Robust Quality

Genichi Taguchi and Don Clausing

Harvard Business Review

90114

HBR
JANUARY—FEBRUARY 1990

Robust Quality

Genichi Taguchi and Don Clausing

When a product fails, you must replace it or fix it. In either case, you must track it, transport it, and apologize for it. Losses will be much greater than the costs of manufacture, and none of this expense will necessarily recoup the loss to your reputation. Taiichi Ohno, the renowned former executive vice president of Toyota Motor Corporation, put it this way: Whatever an executive thinks the losses of poor quality are, they are actually six times greater.

How can manufacturing companies minimize them? If U.S. managers learn only one new principle from the collection now known as Taguchi Methods, let it be this: Quality is a virtue of design. The "robustness" of products is more a function of good design than of on-line control, however stringent, of manufacturing processes. Indeed—though not nearly so obvious—an inherent lack of robustness in product design is the primary driver of superfluous manufacturing expenses. But managers will have to learn more than one principle to understand why.

In 1989, Genichi Taguchi received MITI's Purple Ribbon Award from the emperor of Japan for his contribution to Japanese industrial standards. He is executive director of the American Supplier Institute and director of the Japan Industrial Technology Transfer Association. He is also the author of The System of Experimental Design *(ASI Press, 1987). Don Clausing, formerly of Xerox, is the Bernard M. Gordon Adjunct Professor of Engineering Innovation and Practice at MIT. He has edited Mr. Taguchi's works in English and is a leading exponent of his views. He is the author (with John R. Hauser) of "The House of Quality" (HBR May–June 1988).*

Zero Defects, Imperfect Products

For a generation, U.S. managers and engineers have reckoned quality losses as equivalent to the costs absorbed by the factory when it builds defective products—the squandered value of products that cannot be shipped, the added costs of rework, and so on. Most managers think losses are low when the factory ships pretty much what it builds; such is the message of statistical quality control and other traditional quality control programs that we'll subsume under the seductive term "Zero Defects."

Of course, customers do not give a hang about a factory's record of staying "in spec" or minimizing scrap. For customers, the proof of a product's quality is in its performance when rapped, overloaded, dropped, and splashed. Then, too many products display temperamental behavior and annoying or even dangerous performance degradations. We all prefer copiers whose copies are clear under low power; we all prefer cars designed to steer safely and predictably, even on roads that are wet or bumpy, in crosswinds, or with tires that are slightly under or overinflated. We say these products are robust. They gain steadfast customer loyalty.

Design engineers take for granted environmental forces degrading performance (rain, low voltage, and the like). They try to counter these effects in product design—insulating wires, adjusting tire treads, sealing joints. But some performance degradations come

Taguchi's Quality Imperatives

□ Quality losses result mainly from product failure after sale; product "robustness" is more a function of product design than of on-line control, however stringent, of manufacturing processes.
□ Robust products deliver a strong "signal" regardless of external "noise" and with a minimum of internal "noise." Any strengthening of a design, that is, any marked increase in the signal-to-noise ratios of component parts, will simultaneously improve the robustness of the product as a whole.
□ To set targets at maximum signal-to-noise ratios, develop a system of trials that allows you to analyze change in overall system performance according to the *average* effect of change in component parts, that is, when you subject parts to varying values, stresses, and experimental conditions. In new products, average effects may be most efficiently discerned by means of "orthogonal arrays."
□ To build robust products, set ideal target values for components and then minimize the average of the square of deviations for combined components, averaged over the various customer-use conditions.
□ Before products go on to manufacturing, tolerances are set. Overall quality loss then increases by the square of deviation from the target value, that is, by the quadratic formula $L = D^2C$, where the constant, C, is determined by the cost of the countermeasure that might be employed in the factory. This is the "Quality Loss Function."
□ You gain virtually nothing in shipping a product that just barely satisfies the corporate standard over a product that just fails. Get on target, don't just try to stay in-spec.
□ Work relentlessly to achieve designs that can be produced consistently; demand consistency from the factory. Catastrophic stack-up is more likely from scattered deviation within specifications than from consistent deviation outside. Where deviation from target is consistent, adjustment to the target is possible.
□ A concerted effort to reduce product failure in the field will simultaneously reduce the number of defectives in the factory. Strive to reduce variances in the components of the product and you will reduce variances in the production system as a whole.
□ Competing proposals for capital equipment or competing proposals for on-line interventions may be compared by adding the cost of each proposal to the average quality loss, that is, the deviations expected from it.

from the interaction of parts themselves, not from anything external happening to them. In an ideal product—in an ideal anything—parts work in perfect harmony. Most real products, unfortunately, contain perturbations of one kind or another, usually the result of a faulty meshing of one component with corresponding components. A drive shaft vibrates and wears out a universal joint prematurely; a fan's motor generates too much heat for a sensitive microprocessor.

Such performance degradations may result either from something going wrong in the factory or from an inherent failure in the design. A drive shaft may vibrate too much because of a misaligned lathe or a misconceived shape; a motor may prove too hot because it was put together improperly or yanked into the design impetuously. Another way of saying this is that work-in-progress may be subjected to wide variations in factory process and ambience, and products may be subjected to wide variations in the conditions of customer use.

