In a marketplace starved for practical ways to improve quality, it is only natural that many organizations have feasted on statistical process control (SPC). Use of SPC is critical for eliminating unnecessary variability in a process.

Still, there is a growing awareness that SPC alone cannot lead to the kinds of fundamental improvements needed if American businesses are to make competitive products at competitive prices. That's why companies are augmenting their quality improvement arsenal with design of experiments, a multipurpose tool that can lead to new insights in how to design and manufacture products more effectively. The increased awareness of designed experiments is an encouraging sign, but at the same time there is growing confusion about what it is and when it is useful.

Design of experiments, often abbreviated DOX or DOE, embodies a wide range of efficient, systematic, and flexible experimental protocols all based on testing multiple factors in a single experiment. It is this multifactor approach that makes DOX efficient: testing many factors at the same time means you can get more out of your research and project dollars. Yet it could also lead to confusion, and that's why the factors are systematically tested in predetermined patterns and combinations that allow investigators to keep everything straight. DOX strategies are flexible in that they can be easily adapted to situations ranging from design laboratories to production lines.

The power and the versatility of DOX are main reasons for its rising popularity. In the lab, for example, it enables engineers to progress from "finding something that works" to "making it easy to use and to manufacture." One purpose of using DOX at this stage is to identify product and process designs that will result in high quality products even if the process conditions are less than optimal or components less than perfect. Using DOX also lets companies make products that will work properly even if they are used in extreme environments or a consumer doesn't follow directions exactly.

More generally, the applications of DOX include its use on a production line to identify the process variables that must be carefully controlled (and optimal levels for these variables) and ones that operators need not worry about. This information has led companies to alter the way they operate existing equipment or modify existing procedures, changes that have increased the capability of their production lines. Better still, these increases are inevitably matched by equal improvements in product quality, leading to less rework or scrap. All of these gains lead to a more competitive product and company.

Two fields where DOX is already commonly used are electronic circuit design and chemical manufacturing. In both these areas, there is a clear competitive advantage in being able to use moderate or cheap grades of materials or components to make high-grade final products. Both areas long ago adopted the approach of studying many factors at the same time because they found that was the most effective way to understand what goes on in their production processes.

Does this mean people should think about using DOX everywhere all the time? No. DOX can and should be used more frequently than it is now, but there are times when it will not lead to the kinds of improvements people seek.

Some familiarity with the concepts of DOX will help people decide whether it can help them in their company. The term “design of experiments” is descriptive, but only reflects part of the story. It refers to experiments in the traditional sense of the word—the deliberate manipulation of ingredients, procedures, and conditions in order to test their effects. The key to design of experiments, however, is something not conveyed in the name: many factors are tested at the same time.

To anyone who remembers science classes in school where they were admonished to change just one factor at a time, the idea of testing many factors in one experiment may sound like "bad science." And it would be, if someone were to just fiddle randomly with knobs or materials or ingredients. Such a strategy is unlikely to lead to the discovery of anything important concerning a process or product. Which is why multifactor experiments are designed: careful thought is given to finding efficient ways to extract information from tests of combinations of factors. Groups of designs have been developed over the last 60 years, leading to proven statistical methods for dissecting a response into the effects due to each factor in the combination tested.

In the most common designs, 3 to 15 factors are tested at 2 levels, but the number of factors or levels can exceed this range. The experimenter then selects a set of experimental trials, or runs, that determines how the experiment will be performed.

Each run is one specific combination of the factors and levels being tested. The results of all the runs are compared and examined for patterns that reveal significant information. (See Sidebar 2 on the next page for illustrations and further description of how this works.) Finding the proper patterns and combinations in which to test multiple factors is the foundation of a successful DOX program.

Sometimes, investigators test all the combinations of factors and settings. If a person was studying three factors, each at two levels (or settings), it...
would take eight runs to test all the combinations \(2^3 = 2 \times 2 \times 2 = 8\).

