IMPROVING SUPPLY CHAIN PERFORMANCE
BY USING ORDER FULFILLMENT METRICS

In the past, companies hoping to impress customers on delivery responsiveness did so by allowing inventories to bloat. Today the costs of holding inventory, especially in high-technology markets, prohibit that behavior. Companies must now provide good service while maintaining low inventories. Thus, managers must carefully measure and manage their two conflicting objectives: service and inventory. The challenge is to improve customer delivery service and reduce inventories simultaneously. Such results can be achieved through thoughtful supply chain management.

by M. Eric Johnson and Tom Davis

The more things stay the same, the more they change. This variation on the old saying describes the current dynamic in the computer industry, where rapid change is the norm. The most successful players adapt quickly when new rules appear; for example, when major innovators like Dell and Gateway made the direct purchase of personal computers commonplace. Today you can order a complete computer system with custom options as easily as ordering a sweater from L.L. Bean. The computer is assembled to order at the factory and delivered within a few days. In some cases, the final assembly and configuration may take place in the distribution channel itself, squeezing even more time out of the order fulfillment process.

Changing expectations—especially expectations about customer service—are not unique to the computer industry. Manufacturers in nearly every industry are increasingly aware of the need to satisfy expectations by delivering products to customers in a timely manner. Unfortunately, many highly acclaimed new products have been successfully introduced to the marketplace only to suffer delivery delays. High profile examples—from Beanie Babies to cellular phones—abound in nearly every industry. In these and similar instances, delivery problems erode customer satisfaction and provide footholds for competitors to enter the market.

At Hewlett-Packard, this change in customer expectations has been unsettling. Traditionally an electronic instrument company, HP had succeeded for years in a build-to-order environment where customers ordered equipment through a field representative and waited patiently for delivery—and the privilege of owning an HP product. Not so today! To the dismay of senior management, late deliveries and incomplete orders annoyed customers and undermined the company’s brand equity in each of its markets, from multi-user workstations to desktop printers to gas chromatographs.

Faced with these problems, HP intensified efforts to improve customer satisfaction by understanding and improving supply chain performance. With increasing concern for the needs and perceptions of customers, the company has focused particular attention on the key role played by order fulfillment. Noting that, “HP believes that poor order fulfillment costs it a million dollars a day,” CEO Lew Platt set a five-year goal for the company to achieve a tenfold improvement in its order fulfillment process. Famous for achieving similar goals in product quality, HP has mobilized to meet this challenge. But, challenging the company to improve is one thing; knowing whether it has improved is another.

* * *

M. Eric Johnson, Ph.D., is a faculty member at the Owen Graduate School of Management of Vanderbilt University in Nashville, Tennessee. Associate editor for Operations Research and Production and Operations Management, he is researching supply chain logistics through grants from the National Science Foundation, PepsiCo, and the Hewlett-Packard Company. Tom Davis is responsible for supply chain planning systems at Hewlett-Packard’s Home Products Division in Cupertino, California. His work at HP includes developing and applying new approaches to supply chain modeling and metrics, forecasting, and production and capacity planning.
THE ROLE OF ORDER FULFILLMENT

Taken at face value, order fulfillment sounds like a simple matter—just filling orders. At HP and other leading companies, however, the expression has taken on a meaning that reflects its importance to customers. Order fulfillment is not merely a routine task performed by a material handler, shipping clerk, or cashier; more importantly, it is the desired result of many complex processes undertaken by the entire business.

To the final customer, order fulfillment means many things. It goes beyond getting the right product at the right time—though, of course, that is important. The concept also includes smoothly executed order processing, readily available order status information, and prompt and courteous support after the sale. And, naturally, it means receiving desirable products. Good order fulfillment—an expression increasingly synonymous with complete customer satisfaction—arises from the successful interaction of development, manufacturing, and business processes to deliver what customers really want (see Exhibit 1). In fact, any activity that affects customer satisfaction can be linked to order fulfillment.

Order fulfillment considerations are just as important within the supply chain as they are at the interface between the company and its external customers. A deeper appreciation of order fulfillment arises when the complex position of the internal customer, whose performance can easily affect the efficiency of the entire supply chain, is considered. Exhibit 2 illustrates how an intermediate node plays the role of customer and supplier at once. This reveals that there is an order fulfillment relationship between adjacent nodes in a supply chain. It is evident, however, that the roles of supplier and customer can easily be confused, especially in a complex supply chain. While the “trade customer” or “end user” is generally viewed as the most important, the strength of that final link in the chain depends in large part on what happens upstream. Orders placed with any upstream supplier should arrive in keeping with the expectations of the internal customer immediately downstream. And all upstream activities should be directed toward satisfying the ultimate customer.

