing "factory physics" (Hopp and Spearman 1996) and, despite its preliminary nature, find it extremely useful in structuring our POM teaching.

2. An Historical Perspective

In our factory physics-based course we start with a history of manufacturing. We do this for the following two reasons: (1) to understand how we got to where we are today and (2) to cover key classical models without overemphasizing them. One thread that such an historical analysis reveals is a propensity for too much prescription and not enough description.

For example, a classic model that has had enormous influence on POM is that of Wagner and Whitin (1958). Not only did it spawn a host of follow-on extension papers, but it led to the development of an entire subfield of heuristic approaches to lot sizing that work (almost) as well as the original Wagner-Whitin algorithm. Also, despite the almost 40 years that have passed since Wagner and Whitin’s original paper, the area continues to draw interest. The winner of the 1994 POMS Dissertation Competition was a clever procedure that sped up the original Wagner-Whitin algorithm from order $N^2$ to $N \log (N)$.

However, it is our belief that Wagner-Whitin represents a prescriptive model that was based on a faulty description of the actual system. Why? First, consider the issues addressed: trade-offs between setup costs and inventory carrying costs. Although ordering costs are quite real in an ordering environment (e.g., placing purchase orders), setup costs are less real in the manufacturing setting for which Wagner-Whitin was designed. As several researchers have pointed out [e.g., Karmarkar (1987)] the setup cost used in Wagner-Whitin (and EOQ) is often a surrogate for limited capacity. However, in the limited capacity case, the so-called Wagner-Whitin property (i.e., never produce in a period with entering inventory) usually is not present in an optimal solution (we will illustrate below). Further, although the Wagner-Whitin algorithm is designed to accommodate dynamic setup costs, no set of setup costs is guaranteed to result in even a feasible solution for a given set of demands. Again, this is caused by the Wagner-Whitin property. In spite of this, almost all of the subsequent heuristic approaches preserve the Wagner-Whitin property.\footnote{Baker, Dixon, Magazine, and Silver (1978) do solve the capacitated problem with inventory and setup costs. This is needed when there is an explicit cost for a setup (such as the destruction of parts or a fixture) but not for when the setup cost is being used as a surrogate for capacity. Unfortunately, the problem becomes NP-hard.}

Another research area where prescription has led description is the field of job-shop scheduling. A large percentage of the models in this area assume that all jobs are available at the start (i.e., no arrivals of new work), have deterministic processing times, due dates are exogenously set, and there are no sequence-dependent setups. The objective most frequently is to minimize makespan. When criticized as inappropriate, advocates retort that minimizing makespan is equivalent to maximizing utilization. However, we know from queueing theory (a descriptive field within operations research) that maximizing utilization can lead to big problems by increasing flow times (FT).

The research in the scheduling field has focused on various combinatorial and integer programming techniques to solve larger and larger problems. The 10 machine/10 job problem was first solved to optimality in 1989 (Carlier and Pinson 1989). Because most real-world settings involve considerably more than 10 jobs and 10 machines, it is unlikely that such developments will make much of an impact. Indeed, in their survey paper, Dudek, Panwalkar, and Smith (1992) (themselves, noted researchers in the field) conclude with

At this time, it appears that one research paper (that by Johnson) set a wave of research in motion that devoured scores of person-years of research time on an intractable problem of little practical consequence.
The moral of these stories is that developing prescriptive models without descriptive ones can lead to solving the wrong problem. Of course, the practitioners must work on the real problems whether academics do or not. We believe this is one reason for the development of simple, but not always effective, solution strategies. Material requirements planning and its subsequent incarnations (MRP II, BRP, ERP, etc.) is one such strategy. Although much simpler than approaches to job-shop scheduling, the model underlying MRP is too simple. Its major flaw is the assumption that lead times depend only on the part to be produced and not on shop conditions. In class, we illustrate this flaw by asking how long it would take to go from Chicago’s Loop to O’Hare International Airport, to which students always respond, “it depends on traffic.” The time it takes a job to traverse its routing also “depends on traffic.” Interestingly, as we will demonstrate below, it is not all that difficult to develop a simple approximate model that does depend on traffic.