When all possible combinations are tested, the design is called a "full factorial." The knowledge gained from a full factorial is usually more detailed and precise than if the factors were tested one at a time. Still, running full factorials is not always practical or even desirable for two reasons:

First, often times investigators don't want or need the detailed information provided by a full factorial.

Second, when the factors are tested at two levels, the number of runs doubles each time a factor is added to the design (it triples if they are tested at three levels). For example, testing all the combinations of 8 factors, each at two levels, requires 256 runs; for 9 factors, it is 512 runs. By the time there are 15 factors, the number rises to almost 33,000 runs. Clearly this is impossible.

Fortunately, DOX allows people the option of sacrificing some of its discriminatory powers--the ability to separate out portions of the response due to a specific combination of factors -- in exchange for getting general knowledge about a large number of factors. In these cases, only a fraction of the many possible combinations of factors and settings is tested, hence the name "fractional factorials." This is a very effective way for scientists, engineers, and operators to screen anywhere from 7 to 15 factors (or more) in order to determine which ones are likely to have the greatest impact on the process or product being studied.

**DOX in Action**

The power and logic of designed experiments can be better appreciated by seeing them in action.

Microcircuit Engineering Corporation of Mount Holly, New Jersey, manufactures wire mesh screens used in the production of microcircuits. One of their projects concerned the special coatings, or emulsions, applied to the screens by hand, an activity historically regarded within the company more as an art than a science. The coating department was the company bottleneck—it was the only department that required three shifts, and even when running at full capacity it could not keep up with the output of the other departments.

In 1986, Microcircuit heard of a machine that would allow them to automate the coating process while maintaining their high standards of quality. The machine was expensive, so they arranged for a one-month, in-house trial. At first its performance was less than satisfactory even though the company followed the manufacturer's suggested operating procedures. Faced with a short time in which to make their purchasing decision, the company's engineers and technicians turned to DOX with a twofold purpose: they needed to see if they could modify the machine or procedures and make screen coatings of the desired quality.

The engineers knew many factors could influence the quality of the coat. Most of these factors concerned the setup of the trough containing the emulsion, certain characteristics of the emulsion, and conditions under which the machine was operated. Twelve factors were chosen for the experiment - travel rate, height of the trough, opening of the trough, volume of emulsion dropped, dwell time, and so forth. Each was tested at two levels. A set of 16 runs was chosen for the experiments meaning that 16 different combinations of all 12 factors were tested (making this a fractional factorial).
The engineers measured five quality characteristics on each screen made during the experiment.

The first round of tests showed that 4 of the 12 factors had little or no effect on the quality of the coatings. These factors could therefore be operated at their most convenient settings. Another set of runs on the remaining eight factors led to the discovery of a combination of factor settings that optimized the five quality characteristics measured. The optimal settings were different than those recommended by the manufacturer. When operated in this new way, the result was a mechanized coating process four times more accurate and seven times faster than Microcircuit’s manual process. With this knowledge, the company was in a position to invest in the machine as a way to improve quality and eliminate the bottleneck in the coating department. In past times, Microcircuit would have based the purchasing decision on other information, such as what they read in publications received from the manufacturer. While such information is valuable, it says little about the actual capability of that machine.

Because they used DOX, the Microcircuit engineers not only knew the capability of the machine before putting it on-line, they knew the best way to operate the machine to reach this capability. While they may have been able to get some of this information from traditional experiments, it is unlikely they would have gotten as much detail as they did from their designed experiments. Furthermore, choosing a fractional factorial allowed them to make an informed decision within the one-month trial period, a deadline they would never have made had they tried to test the 12 factors one at a time.

Sidebar 2: Factorial Designs

The term "designed experiments" has become synonymous with the statistical concept called "factorial designs." Factorial designs are frequently employed to efficiently study the effect of many factors simultaneously. Complete factorial designs have the following property: Each level of every factor is studied in combination with each level of every other factor. (At times, an investigator can choose not to test all the possible combinations, a situation described in the accompanying text.)