Why do we insist that most degradations result from the latter kind of failure, design failures, and not from variations in the factory? Because the ambient or process variations that work-in-process may be subjected to in the factory are not nearly as dramatic as the variations that products are subjected to in a customer's hands—obvious when you think about it, but how many exponents of Zero Defects do? Zero Defects says, The effort to reduce process failure in the factory will simultaneously reduce instances of product failure in the field. We say, The effort to reduce product failure in the field will simultaneously reduce the number of defectives in the factory.

Still, we can learn something interesting about the roots of robustness and the failures of traditional quality control by confronting Zero Defects on its own ground. It is in opposition to Zero Defects that Taguchi Methods emerged.

Robustness as Consistency

According to Zero Defects, designs are essentially fixed before the quality program makes itself felt; serious performance degradations result from the failure of parts to mate and interface just so. When manufacturing processes are out of control, that is, when there are serious variations in the manufacture of parts, products cannot be expected to perform well in the field. Faulty parts make faulty connections. A whole product is the sum of its connections.

Of course, no two drive shafts can be made *perfectly* alike. Engineers working within the logic of Zero Defects presuppose a certain amount of variance in the production of any part. They specify a target for a part's size and dimension, then tolerances that they presume will allow for trivial deviations from this target. What's wrong with a drive shaft that should be 10 centimeters in diameter actually coming in at 9.998?

Nothing. The problem—and it is widespread—comes when managers of Zero Defects programs make a virtue of this necessity. They grow accustomed to thinking about product quality in terms of acceptable deviation from targets—instead of the consistent effort to hit them. Worse, managers may specify tolerances that are much too wide because they assume it would cost too much for the factory to narrow them.

Consider the case of Ford vs. Mazda (then known as Toyo Koygo), which unfolded just a few years ago. Ford owns about 25% of Mazda and asked the Japanese company to build transmissions for a car it was selling in the United States. Both Ford and Mazda were supposed to build to identical specifications; Ford adopted Zero Defects as its standard. Yet after the cars had been on the road for a while, it became clear that Ford's transmissions were generating far higher warranty costs and many more customer complaints about noise.

To its credit, Ford disassembled and carefully measured samples of transmissions made by both companies. At first, Ford engineers thought their gauges were malfunctioning. Ford parts were all in-spec, but Mazda gearboxes betrayed no variability at all from targets. Could *that* be why Mazda incurred lower production, scrap, rework, and warranty costs?[1]

That was precisely the reason. Imagine that in some Ford transmissions, many components near the *outer limits* of specified tolerances—that is, fine by the definitions of Zero Defects—were randomly assembled together. Then, many trivial deviations from the target tended to "stack up." An otherwise trivial variation in one part exacerbated a variation in another. Because of deviations, parts interacted with greater friction than they could withstand individually or with greater vibration than customers were prepared to endure.

Mazda managers worked consistently to bring parts in on target. Intuitively, they took a much more imaginative approach to on-line quality control than Ford managers did; they certainly grasped factory conformance in a way that superseded the pass/fail, in-spec/out-of-spec style of thinking associated with Zero Defects. Mazda managers worked on the assumption that robustness begins from meeting exact targets *consistently*—not from always staying within tolerances. They may not have realized this at the time, but they would have been even better off missing the target with perfect consistency than hitting it haphazardly—a point that is illuminated by this simple analogy.

Sam and John are at the range for target practice. After firing ten shots, they examine their targets. Sam has ten shots in a tight cluster just outside the bull's-eye circle. John, on the other hand, has five shots in the circle, but they are scattered all over it—as many near the perimeter as near dead center—and the rest of his shots are similarly dispersed around it (see the "Who's the Better Shot?" diagram).

Zero Defects theorists would say that John is the superior shooter because his performance betrays no failures. But who would you really rather hire on as a bodyguard?

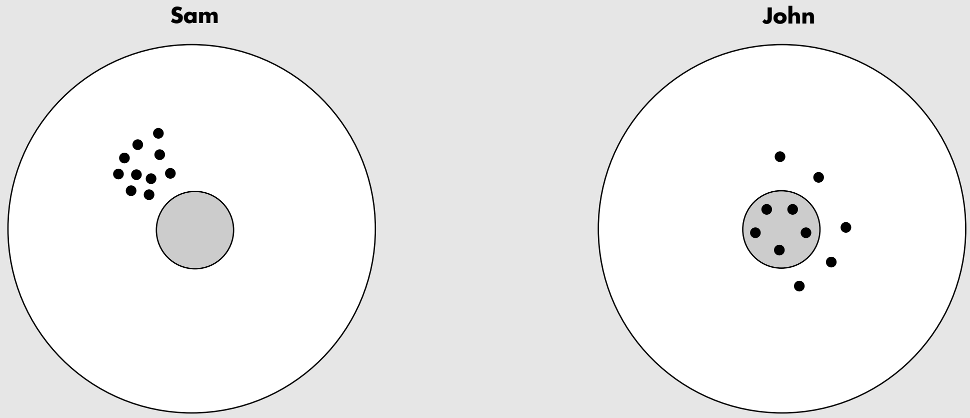
Sam's shooting is consistent and virtually predictable. He probably knows why he missed the circle completely. An adjustment to his sights will give many perfect bull's-eyes during the next round. John has a much more difficult problem. To reduce the dispersion of his shots, he must expose virtually all the factors under his control and find a way to change them in some felicitous combination. He may decide to change the position of his arms, the tightness of his sling, or the sequence of his firing: breathe, aim, slack, and squeeze. He will have little confidence that he will get all his shots in the bull's-eye circle next time around.

