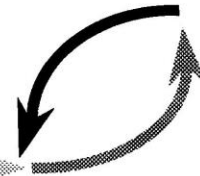




The Systems Thinker



Vol. 3, No. 7

September 1992

Service Quality Excellence: Mastering the "Moments of Truth"

By Daniel H. Kim

In recent years, Total Quality Management (TQM) has moved from a manufacturing improvement process to one that can enhance *all* company operations. While the '80s shook up complacent manufacturers and forced them to compare the quality of their products to a new breed of competitors, the '90s is becoming the decade in which service industries are waking up to the quality challenge. Appropriately, in 1990 Federal Express became the first service company to win the Malcolm Baldrige National Quality Award.

But the importance of service quality is not limited to service industries

like banking, insurance, or package delivery. As the global environment grows fiercer every day, manufacturing companies are realizing the importance of focusing on the quality of their services, not just their products.

vide the customer with high quality. Each of those interactions are the "moments of truth" which determine whether you are seen as a high-quality service operation or not, according to Jan Carlzon in his highly-acclaimed

"You can't touch or feel a service, nor can you inspect it after the service is completed. At the point of 'delivery,' everything that is needed to provide the service must converge in order to provide the customer with high quality."

"Moments of Truth"

Improving service quality promises a myriad of benefits. It costs far less to keep a customer than to win a new one, for example, and perceived high-quality service firms can often charge up to 10% more for their products than competitors. The question is *how* to do it.

In the rush to reap the benefits of improving service quality, companies may be too quick to borrow from past experience in manufacturing. But there are important differences that make improving service quality much more elusive.

One obvious difference between products and services is that one is tangible while the other is not. You can't touch or feel a service, nor can you inspect it after the service is completed. At the point of "delivery," everything that is needed to provide the service must converge in order to pro-

vide the customer with high quality.

Quality, as seen by the customer, is determined by each moment-of-truth encounter with front-line personnel. The net benefit of millions of dollars worth of capital equipment, buildings, salaries, etc. that a company has assembled will be judged, in large measure, by the quality of those interactions. Those moments of truth are numerous, ephemeral and difficult, if not impossible, to measure. And yet, the long-term reputation and success of a service company is largely riding on them. This suggests that many companies are investing far too little in their front-line personnel. How many moments of truth are actually "moments of despair" for their customers?

Just-in-Time vs. Just-in-Case
For some companies, providing

Continued on next page

In This Issue

(Symbol key for diagrams on p. 10)

Service Quality Excellence: Mastering the "Moments of Truth"	1
From the Headlines	4
<i>Utilization Review and Healthcare</i>	
Toolbox	5
<i>Using "Fixes that Fail" to Get Off the Problem-Solving Treadmill</i>	
Feedback/Followup	7
In Review	8
<i>A Systems Thinking Approach to Executive Information Systems</i>	
Calendar	10

Printed on recycled paper

Continued from previous page

high quality service means creating a just-in-time (JIT) operation, where all the necessary ingredients converge at the point of delivery exactly when it is needed. In a JIT production system, inventories are kept to a minimum throughout the factory by making sure there is just enough inventory at each step of the process to supply the next batch. But it is dangerous to carry the JIT philosophy too far in the service arena.

In manufacturing, the product has already been designed; all that remains is to run the production line as smoothly as possible. Variances in the production line can then be controlled and maintained. In a service setting, however, front-line personnel have to be ready to produce a service whose design is not fully complete until they interact with the customer. Unlike the manufacturing setting, customers often introduce variances that cannot be controlled in advance.

Having adequate capacity on-line is critical to providing high quality service. In a JIT production system, if the production line goes down, the downtime does not affect the quality of the next product off the line. If there are enough buffer stocks of finished goods, the customer won't even experience any difference in delivery. If you are under capacity in a service setting, however, there is no way to make it up in real time with "buffered" service time. In addition, it is virtually impossible to "recall" a poorly delivered service. A flight that arrives two hours late and causes people to miss a meeting cannot be changed. A package that is delivered too late for a speech is simply too late. The capacity has to be on-line and available precisely when the customer requests it. This suggests that, unlike JIT, one may need to plan in terms of *Just-in-Case* service capacity—that is, service capacity should be weighted more towards peak volume than average volume.