This property is best illustrated by a picture. In figure 2.1 (below), a factorial design for two factors each with two levels is shown. Each open circle represents one run in the experiment. The position of each run describes the factor level settings for that run.

For instance, the upper right hand corner of the square is a run made with both factor 1 and factor 2 at their high levels; the lower right corner is the run where factor 1 was at its high level and factor 2 at its low level.

Here, the results of a 2x2x2 design are plotted at the appropriate corners of a cube – hence its common name “two-cubed” design. With a little practice it is easy to discern patterns when the results are presented this way.

This basic design layout can be extended to three (or more) factors. For three factors, the design is represented by a cube as shown in figure 2.2 (above). In this example, the factors Time, Temperature and Pressure are each studied at two levels. Since there are three factors studied at two levels each, there are eight possible combinations (2 x 2 x 2 = 8) and we need eight runs.

The response for each of the eight runs is entered onto one of the corresponding eight corners of the cube. One run, for example, was done at the lower temperature (150°C), higher pressure (150 psi), and shorter time period (20 min.). Low temperature takes us to the bottom of the cube; high pressure, to the back; and shorter time, to the left side. The figure at the bottom, back, left corner shows us that the yield for that trial was 76%. We could follow a similar strategy for the other runs.
A Multipurpose Tool

The Microcircuit case study illustrates one of the most common uses of DOX: screening many potentially important factors and pinpointing the few that have the strongest influence on a product or process. This information is valuable no matter what the context, because it tells people what they do and don't have to worry about.

In this particular example, DOX was used to evaluate the capability of new equipment, but it works just as well for evaluating machinery or procedures on production lines. Another situation where DOX is useful is during product scale-up, where it can be used to identify factors critical in making the transition from pilot project or prototype to full-scale go smoothly and achieve the desired results. Moving even further upstream, design engineers can find out what aspects of a product design affect its ease of use and manufacture before the product goes into production. The benefits of using DOX vary in each of these different applications, but the design of products and the transfer to the production line typify the main points.

In typical American organizations, design engineers spend months or years creating a series of prototypes.

Figure 2.3

Determining Factor Effects

By comparison, the results (and graphs) for one-factor-at-a-time experiments are less informative. A typical one-factor-at-a-time experiment studying three factors is illustrated in figure 2.4 (below). Here, the results reflect only half of the "cube" because each factor is only tested when the other two factors remain at a given level. This provides no information about how each factor behaves at different levels of the other two, making it impossible to determine if the factors interact.

Though a one-factor experiment could be repeated in order to provide four estimates, the results would only reflect the effects of each factor at one combination of the other factors.

Cube plots therefore graphically summarize the results of an experiment. The main advantage of depicting the results on a cube is that it allows us to visually examine the data for patterns. For example, knowing that the results of all runs using the high temperature are on the top face and all the low temperature readings on the bottom face, we can easily see whether there are any systematic differences between high and low temperatures. We can extend this idea further to look for the main effects of pressure and time, and for patterns in combinations of particular factors.

These cube plots point out another advantage of factorial experiments; they allow the effect of a factor to be examined repeatedly. For example, as diagrammed in figure 2.3 (above), a two-cube experiment provides four independent estimates of the effect of each factor (under four different conditions). In the example above, we calculate the effect of Time. This repetition applies for the other factors, too. Each run in the experiment thus has the versatility to give information about each of the factors.

Cube plots therefore graphically summarize the results of an experiment. The main advantage of depicting the results on a cube is that it allows us to visually examine the data for patterns. For example, knowing that the results of all runs using the high temperature are on the top face and all the low temperature readings on the bottom face, we can easily see whether there are any systematic differences between high and low temperatures. We can extend this idea further to look for the main effects of pressure and time, and for patterns in combinations of particular factors.

These cube plots point out another advantage of factorial experiments; they allow the effect of a factor to be examined repeatedly. For example, as diagrammed in figure 2.3 (above), a two-cube experiment provides four independent estimates of the effect of each factor (under four different conditions). In the example above, we calculate the effect of Time. This repetition applies for the other factors, too. Each run in the experiment thus has the versatility to give information about each of the factors.