When extrapolated to the factory, a Sam-like performance promises greater product robustness. Once consistency is established—no mean feat, the product of relentless attention to the details of design and process both—adjusting performance to target is a simple matter: stack-up can be entirely obviated. If every drive shaft is .005 centimeters out, operators can adjust the position of the cutting tool. In the absence of consistent performance, getting more nearly on target can be terribly time-consuming.

But there is another side to this. There is a much higher probability of catastrophic stack-up from random deviations than from deviations that show consistency. Assuming that no part is grossly defective, a product made from parts that are all off target in exactly the same way is more likely to be robust than a product made from parts whose deviations are in-spec but unpredictable. We have statistical proofs of this, but a moment's reflection should be enough. If all parts are made consistently, the product

[1]See Lance A. Ealey's admirable account of this case in *Quality by Design: Taguchi Methods® and U.S. Industry* (Dearborn, Mich.: ASI Press, 1988), pp. 61–62.

Who's the Better Shot?



Sam is. His shooting is consistent and predictable. A small adjustment in his sights will give him many perfect bull's-eyes in the next round.

will perform in a uniform way for customers and will be more easily perfected in the next version. If all parts are made erratically, some products will be perfect, and some will fall apart.

So the case against Zero Defects begins with this: Robustness derives from consistency. Where deviation is consistent, adjustment to the target is possible; catastrophic stack-up is more likely from scattered deviation within specifications than from consistent deviation outside. This regard for consistency, for being on target, has a fascinating and practical application.

The Quality Loss Function

Analysis of Ford's *overall* losses as compared with Mazda's suggests that when companies deviate from targets, they run an increasingly costly risk of loss. Overall loss is quality loss plus factory loss. The more a manufacturer deviates from targets, the greater its losses.

From our experience, quality loss—the loss that comes after products are shipped—increases at a geometric rate. It can be roughly quantified as the Quality Loss Function (QLF), based on a simple quadratic formula. Loss increases by the square of deviation from the target value, $L = D_2C$, where the constant is determined by the cost of the countermeasure that the factory might use to get on target.

If you know what to do to get on target, then you know what this action costs per unit. If you balk at spending the money, then with every standard deviation from the target, you risk spending more and more. The greater the deviation from targets, the greater the compounded costs.

Let's say a car manufacturer chooses not to spend, say, $20 per transmission to get a gear exactly on target. QLF suggests that the manufacturer would wind up spending (when customers got mad) $80 for two standard deviations from the target ($20 multiplied by the square of two), $180 for three, $320 for four, and so forth.

This is a simple approximation, to be sure, not a law of nature. Actual field data cannot be expected to vindicate QLF precisely, and if your corporation has a more exacting way of tracking the costs of product failure, use it. But the tremendous value of QLF, apart from its bow to common sense, is that it translates the engineer's notion of deviation from targets into a simple cost estimate managers can use. QLF is especially helpful in the important early stages of new product development, when tolerances are set and quality targets are established.

Sony Televisions: Tokyo vs. San Diego

The compelling logic of QLF is best illustrated by the performance of Sony televisions in the late 1970s. The case demonstrates how engineering data and economic data can (and should) be seen in tandem.

Sony product engineers had ascertained that customers preferred pictures with a particular color density, let's call it a nominal density of 10. As color density deviated from 10, viewers became increasingly dissatisfied, so Sony set specification limits at no less than 7 and no more than 13.

Sony manufactured TV sets in two cities, San Diego and Tokyo. Sets shipped from San Diego were uniformly distributed within specs, which meant that a customer was as likely to buy a set with a color density of 12.6 as one with a density of 9.2. At the same time, a San Diego set was as likely to be near the corporate specification limits of 13 or 7 as near the customer satisfaction target of 10. Meanwhile, shipments from Tokyo tended to cluster near the target of 10, though at that time, about 3 out of every 1,000 sets actually fell outside of corporate standards.

Akio Morita, the chairman of Sony, reflected on the discrepancy this way: "When we tell one of our Japanese employees that the measurement of a certain part must be within a tolerance of plus or minus five, for example, he will automatically strive to get that part as close to zero tolerance as possible. When we started our plant in the United States, we found that the workers would follow instructions perfectly. But if we said make it between plus or minus five, they would get it somewhere near plus or minus five all right, but rarely as close to zero as the Japanese workers did."

If Morita were to assign grades to the two factories' performances, he might say that Tokyo had many more As than San Diego, even if it did get a D now and then; 68% of Tokyo's production was in the A range, 28% in the B range, 4% in the C range, and 0.3% in the D range. Of course, San Diego made some out-of-spec sets; but it didn't ship its Fs. Tokyo shipped everything it built without bothering to check them. Should Morita have preferred Tokyo to San Diego?