Although such aspects of order fulfillment as accurate order configuration and prompt response to customer inquiries should not be neglected, mastery of the physical processes—the path of processes by which the product physically moves through the supply chain to the final customer—is central to order fulfillment excellence. Order fulfillment metrics can be used to reveal the strengths and weaknesses of the link between adjacent nodes in the supply chain. At these critical junctures the proper application of the right metrics can lead to improved supply chain performance—satisfied customers along with improved financial performance.

WHY METRICS?

Each of us performs in accordance with the objectives used in measuring our performance. In sales we must meet quotas. In production we must match the plan. We must submit travel expense reports on time or suffer the consequences. Much of our time on the job is spent responding to...
a steady drumbeat paving the way we go about our work.

And yet, everyone has encountered pointless metrics. We’ve all seen precious energy wasted collecting data that no one ever uses (and puzzle over who could possibly be behind such efforts). These experiences sour us to the idea of yet another directive from management to track this or that. So why harp on metrics now?

Managers know that relevant metrics and performance objectives are fundamental to a successful business. To reconcile fruitless experiences with sound management practice, managers must rise to these challenges:

- Measure only the right things.
- Avoid meaningless efforts.
- Use the results productively.

At Hewlett-Packard, attention to order fulfillment metrics had flagged before Lew Platt’s initiative. The company was measuring the wrong things. In many instances, it was ignoring the data that were collected or, worse still, using the information inappropriately. This point was made very clear during an interorganizational supply chain initiative conducted with Best Buy (a large electronics retailer) and Texas Instruments (an upstream supplier of microchips to HP). After examining the entire supply chain for a set of HP’s popular inkjet printers, the company found that while HP’s internal distribution metrics appeared good, the resulting delivery service to Best Buy was less than stellar. The key issue was metrics: HP distribution was carefully measuring its internal ability to process customer orders, but never compared shipment performance to customer delivery requests. HP realized that the only way to take control of the order fulfillment process was to revisit the management fundamentals of measurement and control. There are a number of reasons for this.

**Meeting Customer Expectations**—As Anil Kumar and Graham Sharman pointed out in an article in the Winter 1992 issue of *Sloan Management Review*, you cannot satisfy a customer without delivering the product. In the personal computer market, delivery performance ranks second only to product attributes in importance to customers. This reality pervades most industries, and a number of other studies support this result. It is not just speedy delivery that customers want, but reliability as well and reliability often ranks ahead of speed when achieving order fulfillment success. This should come as no surprise. Federal Express would not have made its fortune without first convincing customers that overnight delivery really meant *overnight*.

A recent visit to a large HP customer brought home the importance of reliability. The customer had been buying from HP for one key reason—reliability—but other vendors promised five-week delivery, luring the customer away from HP with its six-week delivery. After a short time, however, with the computers’ delivery times varying from four to six weeks, the customer returned to HP. Although the other vendors were faster on average, the variability in their delivery performance wreaked havoc. As the customer’s manufacturing manager pointed out, “We never knew when to expect deliveries. Materials planning became a real chore, and we had to inventory the parts, driving up our material handling costs.

Another HP customer recounted a similar experience, concluding that “fast is nice, but customers expect more.” Like quality in the 1980s, *reliable* delivery has become a strategic weapon—and in some markets, a requirement. Customers increasingly rate their vendors on delivery, so companies must introduce and manage processes to make sure that delivery is reliable. Once again, effective management requires measurement.

**Improving Supply Chain Capability**—A customer’s delivery experience depends largely on the performance of the other players in the product delivery system. The physi-
cal processes in the supplier’s facility and the customer’s intertwine to determine downstream performance. Thus, reliable order delivery depends on competent supply chain management. Increasingly, managers appreciate this fact. Yet when Hau L. Lee and Corey Billington identified 14 pitfalls in managing supply chain inventory, no supply chain metrics topped the list. Instituting appropriate metrics provides the foundation for sophisticated supply chain modeling efforts, but without basic performance data, further work cannot succeed. Fortunately, the mere act of introducing appropriate metrics often leads to significant improvement in performance. Exposing problems is always the first step in creating change.

**Enhancing Asset Performance**—Inventory can be a large, unproductive asset. Making sure the company has the right amount of inventory—and no more than that—is crucial for productive asset utilization. The same is true for capacity. Idle capacity ties up cash that might otherwise be applied to more useful projects, like new product development or an ad campaign. The gurus of just in time have led many managers to believe that reducing inventory, at any cost, is certain to improve production efficiency and asset performance. But this is not true. Inventory is not necessarily evil. In fact, inventory plays a vital role in providing reliable delivery. Like capacity decisions, inventory decisions must be made through careful consideration of the alternatives. The important questions are: Why are we holding inventory? Are we holding the right amount? What problems in the supply chain is excess inventory compensating for? These same questions apply to capacity, as well; the two are intricately related.