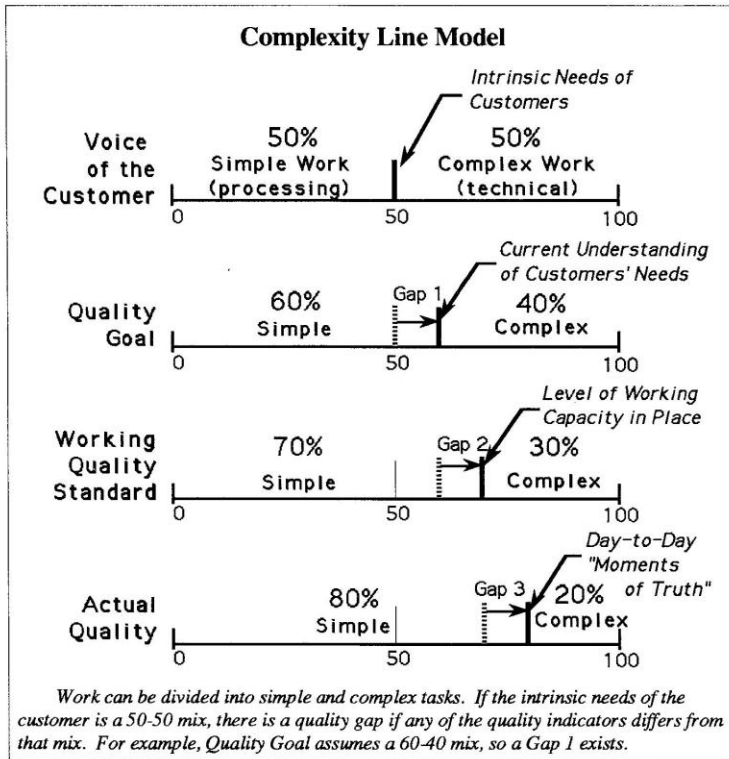
Complexity Line Model

The Complexity Line Model offers a way of looking at service quality in terms of four different capacity requirements: voice of the customer, the quality goal, working quality standard, and actual quality. In the Complexity Line Model, all service work is viewed either as simple (processing) work or complex (technical) work. In reality, there are many gradations, but for ease of use we will work with the two categories. Simple work means things that can be handled by an entry-level person. Complex work, on the other hand, requires a lot more experience and skill. In general, complex work also requires more time.

Suppose we are managing a customer service call center. Ideally, we should be staffed to match the exact needs of the customer as shown on the Voice of the Customer line (see "Complexity Line Model" diagram). Suppose that the intrinsic needs of the customer calls coming into our center are split 50-50 between simple and complex. That is, on any given day, 50% of the calls are routine. The other 50% are complex and require more understanding about the business.

Gap 1: Understanding the Voice of the Customer. In reality, we never know exactly what the customer needs. The customers themselves may not fully know what they need. The *quality goal* line represents our current understanding of the customer's needs. For example, we may be staffed and prepared to handle a call volume that we believe is 60% simple and 40% complex. Gap 1 represents the difference between what the customer actually needs and what we think the customer needs. In this case, 10% of the customers will not receive proper service. Reducing this gap requires an investment in understanding what the voice of the customer is.

Gap 2: Understanding the Voice of the Process. The Working Quality Standard line may be different from the quality goal if the service capacity in place is not sufficient to provide the stated quality goal. In this case, say the people doing the work are capable of handling a call volume where 70% of



the work is simple, and 30% is complex. Gap 2 represents our lack of understanding of what our current system is capable of (voice of the process) relative to what we are asking it to do (quality goal). In this case, we fall short of our own goal by another 10%.

Gap 3: Managing Customer-Generated Variance. The Actual Quality line represents the day-to-day moments of truth in which the customer actually experiences our service quality. If we are staffed to meet a 70-30 complexity mix, and the volume of calls stays relatively constant, customers will experience quality at the working quality standard level. Suppose, however, that incoming call volume suddenly jumps by 20%. What will the pressure in the system do to actual quality? The only way to serve a larger number of customers with the same number of people and skill mix is to reassign complex work as simple. The work can then be given to less experienced staff ("I know Joe's only 2 weeks on the job, but I think he can handle this one") or we can treat it as simple and spend less time on it ("I don't think they need to know about all the other options...").