The answer, remember, must be boiled down to dollars and cents, which is why the conventions of Zero Defects are of no use here. Suppose you bought a TV with a color density of 12.9, while your neighbor bought one with a density of 13.1. If you watch a program on his set, will you be able to detect any color difference between yours and his? Of course not. The color quality does not present a striking problem at the specification limit of 13. Things do not suddenly get more expensive for the San Diego plant if a set goes out at 13.1.

The losses start mounting when customers see sets at the target value of 10. Then, anything much away from 10 will seem unsatisfactory, and customers will demand visits from repairpeople or will demand replacement sets. Instead of spending a few dollars per set to adjust them close to targets, Sony would have to spend much more to make good on the sets—about two-thirds of the San Diego sets—that were actually displeasing customers. (Dissatisfaction certainly increases more between 11.5 and 13 than between 10 and 11.5.)

What Sony discovered is that you gain virtually nothing in shipping a product that just barely satisfies the corporate standard over a product that just fails. San Diego shipped marginal sets "without defects," but their marginal quality proved costly.

Using QLF, Sony might have come up with even more striking figures. Say the company estimated that the cost of the countermeasure required to put every set right—an assembly line countermeasure that puts every set at a virtual 10—was $9. But for every San Diego set with a color density of 13 (three standard deviations from the target), Sony spent not $9 but $81. Total quality loss at San Diego should have been expected to be *three times* the total quality loss at the Tokyo factory.

Deviation: Signal to Noise

If Zero Defects doesn't work, what does? We have said that quality is mainly designed in, not controlled from without. In development work, engineers must discipline their decisions at virtually every step by comparing expected quality loss with known manufacturing cost. On the other hand, the reliability of QLF calculations is pretty obviously restricted by the accuracy of more preliminary measures. It is impossible to discern any loss function properly without first setting targets properly.

How *should* design engineers and manufacturing managers set targets? Let us proceed slowly, reconsidering what engineers do when they test components and subassemblies and how they establish what no particular part "wants to be" in the context of things that get in its way.

When Sony engineers designed their televisions, they assumed that discriminating customers would like a design that retained a good picture or "signal" far from the station, in a lightning storm, when the food processor was in use, and even when the power company was providing low voltage. Customers would be dismayed if the picture degraded every time they turned up the volume. They would reject a TV that developed snow and other annoying "noises" when afflicted by nasty operating conditions, which are themselves considered noises.

In our view, this metaphorical language—signal as compared with noise—can be used to speak of all products, not just televisions. The signal is what the product (or component or subassembly) is trying to deliver. Noises are the interferences that degrade signal, some of them coming from outside, some from complementary systems within the product. They are very much like the factors we spoke of as accounting for variations in product performance—environmental disturbances as well as disturbances engendered by the parts themselves.

And so it seems reasonable to define robustness as the virtue of a product with a high signal-to-noise ratio. Customers resent being told, "You were not expected to use our product in high humidity or in below-freezing temperatures." They want good performance under actual operating conditions—which are often less than perfect. We all assume that a product that performs better under adverse conditions will be that much more durable under normal conditions.

Signal-to-noise ratios are designed into products before the factory ramps up. The strength of a product's signal—hence, its robustness—is primarily the responsibility of the product designers. Good factories are faithful to the intention of the design. But mediocre designs will always result in mediocre products.

Choosing Targets: Orthogonal Arrays

How, then, do product designers maximize signal-to-noise ratios? World-class companies use a three-step decision-making process:

1. They define the specific objective, selecting or developing the most appropriate signal and estimating the concomitant noise.
2. They define feasible options for the critical design values, such as dimensions and electrical characteristics.
3. They select the option that provides the greatest robustness or the greatest signal-to-noise ratio.

This is easier said than done, of course, which is why so many companies in Japan, and now in the United States, have moved to some form of simultaneous engineering. To define and select the correct signals and targets is no mean feat and requires the expertise of all product specialists. Product design, manufacturing, field support, and marketing—all of these should be worked out concurrently by an interfunctional team.

Product designers who have developed a "feel" for the engineering of particular products should take the lead in such teams. They can get away with only a few, limited experiments, where new people would have to perform many more. Progressive companies make an effort to keep their product specialists working on new versions rather than bump them up to management positions. Their compensation schemes reward people for doing what they do best.

But the virtues of teamwork beg the larger question of how to develop an efficient experimental strategy that won't drain corporate resources as you work to bring prototypes up to customer satisfaction. Intuition is not really an answer. Neither is interfunctionality or a theory of organization. Product designers need a scientific way to get at robustness. They have depended too long on art.

The most practical way to go about setting signal-to-noise ratios builds on the work of Sir Ronald Fisher, a British statistician whose brilliant contributions to agriculture are not much studied today. Most important is his strategy for systematic experimentation, including the astonishingly sensible plan known as the "orthogonal array."

Consider the complexity of improving a car's steering. Customers want it to respond consistently. Most engineers know that steering responsiveness depends on many critical design parameters—spring stiffness, shock absorber stiffness, dimensions of the steering and suspension mechanisms, and so on—all of which might be optimized to achieve the greatest possible signal-to-noise ratio.