Practitioners (at least consultants) also have offered some descriptive models. One set of these is the series of Socratic lessons/models discovered by readers of The Goal by Goldratt and Cox (1984). Unfortunately, The Goal is, for many practitioners, description without prescription. Our students typically report that after reading The Goal they are hungry for more details on implementation. Unfortunately, The Race (Goldratt and Fox 1986) and other subsequent books by the same author provide little help along these lines. Today’s practitioner is inundated by hyperbole from different purveyors of manufacturing philosophy all professing to have “the answer.” It appears that to get attention in the buzzword-dominated world of POM practice, ideas must be not only new but also extreme and dogmatic. For instance, “there are 14 points never to be violated” (Deming 1986), “quality is free” (Crosby 1979), “zero inventory is the goal” (Hall 1983), “zero defects is the goal,” “change requires the radical redesign of business processes” (Hammer and Champy 1993), “never use the t (tradeoff) word” (Schonberger 1986) and so on (emphasis ours).

This is indeed a sad state of affairs. Although there certainly has been some very good POM research into the development of a science of manufacturing, this research is not yet the basis for most teaching and practice. So, well-packaged dogma continues to dominate the classroom and the boardroom. But it needn’t be this way. We never hear such shrill exhortations in other applied fields. For one to insist that a particular grade of stainless steel must always be used regardless of the service would not be credible. Instead we find prescribed standards that are based on an empirical understanding of how the systems involved work. Also, the standards are applied with discretion. The term “engineering judgment” often is used to describe knowing when to adhere to and when to violate the standard. We believe that such standards will be common in POM once a science of manufacturing is established.

3. Descriptive Models First

Descriptive models need not be complicated. When teaching this subject we take as our focus a flow, usually represented by a manufacturing line (or routing). This is somewhere between the mechanical engineering focus of a “tool striking a piece of work” and the strategic business focus of a factory as a collection of cash flows. We begin with parameters that describe the line: the bottleneck rate and the “raw process time” (i.e., the time to process the part from any other delays such as queueing).

Often when we teach this subject we quiz the class (mostly people who have returned from at least 3 years of working in manufacturing) on what quantities determine whether a manufacturing system is operating well or not. Several always come to mind: throughput (TH) capacity, flow time (also known as cycle time, sojourn time, and throughput time), finished goods inventory, and work-in-process (WIP) inventory. We then ask the students to graph flow time as a function of WIP for a “perfect” line (i.e., one with no variability)
with a given bottleneck rate and raw process time. The variety of different graphs is astonishing. Although most are increasing some have extrema (both minimum and maximum). The point is made that although we all agree that these are important quantities, most of us do not have a clue as to how they are related.

We then begin to explore this relation using what we call the "penny fab," which consists of four identical processes in series, each requiring exactly 2 hours. On an overhead transparency, the penny fab machines are exactly the size of a copper penny (see Figure 1).

![Figure 1. Penny Fabs with WIP = 2 and 3.](image-url)
We use the penny fab to explore the relations between WIP, TH, FT, capacity (r_b), and raw process time (T_0). To do this, we fix the number of pennies in the system by starting another as soon as one is finished. We "simulate" the system on the overhead with one, two, three, four, five, and 10 pennies, all the while recording WIP, flow time, and throughput amounts. The students plot both flow time as a function of WIP and throughput as a function of WIP. We then point out the relation

$$\text{TH} = \frac{\text{WIP}}{\text{FT}}$$

identify it as Little's Law (tautology), and discuss that it holds in general (i.e., even with variability and randomness present).

The plots also reveal

$$\text{flow time} = \max \{ T_0, \frac{\text{WIP}}{r_b} \}$$

and

$$\text{throughput} = \min \{ \frac{\text{WIP}}{T_0, r_b} \}.$$ We identify these as the "best-case" flow time and throughput.

We then consider a case that illustrates the other performance extreme. In this case, when there are n jobs in the system, the first job takes 2n hours to complete whereas the others require zero time. This preserves the average of 2 hours while maximizing the variance of the process time. It also results in the "worst-case" performance in that throughput is minimized while flow times are maximized for the given average process times. It is interesting to note that this case is also deterministic. Students typically believe this example is contrived and there are no real-world examples of such behavior. Such comments provide a nice entree into a discussion of move batch sizes. It is easy to see that if jobs are moved n at a time, the worst-case performance results.

Finally, we discuss what we call the "maximum randomness" or "practical worst" case. To help understand this case, we observe that in the best case jobs appear to repel each other, resulting in minimum queueing, whereas in the worst case they attract one another, giving rise to maximum queues. In this light, we define the notion of a "system state," that is the number of jobs and the remaining process times at each station, and note that the best and worst cases involved relatively few of the possible states. In the maximum randomness case, all states are equally likely. (In a more technical class, this leads to a discussion of the "memorylessness" property and the exponential distribution.) We observe that the maximum randomness case (practical worst case) corresponds to a line composed of identical single machines (single server queues with the same rate) having exponential process times. Using symmetry arguments alone, we derive an expression for the flow time in the system, viz.

$$\text{flow time} = T_0 + \frac{\text{WIP} - 1}{r_b}$$

Although all of these models are tautological, they provide a framework for bounding the performance of real systems. We then observe some plots of actual systems (see Figure 2) and note that in spite of the added complexities, the same basic relation between flow times and WIP holds.