Increased focus on financial performance, particularly return on assets, highlights again the need for appropriate performance measures in the order fulfillment arena. Without the right measure of service and a history of service performance, how can managers possibly make credible decisions about cutting (or increasing) inventory, adding (or idling) capacity?

**Instilling Workforce Cooperation**—When the beat changes, the workforce changes pace accordingly. Metrics are important in motivating the workforce, and the right metrics can lead to desirable employee behavior. But even seemingly good metrics can motivate undesirable behavior. Salespeople sandbag deals to ensure another good month. Distribution hides inventory in the field, fearful of coming up short. Production rushes to ship jobs early to produce good end-of-quarter results.

One production manager described how he would reduce work-in-process inventory (WIP) at the end of each month by slowing the release of work to the floor. After his monthly inventory audit, shop floor WIP would explode as delayed jobs were dumped on the floor. Jobs were completed late, productivity was lost, and WIP fluctuated wildly because of a directive to reduce shop floor inventory, a performance requirement reported on the production manager’s evaluation, influencing promotion and compensation. In a similar vein, product designers have been known to shave a few pennies from their product’s material cost, and in so doing add a dollar to the unit cost of delivering the product to the customer. As one HP manufacturing manager bluntly put it, “The way you’re measuring us isn’t consistent with your new objectives. Start grading us on the cost of the inventory we hold, and we’ll hold less of it. It’s as simple as that!”

**Enriching the Stakeholder**—There are strong connections among the metrics covered above. Each of them is compelling on its own and serves the stakeholders in the operation, from suppliers to employees to shareholders. A solid ensemble of metrics allows management to weigh the sometimes competing interests of shareholders and customers, and stimulates strategic focus on the trade-off between service and its cost.

**ESTABLISHING PERFORMANCE OBJECTIVES**

Successfully managing the order fulfillment process usually requires several metrics. For example, managers must track things like order-to-delivery cycle time, order defects, and defects on arrival (wrong item received). In addition to numerous tactical measures, they need a general measure of performance—something that quickly indicates the overall performance of the supply chain. A qualitative way to represent performance appears in Exhibit 3. By combining two metrics—one for service and one for inventory—managers can satisfy two constituencies:

- **Customers**—The service measure (reliable delivery to a customer’s desired delivery date, for example) represents the customer’s view. Better service means happier customers.
- **Business Analysts**—Quantifying the inventory investment, from raw materials to finished goods, serves the needs of the business analysts. All else being equal, it is better to hold less inventory in the system.

This view of service and inventory investment offers a ready indication of whether an organization is making effective use of assets. If it is performing in quadrant A, it suffers the double curse of too much inventory and poor customer service. Everything is going wrong. With this much inventory the organization ought to be able to keep its cus-
Improving Supply Chain Performance by Using Order Fulfillment Metrics

Customers happy; yet despite substantial investment in inventory, it is losing sales because of inadequate service. This problem could result from holding surplus stock of unpopular products or components while stocking out of those most demanded by customers. Another explanation might be that the order processing systems cannot turn stock around quickly enough to meet customer demand, and other operational difficulties might contribute to the problem.

For an organization in quadrant B, it should be easy to provide good service by holding lots of inventory. If an organization in quadrant C holds too little inventory, product availability will diminish and customer service will suffer. However, this simple matrix doesn’t capture all aspects of the delivery performance problem. Although the poor service indicated by quadrant C might be explained by insufficient stocks, other factors might contribute. More specific measures of performance might help reveal the cause of the imbalance between inventory and service.

Naturally, it is best to operate in quadrant D. Here relatively low inventory investment does not prevent a high degree of customer service, so inventory assets are being put to highly productive use. Other system attributes like fast cycle times or uniformly accurate order configurations may contribute to this productivity.

The remaining challenge is to add more substance to this simple matrix. Suitable measures for service and inventory must be chosen, and these measures should reflect the nature of the product and the market. After selecting the appropriate measures, managers need to locate the breakpoint between “good” and “bad” performance for the business.

Hewlett-Packard has widely adopted the framework depicted in Exhibit 4. The company most commonly defines “fill rate,” a measure of customer satisfaction with delivery performance, as the percentage of items found in stock when requested. In practice, the definition might vary somewhat; for example, HP might focus on immediate, or off-the-shelf, fill rate. Other times, as in the case of many build-to-order products, the appropriate measure might be the percent of orders filled on the quoted delivery date.

Inventory also can be assessed in many ways. HP often uses a common measure, weeks of finished goods inventory (FGI), based on “average” volumes. However, the company sometimes finds unit volumes, dollar volumes, total pipeline inventories, or inventory turns better suited to describe the situation at hand, especially for the build-to-order businesses.