Figure 2.3

A One-factor-at-a-time Experiment

A one-factor-at-a-time experiment used to study three factors. It lacks many of the desirable properties of the factorial experiments.
During the earliest design stages, the engineers will perform experiments just so they can find something that works—a one-factor-at-a-time approach. For example, at first an engineer may want to develop a feel for what factors influence the product or process, and he or she will experiment to answer questions such as "what happens if this piece is a little longer" or "what effect will turning this knob have."

After a series of such experiments, the engineer will have enough knowledge to put together a prototype. As depicted in the figure above, subsequent tests on the first prototype typically lead to more refined designs, each of which works a little better than the previous one.

Yet despite this careful work, the performance or reliability of the product drops drastically when it goes into production. Everyone wonders what happened, forgetting that the prototypes were tested in a lab under controlled conditions and probably with special high-grade materials. Out on the floor, conditions and materials never come near these ideals, and the design "message" is garbled because operators lack the equipment and materials needed to duplicate laboratory conditions.

Designed experiments can help solve this problem. They enable engineers to test a wide range of options and see which will work best out on the production floor or at a customer's plant (or home). Traditional experiments are simply too slow and inefficient for this purpose. To illustrate, consider an example from a Japanese copier manufacturer that shows how they approached a similar design situation.

Several years ago this manufacturer decided to make a copier to sell in a variety of international markets. That meant the copier would have to work right in radically different climates and varying conditions. Before this copier went into production, the company engineers took prototype and performed about four dozen designed experiments. The result was a copier that was easy to manufacture, easy to service, and reliable under a wide range of operating conditions (high or low humidity, high or low temperatures, thick or thin paper, and so on).

Because variability in production materials is rarely taken into account during design stages, the quality of a new product usually falls off when prototypes are transferred to the production line.

As increasing numbers of companies are learning, it is much more effective to prevent problems than to fix them. Engineers or others can combine their knowledge about a product with efficient DOX experiments to find products or processes robust enough to withstand variation from any number of sources. The time taken up front pays off in products that are easier to transfer to the production line. Waiting until the product is in production and relying on inspection to catch problems squanders resources on scrap and rework.

Some companies may balk at these suggestions, especially if they are in a volatile market where product life cycles are short. In those cases, elaborate design programs like the one described (it took three years to refine the copier design) are not always possible. Even so, DOX can still help because, even from the very first experiments, they provide useful information. As mentioned previously, there are ways to speed up an experiment program by concentrating more on screening large number of factors rather than going into detail on all of them.

In any case, performing several designed experiments in the product and process design stages usually results in much faster transfer onto the production line. The information gained can direct operators and process engineers towards optimal or near-optimal procedures and settings for that particular design.

These examples and applications illustrate what DOX is and what it can do. But it is just as important to understand what DOX isn't and what it can't do. The two tools commonly confused with or seen as alternatives to DOX are statistical process control and one-factor-at-a-time experiments. In fact, these tools are more complementary than competing; each has its own purpose.
SPC and DOX

Designed experiments and statistical process control (SPC) are both useful and necessary quality improvement tools. Their objectives are quite different however. SPC is a process monitoring tool; it raises a flag when something in the process changes, thereby alerting operators of the need to act. Ongoing, repetitive processes benefit the most from SPC because then it is easy to get enough data to detect patterns. The focus of SPC is downstream, on products or components coming off the lines.

Because it is used to react to events, SPC is often referred to as a passive quality improvement tool; since DOX is used to seek out sources of problems and prevent them from influencing the product it is viewed as an active quality improvement tool. Yet calling SPC “passive” is misleading, for the operators still track down the sources of variation—an activity that usually leads to substantial gains in product uniformity (and thus productivity). In fact, through the conscientious use of SPC, operators can stabilize an unpredictable process and often reach its capability limits as it is currently designed.