Figure 3 shows a simple approximation to this behavior. Our approximation relies on two parameters, a practical raw process time T_P and a practical bottleneck rate r_P. Both of these parameters are somewhat less efficient than their best-case counterparts. We call the resulting descriptive model that relates flow time to WIP the "conveyor model" because it corresponds to the behavior of a physical conveyor. The time to go down a
conveyor remains constant until the conveyor is full and there is a queue in front of it. Similarly, the time through a production line is fairly constant until it becomes congested, that is, until the WIP level becomes large. Once full, the time to go down a conveyor is
given by dividing the production rate into the amount of work there is at the conveyor. We find this simple descriptive model extremely useful in the development of useful prescriptive models.

In summary, we find that good descriptive models lead to better insight into the system we are trying to design or control. In the next section we see that a good descriptive model can lead to a better prescriptive model.

4. Building Prescriptive Models

The conveyor model can be used to solve the capacitated lot-sizing problem described earlier. Suppose our line has a practical capacity of 100 units per day and a practical raw process time of 3 days. Currently, there are 50 units in finished goods inventory, 95 units that have been in the line for 3 days (i.e., these will start coming out immediately), 95 units that have been in for 2 days, and 100 that have been in 1 day. Because less than 100 units were started 2 and 3 days ago, output is limited by available WIP. Thus, the maximum output from this point forward is 50 (immediately from finished goods inventory or FGI), 95 today, 95 tomorrow, and 100 from that point on. Demands for the next 10 days are as follows: 100, 120, 100, 0, 200, 0, 200, 120, 0, and 80. The 340 units in finished goods and currently in WIP cover demand for periods 1, 2, and 3 as well as 20 units of the 200 due in period 4. Thus, the netted demand is: 0, 0, 0, 0, 180, 0, 200, 120, 0, and 80. If we offset these by 3 days to find out what the “starts” should be, we obtain starting demands of 0, 180, 0, 200, 120, 0, and 80. Our task at hand is to find a start schedule that minimizes inventory and is capacity feasible.

First, we consider the Wagner-Whitin algorithm and vary the ratio of the setup cost and the carrying cost in hope of arriving at a feasible schedule. The schedule becomes feasible only when the setup cost is so high that all production is started in the first period. To do this, the cost ratio must be at least 1,200. If we project forward using finite capacity,
this results in 340 part-periods of inventory being carried. If we reduce the ratio to 1,199, the schedule becomes infeasible with a shortage of 120 units in period 9. Further reductions of the ratio only makes the infeasibility worse.

An extremely simple algorithm yields a capacity feasible schedule that minimizes inventory [see Tardif and Spearman (1995) for details]. Let

\[ D_t = \text{demand in period } t, \]
\[ X_t = \text{production in period } t, \text{ decision variable}, \]
\[ I_t = \text{ending inventory for period } t, \]
\[ C_t = \text{capacity available in period } t. \]

The algorithm works backward from period 10. We set the desired ending inventory for period 10 to zero, \( I_{10} = 0 \). Then, the production quantity for period \( t \) is given by

\[ X_t = \min \{ C_t, D_t + I_t \} \]

The ending inventory for period \( t - 1 \) is then set to

\[ I_{t-1} = I_t + D_t - X_t \]

If the ending inventory for period 0 is greater than zero, then demands cannot be met within the current capacity.

Applying the algorithm to the given netted demands yields a start schedule of 100, 100, 100, 100, 100, 0, and 80 with a total of 260 part-periods of inventory carried. This is 24% less than the "optimal" Wagner-Whitin algorithm. The algorithm can be extended easily to accommodate multistage production systems. The "conveyor model" also can be applied to due date quoting and throughput tracking.

Students find this model easy to understand and implement (using a spreadsheet). Assignments can range from simple problems such as the one above to extensive case studies involving multiple plant locations. Things become more difficult (and nonoptimal) when multiple products with WIP in the line are considered. However, using some of the predecessor and successor features found in many spreadsheet packages students can develop good feasible schedules by hand.

To summarize, a good descriptive model will characterize the actual system without overwhelming the analyst in details. Such models can be combined with algorithms to provide effective prescriptive models.