Exhibit 4 plots individual points summarizing performance for different product lines. The product lines are similar, so they can be judged by using consistent measures of acceptable and unacceptable performance. In fact, this chart looks very much like a monthly summary management report. (A brief legend might explain any special circumstances leading to the month’s results; for example, a product mix problem for product line 2.)

In contrast, Exhibit 5 shows the "efficient frontier" for
a single product line. This term, familiar to many from economics, indicates the limit of what is achievable, given current operating conditions—things like manufacturing cycle time, supplier performance, and forecasts of varying customer order patterns. The solid line in Exhibit 5 indicates the best that product line managers can hope to do under the circumstances. With an inventory investment of two weeks of supply, they can expect at best about a 93 percent fill rate. At seven weeks of supply, the fill rate they can reasonably strive for is 98 percent. In practice, however, operations generally find it difficult to achieve the customer service target that theory predicts for their level of inventory. The day-to-day “friction” of a real factory or warehouse catches up to them. Hence, the actual performance levels
are expected to rise above the boundary defined by the efficient frontier. But the theoretical operating line encourages managers to move their performance achievement ever closer to the frontier.

The exchange of inventory for service depicted in Exhibit 5 follows from analytical models of supply chain physics. Given a supply chain, the curve shows all the feasible points of operation. But the curve does not indicate the "best" operating point. Managers must make that decision based on the costs of holding inventory and the costs of achieving a given service level.

In a perfect economic world, managers would all find their way to the best operating point. However, many forces prohibit them from taking that route. Operational realities can bounce organizations around; if orders are weak, managers are likely to see exceptional service levels, even in cases when stocks are low. Thus, sometimes they might even see performance that looks better than the so-called theoretical optimum.

Over the long haul, another dynamic affects performance. Imagine the following scenario. Orders begin to drop somewhat, so management brings in a new marketing director. To stimulate sales, the new director decides that the firm needs to improve customer service. A pointed memorandum makes the rounds. Customers must be delighted with our responsiveness and reliability! After a few months, orders are back on track and service ratings are higher than ever. Nevertheless, a quarter or two later, inventory levels have reached the point where cash flow has become a real issue, so the controller articulates an equally firm memo: Inventories must shrink!

What happened? To improve customer service, management authorized an increase in stock levels. Absent other changes to the supply chain, inventory must increase to improve customer service. Similarly, if managers want to reduce inventory levels, they must expect service to drop. Sometimes when a sweeping change in policy is pursued (Service must improve), some of the inefficiencies in the system are cleaned up. Overall performance really does get better. But all too often the simple dynamic illustrated in Exhibit 6 is at work. The pendulum swings from a state of high inventory and high service to a state of low inventory and low service. In essence, power oscillates between the marketing (service) camp and the finance (inventory) camp.

In fact, this dynamic caught up with Hewlett-Packard in 1993. Almost a year into a focused effort to improve order fulfillment, the company nearly ran out of cash. HP reached its credit limit with the banks, a shocking turn of events for the conservative firm. In response to the organizational push to improve service, attention drifted from the problem of managing inventory. Stocks accumulated to meet customer service goals. Now, it is safe to say, the company is much more keenly focused on the measurement scheme advocated above. Management better understands the natural exchange between service and inventory. Managing one without paying attention to the other just doesn’t work.

**IMPROVING ORDER FULFILLMENT**

There are as many ways to fight this counterproductive dynamic as there are managers. Fundamentally, however, there are only two solutions. The curve shown in Exhibit 5 is defined by theory for a given supply chain, for a given set
of policies, and for the prevailing operating environment. Constrained to that curve, managers can never improve service and reduce inventory investment at the same time (aside from relatively minor gains by squeezing out latent inefficiencies). But what if the supply chain itself changes? This means exchanging the "physics" of one system for those of another by changing the way the product is designed, produced, and delivered. Taking this approach, managers can make great gains at relatively low cost, leading to improvements in both realms.

The measurement scheme depicted in Exhibit 3 really took hold at HP and the inventory scare. The severity of the situation—at a time when the company was struggling to fund rapid growth in a number of businesses—drove home the importance of controlling both dimensions simultaneously. By setting guidelines for acceptable performance, upper management sent a strong message to the product lines: Find your way into the golden zone. Previously, managers had responded to pressure to improve in one dimension or the other by sliding up and down the exchange curve. Now more was required.