So now work is handled as if it is 80% simple, 20% complex, which represents the *actual* quality. Gap 3 represents the daily adjustments that have to be made when customer volume and special requests exceed the capacity established by the working quality standard. Although our working quality standard has not changed, actual quality has grown worse. If this becomes a frequent occurrence, the quality stan-

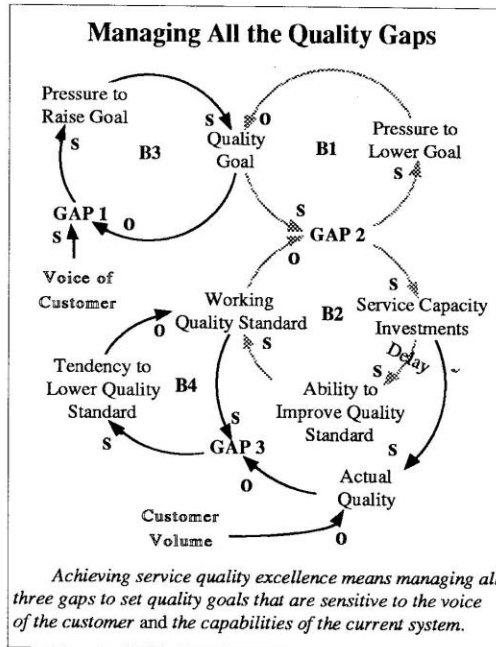
dard can be pulled downward toward actual quality. As the quality standard adjusts to a lower level, actual quality can get pulled down still further the next time the call volume exceeds the already lowered capacity (see "Quality Erosion over Time" diagram).

Drifting Goals Structure

The dynamics of service quality can be captured in a "Drifting Goals" archetype (B1 and B2 in "Managing All the Quality Gaps"). A gap between the quality goal and working quality standard can be closed in one of two ways—lowering the goal (B1) or raising the standard (B2). Lowering the goal is easy and quick; raising the standard takes time and investment (see "Drifting Goals: The Boiled Frog Syndrome," *Toolbox*, October 1990).

Focusing on the needs of the customer can help balance the pressure to reduce the quality goal. The voice of the customer increases Gap 1 which increases the pressure to raise the quality goal (B3). In this model, we see that the art of *setting quality goals* requires balancing the voice of the customer with the voice of the process. In terms of TQM, this means continually trying to identify the intrinsic needs of the customer (voice of the customer) and understand the systems and processes enough (voice of the process) to design them to be in line with those needs.

Maintaining the working quality standard



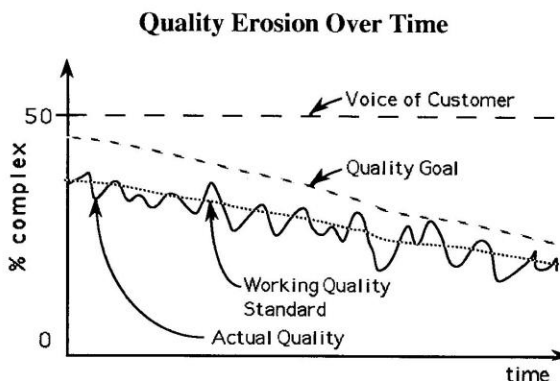
without losing ground requires managing the gap between actual quality and working quality standard. A high-quality operation should have adequate capacity to handle the majority of the variance it encounters and should keep actual quality within a narrow band around the working quality standard. An operation that is out of control would have a wildly-fluctuating and persistent gap.

Managing All the Quality Gaps

Achieving service quality excellence means managing all three gaps simultaneously. Focusing exclusively on the customer and making your quality goals aligned with their needs will reduce Gap 1, but it will only make Gap 2 worse. If you raise the quality goal without investing in the requisite training, personnel and systems, employees will see it as nothing more than banner-waving and go on with business as usual.