But suppose that even when the system operates at its current best there is more variation, less reliability, or lower yield than you would like. Another example from Microcircuit Engineering Corporation illustrates this situation.

An important customer asked Microcircuit to make mesh screens that met very strict tension requirements. The Microcircuit workers knew that the products then coming off the line rarely met this specification - there was just too much variability even when everything went smoothly. They may have been able to eliminate some of the variability with SPC techniques, but could not have changed the production line's inherent capability.

Instead they used DOX to test the influence of 15 factors on the quality of the mesh screens. They discovered that most of the factors had little impact on the screen tension, but one factor, the type of clamp pad, was critical. By changing the type of pad and regularly checking the clamp tension, they were able to meet some of the more exacting requirements put forth by their customer. In fact the overall quality of the screens improved in ways that benefited all of Microcircuit's customers.

There is no doubt that SPC has a vital role to play in stabilizing processes, an accomplishment that leads to more uniform products. Yet it rarely if ever leads to breakthroughs in quality or production capability.

One-at-a-time and DOX

It is also enlightening to compare DOX with the more common type of experimentation. Classroom science courses still stress the one-at-a-time approach to experimentation: you change the one factor you're interested in while holding everything else constant.

Like SPC, such one-factor-at-a-time experiments are useful, particularly in very early design stages where lots of configurations have to be tried just to find a design that works. Sometimes, too, it is helpful to run a quick one-factor experiment when testing a theory or trying to see if a minor adjustment will solve a problem.

Very simply, this is the strategy we all adopt when for example, we have an appliance that doesn't work. We try one thing after another until it works or we give up: checking plugs, turning knobs one at a time, checking other connections, and so on. In these cases, and the analogous "getting something that works" design situation already mentioned, this strategy is not a bad way to go.

However, in most other situations where experimentation is called for, DOX is a more strategic choice because it makes more effective use of time and dollars, and allows you to more easily discover how all the factors interact.

Unfortunately, our classroom training prejudices our viewpoints in ways that are difficult to overcome. In the first place, students are never exposed to the concepts of DOX. What's worse, they are taught that trying to test more than one thing at a time is wrong. It is true that you can't just randomly change settings, procedures, ingredients, dimensions, or fixtures and hope to learn very much about a product or process. And it was to conquer the inadequacy of such random exploration that experimentation was developed. However, the revolutionary insights embodied in DOX offer an even better approach: to manipulate many factors in a few specific, carefully chosen patterns that allow you to get a great deal of information from the results.

In addition, the emphasis on one-factor-at-a-time experiments in the classroom sometimes leads people to expect that there is only one cause for each problem, and that if they can pinpoint that one thing, they will have the solution. In reality, problems are often caused by the interaction of several factors, meaning the effect of one factor depends on the level, or setting, of another factor. When interactions occur, the effect of one factor is either enhanced or limited by the presence of another factor. As explained in Sidebar 3, interactions abound in the world around us: neither sunlight nor water alone does much for the growth of a plant but the two together have a significant effect. It is
difficult to detect interactions using one-factor-at-a-time experiments. The failure to detect interactions has contributed greatly to the failure of many "solutions" in the past.

Even when detecting interactions between factors is not the primary issue - such as when an engineer needs to screen a large number of factors in order to find the one or two most important - DOX is still a better choice of strategy. It allows researchers to test multiple factors much more quickly than one-at-a-time experiments, thereby saving both time and money, and provides greater assurance that the effects they detect are real. This latter statement may appear counterintuitive: one-factor-at-a-time thinking leads us to assume that we would have to sacrifice precision and accuracy to test many variables in the same experiment. This, too, is false.

For an explanation of how DOX can be more cost efficient and more accurate than other experiments, we look to underlying design concepts. In a traditional experiment, each factor is typically tested in only one or two trials. The rest of the time it is held constant while a different factor is tested. By contrast, in designed experiments each factor is tested and monitored in each run. So you can judge the effect of each factor (and each factor combination) based on four or more estimates rather than one or two. (See Sidebar 2.)