Consider the two product lines depicted in Exhibit 7. The supply chain for product line 1 is positioned for success. Its exchange curve crosses squarely across the golden zone in the lower right of the chart. Managers for this product should find it relatively straightforward to post high-performance scores without major structural change. However, the managers of product line 2 face radical change if they hope to meet the company’s goals for service and inventory investment. Their curve doesn’t even cross through the golden zone. Even if their supply chain performs at its theoretical maximum level of efficiency, they cannot perform acceptably. They have no choice but to jump from their current curve to one more closely resembling that of product line 1.

Techniques for moving a product line from one curve to another include process improvement and process change. Process improvement means refining the supply chain to reduce the level of uncertainty in its performance. Process change means changing the supply chain to eliminate the uncertainty altogether.

A CLOSER LOOK AT PROCESS IMPROVEMENT

Like other processes, order fulfillment can be controlled and improved using well-known techniques of quality management. Decades ago, W. Edwards Deming taught that process variability was the root of quality problems. Nevertheless, it wasn’t until the 1980s that the focus on statistical process control, total quality management, total quality control, and Motorola’s six-Sigma program demonstrated that quality could be ensured by identifying and controlling key processes. The same ideas and tools used to improve manufacturing processes can be applied to supply chain processes. The application of statistical process control to the order fulfillment process is a natural extension of the strategic objectives of the enterprise. Throughout the supply chain, key process variables can be identified and tracked with the goal of improving the order fulfillment process. Quality tools like fishbone and Pareto diagrams can be very helpful in identifying key variables.
There are four major categories of variables related to the order fulfillment process: forecast, supply, process, and transportation. Forecast variables track the ability to estimate future product demand for different forecast horizons. Supply variables track the vendor's ability to deliver components reliably enough for the manufacturing process. Process variables measure the ability to maintain cycle time targets. Transportation variables measure transportation times and reliability up and down the supply chain.

Determining which process variables to monitor is a challenge when using techniques like statistical process control. As in monitoring a manufacturing process to ensure quality control, some process variables are very important while others may have little impact on defects. Finding the sensitive variables requires historical operating data where defects can be linked to variables in the process. In the order fulfillment process, defects are late deliveries to the customer, product unavailability, and inventories that are off target. Linking these occurrences with process variables like transportation time requires managers to track orders and inventories, isolate defects, and track defects to their root causes.

HP has put a system in place to do just that for all its businesses. The system tracks order information, including customer contact, promised dates, and delivery dates. With this information, individual order histories can be traced and compared with operating data like inventory levels to identify the root causes of defects.

**PROCESS CHANGE ACROSS THE FUNCTIONS**

Supply chain management is crossfunctional. Although traditional functions like R&D, manufacturing, and marketing each play a role, the issues also cut across entity lines, with players elsewhere in the system often championing changes to the supply chain. Too often, inappropriate metrics obstruct close cooperation between members of the supply chain. Within a vertically integrated company, an upstream supplier division may have little incentive to do something that helps a downstream division—even if those downstream improvements are substantial. As many have joked, just in time often means little more than trying to force others to hold inventory you don't want to pay for. When the different players do make decisions in the best interest of the entire system, radical improvements frequently result. The most valuable supply chain decisions lead to substantial reductions (or complete avoidance) of uncertainty, greatly reducing the amount of inventory required as insurance against an uncertain environment.

Although the decisions involve players from different functions, each function generally takes the lead for different kinds of supply chain changes—for example, manufacturing changes that affect where products are made. All else being equal (production costs, for example), situating production close to the customer improves supply chain performance. For a product shipped by boat around the world, this could mean a lead time reduction of six weeks or more. This sharply reduces the impact of variable demand, which means lower safety stock levels. Moreover, local production often leads to reduced tariff expense. And in many cases, marketing likes the extra benefit of having a strong local presence. The design of the physical manufacturing network has a profound effect on order fulfillment.

Marketing, in turn, controls which products are made. Given the opportunity, most marketers would like to have countless differentiated products available to fill every niche in the market. But this generally comes at a substantial cost to the product delivery system. By carefully managing the proliferation of products, marketing can exert its own influence on the efficient frontier. In fact, one of the ways marketing functions can strip uncertainty out of the system is by terminating old products instead of letting them linger on the corporate price list. Another way that marketing influences the order fulfillment process is by managing distribution channels. By helping to set the right expectations in the minds of the customers, marketing can further aid customer service improvements.

R&D ultimately controls product design, and their design decisions can have a critical impact on downstream manufacturing and distribution processes. One newly rediscovered technique to aid supply chain performance is “postponement.” This strategy calls for product design changes that permit the product to be only partially assembled at the factory into generic units. The final assembly is performed in a light manufacturing process conducted at the distribution center. This technique is especially effective when the same, generic base unit can be quickly transformed into a number of alternative final products. Differentiation of the product is postponed until firm orders arrive, greatly reducing dependence on accurate forecasts. For example, HP developed a desktop printer that could be quickly configured for Windows or Macintosh markets. Supply chain-friendly design decisions may add to the product’s material cost, but they can mean big savings in manufacturing, transportation, inventory, or even service after the sale.