5. Structure of the Course

Our preferred structure for a POM course based on factory physics is a two-quarter or one-semester sequence in which we cover manufacturing operations management along with certain quantitative methods (linear programming, simulation, and regression). The course (and the book) is divided into the following three sections: I. The Lessons of History, II. Factory Physics, and III. Principles in Practice.

In Part I we cover a history of American manufacturing along with "classic" approaches such as EOQ and \((Q, r)\) inventory policies, lot sizing (including Wagner-Whitin), MRP and MRP II, and just-in-time. Each of these techniques is presented from an historical perspective. In other words, we do not judge the technique as good or bad, only that it has been done. The last chapter in Part I is titled "What Went Wrong" in which we critique the approaches and explain the need for better descriptive models.

Part II covers the basic descriptive models and states 20 "laws" of factory physics. In addition to the penny fabs, we make use of the Kingman approximation of the waiting time in a GI/G/1 queue and some of the techniques used in Whitt's queueing network analyzer (Whitt 1983) to analyze networks of workstations. These approximations are useful to illustrate the effect of how increasing variability increases congestion and how variability propagates in manufacturing systems.
We combine the relation between flow time and utilization from queueing theory with the conveyor model to show that pull systems (such as kanban) appear to be more effective than push systems (e.g., MRP) because they limit WIP (not because they "pull"). Given this, we offer a simpler alternative to kanban known as constant WIP (CONWIP) in which, instead of limiting the amount of WIP at each workstation, work is limited over the entire routing. This solves several implementation problems facing kanban such as being able to accommodate changing product mixes and short runs of small lots. Other topics include the following:

- The impact of random outages on "effective" process times
- The effect of lot sizing on waiting times [a la Karmarkar (1987)]
- The relationship of move batch sizes to flow time
- The need to buffer variability with either longer lead times, more inventory, and/or extra capacity
- The human element in operations management

Part III then applies the descriptive models of Part II to develop prescriptive models. We start with a discussion of "total quality manufacturing" whose theme is that one cannot have good quality without good logistics (i.e., low WIP levels and short flow times) and vice versa.

We then describe a hierarchical production control framework that accommodates pull systems. This framework provides the basic structure for the remainder of the book. Separate chapters are devoted to portions of this hierarchy: aggregate production planning, production tracking, forecasting, shop floor control, production scheduling, advanced inventory management, and capacity management.

6. Making It Relevant

Making POM and/or production control relevant to a group of students can be difficult. One way is to analyze current technology or trends using models from the course. Usually, a number of the students have significant work experience in a manufacturing setting (this is true even for undergraduate classes through co-op programs). Many times these students have been involved in (or, at least, have heard of) various manufacturing trends. Some of these experiences involve the installation of mature systems (e.g., an ERP system) whereas others are more of a "revolutionary" and buzzword nature. One way to bring in the real world is to assign a case study along with a reading from a current "popular" management book. The situation presented in the case can then be analyzed using both factory physics and the buzzword approach.

In the past we showed how much of the philosophy of TQM is devoted to variability reduction and, using simple models, why this is effective (and also when it is not). More recently, a trend known as "lean manufacturing" has come on the manufacturing scene. Although some of the "lean principles" offered in Lean Thinking by Womack and Jones (1996) are quite useful (e.g., their discussion of batch and queue versus continuous flow), others are more muddled. Part of becoming lean is to eliminate all waste or muda defined as "any activity which consumes resources but creates no value." In spite of the appeal of this simple definition, its practical use often conflicts with many of the "flow" principles.

Womack and Jones point out that the first step to "making value flow" is to "... focus on the actual object ... and never let it out of sight from beginning to completion." This is quite true and is one motivation for cellular manufacturing. The machines are arranged in a way so that there is no time wasted in queueing and in moving from place to place. Of course, if there are different products made in the cell, some requiring different processes than others, machines for all of the processes must be part of the cell all of the time (otherwise we would have muda by adding to and removing machines from the cell).
However, in this case we also have the muda of wasted capacity (i.e., machines dedicated to the cell that create no value for some products). In a situation with a wide variety of products this can be prohibitive. Cellular manufacturing typically is not used in job-shop situations, process-oriented manufacturing is. In a process orientation machines are grouped by process instead of by product. The result is less wasted capacity muda but more long flow time and high inventory muda. A way to reduce inventory and queue time muda is to move one part at a time between the processes. Unfortunately, this results in the muda of movement that creates no value and the larger muda of paying workers to only move product.