It makes sense, moreover, to compare the initial design value to both a larger and a smaller value. If spring stiffness currently has a nominal value of 7, engineers may want to try the steering at 9 and at 5. One car engineer we've worked with established that there are actually 13 design variables for steering. If engineers were to compare standard, low, and high values for each critical variable, they would have 1,594,323 design options.

Proceed with intuition? Over a million possible permutations highlight the challenge—that of a blind search for a needle in a haystack—and steering is only one subsystem of the car. In Japan, managers say that engineers "like to fish with one rod"; engineers are optimistic that "the next cast will bring in the big fish"—one more experiment and they'll hit the ideal design. Naturally, repeated failure leads to more casts. The new product, still not robust to customers' conditions, is eventually forced into the marketplace by the pressures of time, money, and diminishing market share.

To complete the optimization of robustness most quickly, the search strategy must derive the maximum amount of information from a few trials. We won't go through the algebra here, but the key is to develop a system of trials that allows product engineers to analyze the *average* effect of change in factor levels under different sets of experimental conditions.

And this is precisely the virtue of the orthogonal array (see the insert, "Orthogonal Arrays: Setting the Right Targets for Design"). It balances the levels of performance demanded by customers against the many variables—or noises—affecting performance. An orthogonal array for 3 steering performance levels—low, medium, and high—can reduce the experimental possibilities to 27. Engineers might subject each of the 27 steering designs to some combination of noises, such as high/low tire pressure, rough/smooth road, high/low temperature. After all of the trials are completed, signal-to-noise values may be used to select the best levels for each design variable.

If, for example, the average value for the first nine trials on spring stiffness is 32.4, then that could characterize level one of spring stiffness. If the average value for the second group of trials is 26.7, and the average for the third group 28.9, then we would select level one as the best value for spring stiffness. This averaging process is repeated to find the best level for each of the 13 design variables.

The orthogonal array is actually a sophisticated "switching system" into which many different design variables and levels of change can be plugged. This system was conceived to let the relatively inexperienced designer extract the *average* effect of each factor on the experimental results, so he or she can reach reliable conclusions despite the large number of changing variables.

Of course, once a product's characteristics are established so that a designer can say with certainty that design values—that is, optimized signal-to-noise ratios—do not interact at all, then orthogonal arrays are superfluous. The designer can instead proceed to test each design variable more or less independently, without concern for creating noise in other parts or subassemblies.

System Verification Test: The Moment of Truth

After they've maximized signal-to-noise ratios and optimized design values, engineers build prototypes. The robustness of the complete product is now verified in the System Verification Test (SVT)—perhaps the most critical event during product development.

In the SVT, the first prototypes are compared with the current benchmark product. Engineers subject both the prototype and the benchmark to the same extreme conditions they may encounter in actual use. Engineers also measure the same critical signal-to-noise ratios for all contenders. It is very important for the new product to surpass the robustness of the benchmark product. If the ideal nominal voltage is 115 volts, we want televisions that will have a signal of 10 even when voltage slips to a noisy 100 or surges to an equally noisy 130. Any deviations from the perfect signal must be considered in terms of QLF, that is, as a serious financial risk.

The robust product, therefore, is the one that minimizes the average of the square of the deviation from the target—averaged over the different customer-use conditions. Suppose you wish to buy a power supply and learn that you can buy one with a standard deviation of one volt. Should you take it? If the mean value of output voltage is 1,000 volts, most people would think that, on average, an error of only one volt is very good. However, if the average output were 24 volts, then a standard deviation of one seems very large. We must always consider the ratio of the mean value divided by the standard deviation.

The SVT gives a very strong indication, long before production begins, of whether customers will perceive the new product as having world-class quality and performance. After the new design is verified to have superior robustness, engineers may proceed to solve routine problems, fully confident that the product will steadily increase customer loyalty.

Back to the Factory

The relationship of field to factory proves to be a subtle one—the converse of what one might expect. We know that if you control for variance in the factory, you reduce failure in the field. But as we said at the outset, a concerted effort to reduce product failure in the field will simultaneously reduce the number of defectives in the factory.

Strive to reduce variances in the components of the product and you will reduce variances in the production system as a whole. Any strengthening of a design—that is, any marked increase of a product's signal-to-noise ratio—will simultaneously reduce a factory's quality losses.

Why should this be so? For many reasons, most importantly the symmetries between design for robustness and design for manufacture. Think of how much more robust products have become since the introduction of molded plastics and solid-state circuitry. Instead of serving up many interconnected wires and tubes and switches—any one of which can fail—engineers can now imprint a million transistors on a virtually indestructible chip. Instead of joining many components together with screws and fasteners, we can now consolidate parts into subassemblies and mount them on molded frames that snap together.