Finally, distribution has dominion over its own share of crucial supply chain issues. For example, this group often decides how a product is transported—a fundamental supply chain issue. To solve a delivery problem, distribution might propose shipping the product by air instead of by sea. The immediate response to this is generally negative.
since air shipments are costly, but (with co-locating factories near customers) air shipment can sharply cut lead times and ameliorate the effect of uncertain demand. For the right products, this benefit outweighs the additional cost. Distribution also provides the link to the customer; this group can make changes to improve order processing. And in some industries, the manufacturer’s distribution team manages the stock levels of their customers. Forging this link through closer cooperation can eliminate much more of the variability that plagues the supply chain.

Recently a crossfunctional team at HP conducted a thorough supply chain analysis in support of a new product development effort. Led by the new product introduction manager (resident in manufacturing), the core team was comprised of players from manufacturing, materials, marketing, finance, and R&D, and it was backed up by a supporting cast from the worldwide distribution centers. Its initial charter was to find a better way of making and distributing HP products.

By the end of the project, the team had influenced product design decisions as well as the design of the manufacturing network. In particular, it persuaded R&D to spend a bit more on the power supply so that the same generic unit could be sold on all continents (different models had been required to match different nations’ voltage and frequency standards). This move protected them against demand in the different regions going against predictions, so they wouldn’t be left holding too much of the wrong model. The project team also identified a cost-effective postponement strategy that further served order fulfillment needs. This is a shining example of how changes to the process—a new set of operating conditions—can move performance from one service/inventory-efficient frontier to another.

To implement a measurement scheme like the one proposed here requires specifics. Unfortunately, each business situation calls for a unique combination of attributes—different metrics applied to different products in different markets. But here are some general pointers to help ensure success.

**Product Aggregation**—Most instances of measuring and improving order fulfillment address product lines; this is the highest level of aggregation commonly used for studying inventory. But that’s a pretty high degree of aggregation. Where is the focus on individual products, assemblies, and materials? After all, that is what companies really stock.

What managers capture in their measurement and control system depends on their role. For senior managers, a product line focus is perfectly appropriate. They don’t need to trouble themselves with the details of single items at remote depots. If they are managing several businesses, they need to see at a glance how each of those businesses is currently performing. Doing this means combining data on hundreds of different items. Fortunately, everyone understands that some precision is lost when results are aggregated to such a high degree. Simple weighted averages work best. The same logic can be implemented at lower levels in the organization. Throughout the data reduction process, the most important consideration is consistency. When senior managers focus on specific performance measures for service and inventory investment, the responsible product line managers should track things in the same way. Likewise, product managers should use the same metrics. Product line managers should look in detail at the performance of their whole portfolio of products. And individual product managers should concern themselves with product-specific results, like the inventory/service picture at each warehouse worldwide.

If managers up and down the chain look at things in different ways, someone, sometime, will get burned. Different managers will find themselves working at cross purposes. For example, if warehouse managers are asked to focus on item fill rate but senior managers look at order fill rate, there is a mismatch. To meet the goal, the warehouse manager might implement a change that actually hurts order fill rates (for example, concentrating on filling as many partial orders as possible). But if the same metrics are used throughout the organization—and everyone knows what they mean and why they’re looking at them—the business’s overall performance will improve.

Likewise, it is important to use the same performance objectives at all levels of the business. A manufacturing manager is unlikely to achieve a 98 percent service goal if the distribution center gets the message that 95 percent is good enough. Of course, for different products and different businesses, the goals will vary. But within the individual product supply chain, the players should align themselves with a common goal.

**Exhibit 8** illustrates the cascading effect of a consistent measurement scheme. Performance reports for the product managers should look the same as those for the product line managers, whose reports, in turn, should be consistent with those seen by senior managers. The only differences should be in the level of detail.

**Service Measures**—Service measures typically address one of two situations “immediate” shipment of a product and time to complete an order. First, consider metrics that capture the ability to fill an order for a particular item. Line item fill rate (LIFR) is the most popular metric of this type; however, there are many related metrics. For example, fill...
Improving Supply Chain Performance by Using Order Fulfillment Metrics

EXHIBIT 8. A Measurement Hierarchy

rates may be tracked for a particular product or product family; this is sometimes referred to as item fill rate. Rather than track lines or items, some firms track dollar-weighted fill rates by revenue or by profit contribution. Dollar fill rates are sometimes popular among executive management because they appear “financial” in nature and relate the possible financial severity of an order fulfillment problem. This can be deceiving, since high dollar fill rates may still generate little customer satisfaction. Surprisingly, few managers seem to use order fill rate, which measures the successful completion of a customer order. Although there are situations in which order fill rate may not be appropriate, in many cases it is the most important metric.