If you focus exclusively on reducing Gap 2, however, you may encounter tremendous pressure in the system to close the gap by lowering the goal.

Continued on next page



Continued from previous page

The history of the quality standard can often provide compelling evidence that the quality goal is out of line with the "real" system, and lead to "Drifting Goals." Focusing exclusively on eliminating Gap 3 will create an identical tendency for the standard to float with actual quality.

The customer's experience of quality is determined by the sum of all three gaps. The challenge for service

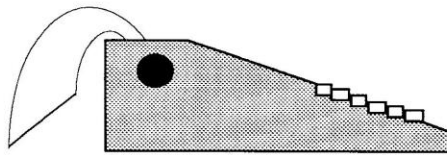
companies (as well as manufacturing firms with service operations) is to develop the ability to identify and eliminate all three gaps even as the voice of the customer continually changes. It requires investing in service capacity ahead of the current requirements in order to be able to treat each moment of truth with the quality that the customer intrinsically needs or wants.

The Complexity Line Model is based on the work of Bob Bergin and Gerri

Prusko at Hanover Insurance Co. (Worcester, MA). The author has developed and used a Service Quality Management computer simulator to provide practice fields for managers to understand the complexity line concepts. The software runs only on Macintosh computers. If you wish to acquire a copy, please write to Daniel H. Kim, MIT Organizational Learning Center, E40-294, 1 Amherst St., Cambridge, MA 02139.

Further reading: "Now Quality Means Service Too," Fortune, April 22, 1991; Jan Carlzon, Moments of Truth (New York: Harper & Row, 1987).

From The Headlines



Some generic structures occur so frequently that they have been identified as a generic set of tools called "systems archetypes." This column is designed to help readers recognize archetypes at work in newspaper and magazine articles.

Utilization Review and Healthcare

"[M]edical cost cutters known as utilization review companies...have become immensely popular with employers seeking to control their health-care costs...[D]oubt is growing about the new industry's most basic promises—lower medical costs...For every form a utilization firm sends out and every phone call one of its employees makes, someone in the medical community must respond, usually with equal or greater effort...To some industry specialists, the issue embodies a Catch-22 of medical economics: Efforts to police costs add to the industry's administrative burden." ("Firms that Promise Lower Medical Bills May Increase Them," *The Wall Street Journal*, July 28, 1992)

* * * *

Utilization review (UR) procedures have become common in healthcare, as employers try to manage costs by restricting and limiting unnecessary care. For a while, such procedures worked: a study in *Medical Care* that analyzed 223 insured groups from 1984 to 1986 concluded that hospital admissions decreased 13% and overall medical costs were cut 6%, as a result of UR.

But *The Wall Street Journal* article raises concerns about the overall finan-

cial benefits of these procedures, because of the great administrative burden they place on doctors and hospitals. Some analysts estimate that 20% of the current healthcare bill can be attributed to excess paperwork and other administrative-related tasks. By the year 2020, such spending could amount to half of

total healthcare dollars.

Although the UR system worked well at cutting costs at first, recent problems and increased administrative costs for providers suggest that the UR industry is suffering from a "Fixes that Fail" situation.

UR is popular because the cost savings far outweigh the fees charged to employers (B1 in "The Utilization Review Fix that Failed"). But over the long term, the increased costs to providers will lead to increased costs of services, further cost-shifting to patients, and in the end, greater costs for employers (R1). The increased costs are invisible for the most part, since the extra time it takes to answer a phone call or chase down information for a report is not measured.

Clearly, UR firms have helped employers deal with their biggest concern with healthcare: increasing costs. But the "Fixes that Fail" archetype suggests that any successes will be short-lived. As the healthcare system is revamped, it is critical that employers and providers focus on the long term and make sure any proposed "fixes" will not worsen the healthcare crisis (see "No More Band-Aids for Health-care Reform," *Cover Story*, August 1992).

The Utilization Review Fix that Failed



As healthcare costs skyrocket, utilization review helps businesses reduce their spending (B1). Over time, however, the additional administrative burden on providers caused by UR leads to higher costs, which are eventually shifted back to the employer (R1).