All of these improvements greatly reduce opportunities for noise interfering with signal; they were developed to make products robust. Yet they also

Orthogonal Arrays: Setting the Right Targets for Design

U.S. product engineers typically proceed by the "one factor at a time" method. A group of automotive-steering engineers—having identified 13 critical variables governing steering performance—would begin probing for design improvement by holding all variables at their current values and recording the result. In the second experiment, they would change just *one* of the variables—spring stiffness, say—to a lower or higher value, and if the result is an improvement (skidding at 40 mph, not 35), they will adopt *that* value as a design constant. Then comes the next experiment in which they'll change a different variable but not reconsider spring stiffness. They'll continue in this manner until they have nudged the steering system as close as possible to some ideal performance target (skidding at 50 mph, not 40).

The obvious trouble with such results is that they fail to take account of potentially critical interactions among variables—let alone the real variations in external conditions. While a certain spring stiffness provides ample performance when tire pressure is correct, how well will this stiffness work when tire pressure is too low or too high?

What these engineers need, therefore, is some efficient method to compare performance levels of all steering factors under test—and in a way that separates the *average* effect of spring stiffness at its high, low, and current settings on the various possible steering systems. The engineers could then select the spring-stiffness setting that consistently has the strongest positive effect on the best combination of variables.

If a particular spring-stiffness setting performs well in conjunction with each setting of all 12 other suspension factors, it stands a very good chance of reproducing positive results in the real world. This is an important advantage over the "one factor at a time" approach.

The orthogonal array can be thought of as a distillation mechanism through which the engineer's experimentation passes. Its great power lies in its ability to separate the effect each factor has on the average and the dispersion of the experiment as a whole. By exploiting this ability to sort out individual effects, engineers may track large numbers of factors simultaneously in each experimental run without confusion, thereby obviating the need to perform all possible combinations or to wait for the results of one experiment before proceeding with the next one.

Consider the orthogonal array for steering. Each of the rows in the array shown here constitutes one experiment, and each vertical column represents a single test factor. Column 1, for example, could represent spring stiffness. Engineers test each of the 13 steering factors at three different settings. (For spring stiffness, then, these would be the current setting, a stiffer setting, and a softer setting, notated 1, 2, and 3 in the array.)

Engineers perform 27 experiments on 13 variables, A through M. They run 27 experiments because they want to expose each performance level (1, 2, and 3) to each other performance level an equal number of times (or 3 × 3 × 3). The engineer must perform all 27 experiments, adhering to the arrangement of the factor levels shown here and drawing lots in order to introduce an element of randomness to the experimentation.

have made products infinitely more manufacturable. The principles of designing for robustness are often indistinguishable from the principles of designing for manufacture—reduce the number of parts, consolidate subsystems, integrate the electronics.

A robust product can tolerate greater variations in the production system. Please the customer and you will please the manufacturing manager. Prepare for variances in the field and you will pave the way for reducing variations on the shop floor. None of this means the manufacturing manager should stop trying to reduce process variations or to achieve the same variations with faster, cheaper processes. And there are obvious exceptions proving the rule—chip production, for example, where factory controls are ever more stringent—though it is hard to think of exceptions in products such as cars and consumer electronics.

The factory is a place where workers must try to meet, not deviate from, the nominal targets set for products. It is time to think of the factory as a product with targets of its own. Like a product, the factory may be said to give off an implicit signal—the consistent production of robust products—and to be subject to the disruptions of noise—variable temperatures, degraded machines, dust, and so forth. Using QLF, choices in the factory, like choices for the product, can be reduced to the cost of deviation from targets.

Consider, for instance, that a cylindrical grinder creates a cylindrical shape more consistently than a lathe. Product designers have argued for such dedicated machines; they want the greatest possible precision. Manufacturing engineers have traditionally favored the less precise lathe because it is more flexible and it reduces production cost. Should management favor the more precise cylindrical grinder? How

Factor Levels

			Column												
	No.	A	B	C	D	E	F	G	H	I	J	K	L	M	
	1	1	1	1	1	1	1	1	**1**	1	1	1	1	1	Factor Levels
	2	1	1	1	1	2	2	2	**2**	2	2	2	2	2	for First
	3	1	1	1	1	3	3	3	**3**	3	3	3	3	3	Experiment
	4	1	2	2	2	1	1	1	**2**	2	2	3	3	3	
	5	1	2	2	2	2	2	2	**3**	3	3	1	1	1	
	6	1	2	2	2	3	3	3	**1**	1	1	2	2	2	
	7	1	3	3	3	1	1	1	**3**	3	3	2	2	2	
	8	1	3	3	3	2	2	2	**1**	1	1	3	3	3	
	9	1	3	3	3	3	3	3	**2**	2	2	1	1	1	
Experiments	10	2	1	2	3	1	2	3	**1**	2	3	1	2	3	
	11	2	1	2	3	2	3	1	**2**	3	1	2	3	1	
	12	2	1	2	3	3	1	2	**3**	1	2	3	1	2	
	13	2	2	3	1	1	2	3	**2**	3	1	3	1	2	
	14	2	2	3	1	2	3	1	**3**	1	2	1	2	3	
	15	2	2	3	1	3	1	2	**1**	2	3	2	3	1	
	16	2	3	1	2	1	2	3	**3**	1	2	2	3	1	
	17	2	3	1	2	2	3	1	**1**	2	3	3	1	2	
	18	2	3	1	2	3	1	2	**2**	3	1	1	2	3	
	19	3	1	3	2	1	3	2	**1**	3	2	1	3	2	
	20	3	1	3	2	2	1	3	**2**	1	3	2	1	3	
	21	3	1	3	2	3	2	1	**3**	2	1	3	2	1	
	22	3	2	1	3	1	3	2	**2**	1	3	3	2	1	
	23	3	2	1	3	2	1	3	**3**	2	1	1	3	2	
	24	3	2	1	3	3	2	1	**1**	3	2	2	1	3	
	25	3	3	2	1	1	3	2	**3**	2	1	2	1	3	Factor Levels
	26	3	3	2	1	2	1	3	**1**	3	2	3	2	1	for 27th
	27	3	3	2	1	3	2	1	**2**	1	3	1	3	2	Experiment