Generally, the customer need not see the inventory to place an order. This distinction is important because fill rates have little meaning if customers don’t order when the product is not available. For example, in grocery distribution, it is common for distributors to inform stores of products that are currently unavailable. Stores then order substitutes or delay orders until the product is available again. Tracking fill rates on such products is meaningless because they would always appear high. In fact, stores simply quit buying when supplies are exhausted.

A metric better suited to this situation is the proportion of time a product is available—or, conversely, the probability of stocking out within a certain time period. Availability metrics may be particularly useful in retailing environments, where the primary concern is maintaining just enough inventory on the shelf to prevent lost sales. One key characteristic of this environment is the urgency of having the inventory on the shelf. If customers provide no warning of a pending purchase, and are likely to go elsewhere should they not find what they are looking for, then availability is critical.

Time-based measures make up the second class of metrics. Some relate to the average speed of processing and delivering an order. Others relate to exceptions, such as the expected time before a back-ordered item can be delivered. Many companies use turn-around time as an internal operations measure; the customer may never know what the target turn-around time is. A related metric looks only at backordered items—average time to fill a back-ordered item. While these metrics are important, they are often internally focused and largely ignore the customer’s needs, so they fall short in measuring the true goal of the firm.

One need not reflect long on the traditional metrics to see their limitations. For example, as noted earlier, traditional fill rate measures rarely track completed orders. Thus, they are useful in situations where complete orders are not important or in cases where a single product represents a complete order. A printer manufacturer may be successful at using item fill rates as a measure of service to downstream computer resellers. But using an item or line fill rate measure in situations where customers order a bundle of related products can lead to problems. For example, if each customer ordered 10 to 20 items per order and the company consistently shipped all but one item, line item fill rates would look okay. If the shorted item was inexpensive relative to the others, dollar weighted fill rates would look even better. However, if the shorted item was critical to the performance of the products ordered—like the mouse in a computer order—customer satisfaction would be far lower than indicated by the fill rate measures.

Many would argue that order fill rates are not very important when the customer is actually a distributor and not the final customer. For example, HP ships printers to Best Buy’s central distribution center, which then distributes them to the stores. An incomplete order for Best Buy might not have a huge impact because its distribution center inventorizes products, further protecting their stores from stockouts. However, Best Buy must hold more inventory to protect against partial shipments—a source of supply variability. Moreover, the cost of tracking orders that trickle in rather than arrive together is higher. In the end, complete order fill rate is important to Best Buy and, therefore, to HP.

A compounding problem with the use of fill rates in
many organizations is the time frame. Although an order may not ship immediately because of a missing item, it may be ready for shipment within hours, causing little erosion in customer satisfaction. Thus, in some cases a fill rate measure may drive a manager to carry too much inventory. For example, one of HP’s major distribution centers tracked LIFR religiously. However, measures were based on the ability to ship a product on the first try, after the order cleared administrative checks. At the time, an order took a few days to work its way through credit departments before reaching the warehouse. If the items were not available then, HP penalized itself in the LIFR calculation. However, the order might be filled later that day after resupply from the factory, still in time to make the last truck out. This would not compromise on-time delivery to the customer. The point is this: When did the customer want the order?

A survey of common service metrics found two principal shortcomings. First, few metrics are truly driven by the customer. Second, most metrics ignore variability. Naturally, metrics that capture these attributes without sacrificing the appeal of more traditional measures like fill rates are preferred.

Order windows—derived from the concept of an order aging curve—can be used to track the performance of the delivery process by measuring the percentage of orders delivered within the customer’s delivery target. The aging curve shows the cumulative percentage of orders filled over the time horizon. For example, Exhibit 9 shows that 20 percent of the orders were delivered on time or early, 60 percent within one day, and 95 percent within two days.

The curves present a wealth of information—much more than a single fill rate number. Because they are based on complete customer orders, they provide a better indication of performance as perceived by the customer than item, dollar, or line fill rates. They also help a manager ascertain the system’s variability.

For example, Exhibit 10 shows the order performance for two different companies, with both companies filling about half their orders immediately. However, the aging curve shows that company A is more reliable, with nearly all its orders delivered within one day, whereas only three-quarters of the orders are filled in that window by company B.

A set of internal performance metrics could be generated simply by looking at the percentage of orders delivered within various order windows. For example, in cooperation with marketing, three windows might be specified as green (on time), yellow (a little late), and red (very late). Then the percentage of orders delivered within each category could be tracked with the goal of shrinking the size of the red and yellow categories.