The problem with this analysis is not the need to eliminate muda or waste; it is the temptation to overly simplify a complex situation. Queueing is not an "activity that consumes resources." The product just sits there. If we define waste as anything that does not contribute the profitability of the firm, then the admonition to eliminate waste is true but empty.

Now consider the situation using factory physics. Unfortunately, the strength of factory physics also is its drawback—it offers no easy solutions. Instead of focusing on muda, factory physics makes much of "variability" (which differs from "randomness") and the need for buffers. Some variability is like muda in that it is caused by things that should be eliminated such as rework and scrap loss. Other variability is unavoidable such as changing customer demands and evolving process technology. For example, in the semiconductor industry, if one is to offer state-of-the-art technology, then yields, throughput, etc. will be lower than in a more mature industry. At any rate, a factory physics law states that in order to have high throughput in the presence of variability there will be some sort of buffer. There are only three kinds of buffers: inventory, lead time, and capacity. If you cannot eliminate the variability and want low inventory and short lead times, you will have wasted capacity. This is exactly what happens in cellular manufacturing systems with high product variety.

Combining our historical overview, we then show how many of the examples found in the lean philosophy are from Toyota, circa 1983. During that time, Toyota eliminated much variability by establishing a production schedule for a limited product mix that was constant for months at a time. In such a situation, the only variability is found in the process itself. Toyota spent decades reducing setups and eliminating scrap and rework. Once the variability was reduced, buffers could be eliminated. Indeed, Toyota, in the mid-80s, had very small inventories and high utilization of its resources. It was very lean.

However, today, the situation is different. In a recent presentation by the Nomura Research Institute (Fujino 1997), it became clear that Japanese manufacturing firms are having difficulty competing in a marketplace with highly customized products having brief life cycles with short lead times. In the Q&A that followed, it also became clear that the "lean manufacturing" principles exhibited by the Toyota Production System (TPS) are no longer in use even by Toyota.

The negative impact of overly simplified approaches such as lean manufacturing has been significant. At a recent visit to an in-house conference at a large aluminum company, we learned that not only had the company embarked on an implementation of lean manufacturing but would do it according to the TPS model. When an engineer pointed out that it is physically impossible to change over a casting operation in less than 10 minutes (as Toyota had done), he was dismissed as a nay sayer and "not with the program." Later another engineer remarked that the use of 2 hours between 10-hour shifts to provide an opportunity to "catch up" (i.e., a 20% capacity buffer) in the TPS would not be financially feasible in an industry with so large a capital base. This too was dismissed as something that Toyota did and would not have to be done in the present case (although how it could be avoided was not discussed).

The problem with buzzword programs is that there is no way to challenge such statements. They tend to be backed up with remarkable success stories but no real analytical
basis. This is why a recognized science of manufacturing is so important. In the case of the aluminum company, variability, caused by long setups and process variations, is a fact of life. Of course, some of this variability can and should be eliminated. However, the only way to become more responsive to customers will be to break down the operation into flows and analyze each one (exactly what one plant has done in spite of the corporate guidelines). In some cases, capacity needs to be added. In others, inventory needs to be added whereas in others, it needs to be reduced.

These cases drive home an important point: there are no simple solutions. Good decisions are the result of sound analysis and well-developed intuition. The purpose of teaching factory physics is to provide the student with a means of performing sound analysis and to build his or her intuition for approaching real problems.

7. Conclusion

Our two-course sequence was designed to be part of the core curriculum in the Master of Management in Manufacturing (MMM) program offered jointly by the McCormick School of Engineering and the Kellogg Graduate School of Management at Northwestern University. The reaction on the part of the MMM students has been extremely positive. Many students considered this course as their most useful course in the MMM curriculum—a remarkable feat for a POM course in a business school.

In recent years, there has been much "model bashing" in OM, OR, and IE. We believe this is misguided. The problem is not with models but with the premature use of prescriptive models before good descriptive ones have been developed. The paradigm of using descriptive models to characterize manufacturing systems before teaching prescriptive models appears to give good results. Students have better intuition of why different techniques work and when they might be expected to work. Overall, the students are more confident and better prepared to accept the challenges of a career in production and operations management.

References


BOURLAND, K. AND R. SURI (1993), Spartan Industries, Case Study and Teaching Notes, Amos Tuck School, Hanover, NH: Dartmouth College.


FUJINO, N. (1997), Agility Solutions for Global Supply Chain Management, Nomura Research Institute, Ltd., Presentation made at Georgia Institute of Technology, Atlanta, GA.