Note that the three levels of factor A (1, 2, and 3) each are exposed to the three levels of the other 12 factors an equal number of times. This is true for all factors, A through M. Though not immediately obvious, the three levels of factor H, for example, are exposed equally, each appearing nine times for a total of 27.

It is important to note that each performance value is not exposed to every other possible combination of performance values. In this case, for instance, there is no experiment in which spring stiffness at 2 is exposed to all the other variables at 2; there is no row that shows 2 straight across. But if you look down each column, you will see that each performance level of each variable is exposed to each other variable *an equal number of times.* Before the experiment is completed, all three levels of spring stiffness will be exposed equally to all three levels of tire pressure, steering geometry, and so forth.

The engineer who uses an orthogonal array isn't primarily interested in the specific values yielded by the 27 experiments. By themselves, these values may not produce large improvements over existing performance. Instead, he or she is interested in distilling the effect that each of the three spring-stiffness settings has on the system as a whole. The "essence" of this approach is extracted along the array's vertical rather than horizontal axis.

The array allows the engineer to document results: those factor levels that will reduce performance variation as well as those that will guide the product back to its performance target once consistency has been achieved. The array also may be used to find variables that have little effect on robustness or target values: these can and should be set to their least costly levels.

—Lance A. Ealey

Lance A. Ealey, a member of the consulting staff of McKinsey & Co., is the author of Quality by Design: Taguchi Methods® and U.S. Industry *(ASI Press, 1988).*

do you compare each group's choice with respect to quality loss?

In the absence of QLF's calculation, the most common method for establishing manufacturing tolerances is to have a concurrence meeting. Design engineers sit on one side of the conference room, manufacturing engineers on the opposite side. The product design engineers start chanting "Tighter Tolerance, Tighter Tolerance, Tighter Tolerance," and the manufacturing engineers respond with "Looser Tolerance, Looser Tolerance, Looser Tolerance." Presumably, the factory would opt for a lathe if manufacturing chanted louder and longer. But why follow such an irrational process when product design people and manufacturing people *can* put a dollar value on quality precision?

Management should choose the precision level that minimizes the total cost, production cost plus quality loss—the basics of QLF. Managers can compare the costs of competing factory processes by adding the manufacturing cost and the average quality loss (from expected deviations) of each process. They gain economical precision by evaluating feasible alternative production processes, such as the lathe and cylindrical grinder. What would be the quality loss if the factory used the lathe? Are the savings worth the future losses?

Similar principles may be applied to larger systems. In what may be called "process parameter design," manufacturers can optimize production parameters—spindle speed, depth of cut, feed rate, pressure, temperature—according to an orthogonal array, much like the spring stiffness in a steering mechanism. Each row of the orthogonal array may define a different production trial. In each trial, engineers produce and measure several parts and then use the

On-Line Quality: Shigeo Shingo's Shop Floor

Does tightening tolerances necessarily raise the specter of significantly higher production costs? Not according to Shigeo Shingo, the man who taught production engineering to a generation of Toyota managers. Now over 80, Shingo still actively promotes "Zero Quality Control," by which he aims to eliminate costly inspection processes or reliance on statistical quality control at the shop-floor level.

Shingo advocates the use of low-cost, in-process quality control mechanisms and routines that, in effect, incorporate 100% inspection at the source of quality problems. He argues for checking the causes rather than the effects of operator errors and machine abnormalities. This is achieved not through expensive automated control systems but through foolproofing methods such as *poka-yoke.*

Poka-yoke actually means "mistake proofing"; Shingo resists the idea that employees make errors because they are foolishly incompetent. Shingo believes all human beings have lapses in attention. The system, not the operator, is at fault when such defects occur. The poka-yoke method essentially builds the function of a checklist into an operation so we can never "forget what we have forgotten."

For example, if an operator needs to insert a total of nine screws into a subassembly, modify the container so it releases nine at a time. If a screw remains, the operator knows the operation is not complete. Many poka-yoke ideas are based on familiar foolproofing concepts; the mechanisms are also related in principle to *jidohka,* or autonomation—the concept of low-cost "intelligent machines" that stop automatically when processing is completed or when an abnormality occurs.