Likewise, an externally focused metric could be developed based on customer needs. Each time an order is placed, the customer could specify the order window defined by the earliest delivery date and the latest ship date. Often customers do not want the product immediately and, thus, the earliest delivery date is specified as the earliest date the customer would accept the product. If the product were needed immediately, the customer could specify the earliest date.
Improving Supply Chain Performance by Using Order Fulfillment Metrics

based on the minimum shipping time. The latest delivery date defines the boundary of acceptable performance. Using this window, delivery performance can be tracked. (Bear in mind the importance of tracking shipment date and arrival date independently; this helps identify variability introduced in transit.) Metrics based on order windows capture the most important aspect of the delivery process: reliability. As many have noted, in the world of quality management, reliability is the key to process control. Genichi Taguchi has argued this point best. He notes that the mean of a process is not difficult to change; reducing the variability is the more difficult problem. As noted earlier, customers expect and require reliability.

Inventory Measures—Inventory can be measured in many different ways, but most methods are fairly straightforward. If you have created logical groupings of products as described above, most of the different measures relate linearly; units tie neatly to dollar volumes and to weeks of supply.

Most of HP's work has centered on distributed products. Consequently, it has grown accustomed to focusing on finished goods inventory ready for distribution to customers. Finished goods inventory has the highest value of all the materials in the supply chain, so between volume and value, it carries the most weight. HP generally prefers to measure finished goods inventory in terms of weeks of supply. The principal advantage of using weeks over unit or dollar measures is that it makes it easier to compare performance for two different products. By normalizing the size of the inventory against the average demand, the confusion of different scales is avoided.

For example, if the company has $8 million in inventory of printer A and $1.4 million of similarly priced printer B, which product has the bigger inventory problem? Printer B does because demand for that printer is one-twelfth that of printer A. The situation is clearer if we say we have four weeks of supply of printer A and over eight weeks of supply for B.

A situation in which the product is built to order and there is no FGI requires a slightly different metric. Logical candidates include the combined dollar value of the raw materials (and WIP) and the weeks of supply of critical components. As with service metrics, managers probably know already what the "right" measure of inventory is for their business. The important thing is to maintain consistency in the metrics applied, so that performance measurements compare readily.

Although inventory turns are widely understood and applied—and for many the only measure of inventory—this metric has one important disadvantage. When displayed graphically, it can be misleading. Consider Exhibit 11, which uses turns as the measure of inventory. (Remember, more turns mean more productive inventory, so the desirable operating region on this chart is the upper right.) The two product lines depicted each improved their inventory performance from May to June. Line A went from four turns to five. Line B went from nine to ten. Their performance looks the same on the chart. But line A reduced inventory by 20 percent, whereas line B only improved by 10 percent. It is...
difficult to say a priori which product line accomplished the most in this case; the 10 percent gain may be better for the company than the 20 percent improvement. And many old hands have an intuitive feel for the real impact of different numbers of inventory turns. But for the rank and file, a linear measure—like weeks of supply—makes better sense as a visual management tool. You can calculate weeks of supply as 52 weeks divided by the number of turns.

There is another interesting difficulty with inventory turns. Many managers mistakenly believe that high inventory turns benefit the customer. This may be true when product freshness is an issue, but inventory turns say nothing about how well the order fulfillment process is working in terms of customer satisfaction. All else being equal, higher turns should lead to lower service. “Good” inventory performance, be it high turns or low weeks of supply, judges only financial performance of the inventory asset. Seldom do people confuse inventory reductions measured in weeks of supply with improvements in service.

THE IMPORTANCE OF CLARITY AND CONSISTENCY

Robust metrics, applied consistently when establishing performance objectives, are essential to improving order fulfillment. But developing refined metrics to improve supply chain performance requires a clear understanding of the process. At Hewlett-Packard, managers found that reaching this understanding was itself a challenge. CEO Lew Platt summarized HP’s order fulfillment breakthrough goal in a talk with employees: “You’ve heard the story of the blind man walking around an animal trying to find out what it is. I think this was the year we discovered our order fulfillment problem was an elephant . . . .”

Although the specific metrics chosen will vary according to the situation, all organizations will benefit if they:

- Establish performance objectives with the customer in mind.
- Consider using order windows as the basis of order fulfillment metrics.
- Reflect reliability issues in the metrics they choose.
- Implement metrics consistently throughout a single supply chain.
- Aggregate results as they move up the management chain.
- Apply process control techniques to the business processes.
- Avoid pitting players in the system against one another.
- Collect only the data they really intend to use.
- Communicate their actions and rationale to everyone.

The measurement scheme proposed here—simultaneously tracking the relationship between service and inventory while monitoring and managing delivery reliability using order windows—provides an approach for individual businesses to address order fulfillment problems and improve supply chain performance.