Shingo recommends four principles for implementing poka-yoke:

1. *Control upstream, as close to the source of the potential defect as possible.* For example, modify the form of a symmetrical workpiece just slightly to assure correct positioning with a jig or sensor keyed to an asymmetrical characteristic. Also, attach a monitoring device that will detect a material abnormality or an abnormal machine condition and will trigger shutdown before a defect is generated and passed on to the next process.
2. *Establish controls in relation to the severity of the problem.* A simple signal or alarm may be sufficient to check an error that is easily corrected by the operator, but preventing further progress until the error is corrected is even better. For example, a control-board counter counts the number of spot welds performed and operates jig clamps; if one is omitted, the workpiece cannot be removed from the jig until the error is corrected.
3. *Think smart and small.* Strive for the simplest, most efficient, and most economical intervention. Don't overcontrol—if operator errors result from a lack of operations, improve methods before attempting to control the results. Similarly, if the cost of equipment adjustment is high, improve equipment reliability and consider how to simplify adjustment operations before implementing a costly automated-inspection system.
4. *Don't delay improvement by overanalyzing.* Poka-yoke solutions are usually the product of decisiveness and quick action on the shop floor. While design improvements can reduce manufacturing defects in the long run, you can implement many poka-yoke ideas at very low cost within hours of their conception, effectively closing the quality gap until you develop a more robust design.

Developed cooperatively by operators, production engineers, and machine-shop personnel, poka-yoke methods are employed extensively in processing and assembly operations in Japan and represent one of the creative pinnacles of continuous shop-floor improvement. In its most developed state, such improvement activity can support off-line quality engineering efforts by feeding back a continuous flow of data about real, in-process quality problems.

—Connie Dyer

Connie Dyer is senior editor at Productivity Press. She worked with Shigeo Shingo on, among other books, Zero Quality Control: Source Inspection and the Poka-Yoke System *(Productivity Press, 1986).*

data to calculate the signal-to-noise ratio for that trial. In a final step, they establish the best value for each production parameter.

The result? A robust process—one that produces improved uniformity of parts and often enables managers to simultaneously speed up production and reduce cycle time.

How Much Intervention?

Finally, there is the question of how much to intervene *during* production.

Take the most common kind of intervention, on-line checking and adjusting of machinery and process. In the absence of any operator monitoring, parts tend to deviate progressively from the target. Without guidance, different operators have widely varying notions of (1) how often they should check their machines and (2) how big the discrepancy must be before they adjust the process to bring the part value back near the target.

By applying QLF, you can standardize intervention. The cost of checking and adjusting has always been easy to determine; you simply have to figure the cost of downtime. With QLF, managers can also figure the cost of *not* intervening, that is, the dollar value to the company of reduced parts variation.

Let's go back to drive shafts. The checking interval is three, and the best adjustment target is 1/1,000th of an inch. If the measured discrepancy from the target is less than 1/1,000th of an inch, production continues. If the measured discrepancy exceeds this, the process is adjusted back to the target. Does this really enable operators to keep the products near the target in a way that minimizes total cost?

It might be argued that measuring every third shaft is too expensive. Why not every tenth? There is a way to figure this out. Say the cost of intervention is 30 cents, and shafts almost certainly deviate from the target value every fifth or sixth operation. Then, out of every ten produced, at least four bad shafts will go out, and quality losses will mount. If the seventh shaft comes out at two standard deviations, the cost will be \$1.20; if the tenth comes out at three standard deviations, the cost will be \$2.70; and so on. Perhaps the best interval to check is every fourth shaft or every fifth, not every third. If the fourth shaft is only one standard deviation from the target value, intervention is probably not worth the cost.

The point, again, is that these things can and should be calculated. There isn't any reason to be fanatical about quality if you *cannot* justify your fanaticism by QLF. Near the target, production should continue without adjustment; the quality loss is small. Outside the limit, the process should be adjusted before production continues.

This basic approach to intervention can also be applied to preventive maintenance. Excessive preventive maintenance costs too much. Inadequate preventive maintenance will increase quality loss excessively. Optimized preventive maintenance will minimize total cost.

In Japan, it is said that a manager who trades away quality to save a little manufacturing expense is "worse than a thief"—a little harsh, perhaps, but plausible. When a thief steals \$200 from your company, there is no net loss in wealth between the two of you, just an exchange of assets. Decisions that create huge quality losses throw away social productivity, the wealth of society.

QLF's disciplined, quantitative approach to quality builds on and enhances employee involvement activities to improve quality and productivity. Certainly, factory-focused improvement activities do not by and large increase the robustness of a product. They can help realize it, however, by reducing the noise generated by the complex interaction of shop-floor quality factors—operators, operating methods, equipment, and material.

Employees committed to hitting the bull's-eye consistently cast a sharper eye on every feature of the factory environment. When their ingenuity and cost-consciousness are engaged, conditions change dramatically, teams prosper, and valuable data proliferate to support better product and process design. An early, companywide emphasis on robust product design can even reduce development time and smooth the transition to full-scale production.

Too often managers think that quality is the responsibility of only a few quality control people off in a factory corner. It should be evident by now that quality is for everyone, most of all the business's strategists. It is only through the efforts of every employee, from the CEO on down, that quality will become second nature. The most elusive edge in the new global competition is the galvanizing pride of excellence.