Traditional experiments can drag on indefinitely, tying up resources for an interminable length of time. With designed experiments, on the other hand, the total costs are more evident from the start.

Another bias imparted from experience with traditional experimentation is the perception that testing two factors is twice as much work as testing one factor, three factors is three times as much work, and so on. This is true when you test one factor at a time, but is far from true when you test multiple factors simultaneously. One of the biggest benefits of using DOX is the time saved by investigating many factors all at once. For example, the two most common designs involve only 8 and 16 trials, respectively, and allow you to test anywhere from 3 to 15 factors. By comparison, trying to get as much information on three factors the old way would take at least several dozen trials, and the number would quickly reach into the thousands as more factors were added.

Few people appreciate this argument because it is all too easy to underestimate the time and resources spent on traditional experimentation. When a solution is finally found, we tend to think of the costs only in terms of the last experiment/solution tried. But the true expense should include all the solutions that didn't work as well as lost time and lost opportunities. Traditional experiments can drag on indefinitely as a succession of factors are tested, tying up resources for interminable lengths of time. Experimenters often can't predict when they will find a "solution" and therefore have to just plan to keep going until they do. One classic example of this situation is the 26 years it took for Frederick Taylor and his colleagues to determine the best way to get steel of uniform hardness. In his words, "I want to explain why twenty-six years were necessary to carry out these experiments. Time after time we would have to throw away six months' work because eleven of the twelve elements had slipped while we were experimenting with the twelfth. If hard spots appeared in the steel, a whole line of experiments was thrown out and we would have to get a new forging and start all over again. It was the difficulty of that sort of thing, holding eleven elements constant while we were getting the twelfth, which made that problem as difficult as it was."

Had Taylor known about DOX strategies for testing 12 factors, he could have completed the experimental program in a fraction of the time, and would have gotten even more information about how the factors interacted.

In general, designed experiments are easier to plan for than one-factor-at-a-time experiments because the costs are more visible and calculable from the start. With designed experiments, everything is spelled out up-front: the number of runs, the protocol for each run, needed supplies and personnel, and the time needed to perform the experiment. Closer monitoring of the costs of traditional experiments should help in developing more realistic comparisons with designed experiments.

In addition to raising objections to the perceived expense of DOX programs, people often say it sounds impractical (if not impossible) to run an experiment while juggling anywhere from 3 to 15 variables or more. After all, if all these things are changing at once, how do you know which factors cause which effect?

Without wanting to make DOX sound too mysterious or magical, the best answer is that the designs are chosen specifically to allow investigators to apportion the effects of each factor. These patterns have been studied for 60 years, so their properties are understood in some detail.

IF DOX IS SO GREAT...

If, as the above examples may lead you to believe, DOX can work such miracles, then why hasn't it been used more? We think the reasons fall into four categories.

(1) Historical
   Though DOX has been actively studied for over 60 years, it is still the new kid on the block in relation to most scientific methods. The principles of DOX were developed for agricultural use in the 1920s, and were not used extensively outside agriculture until World War II. The need to
simultaneously test many factors is, perhaps, most evident in agriculture: testing a fertilizer under controlled laboratory conditions may tell you something, but it won't tell you how that fertilizer will perform in the real world where plants in the same field are exposed to different moisture conditions, receive varying degrees of sunlight or damage from animals, and so forth. In agricultural research, use of DOX is now standard, and to experiment otherwise is "unthinkable."

Another roadblock to use of DOX is simply that few people know it exists — as mentioned, it isn't taught in the classroom, so where would they hear about it? This is changing, but only slowly.

Because of these factors, use of DOX in this country is patchy. Some large corporations use it extensively; other companies, and particularly small organizations, use it very little if at all.

Sidebar 3: Interactions

Interactions are common in the world around us. For example, fertilizers may have little effect if applied during a dry spell, but a major effect when adequate water is available. On the production line, ingredient A may have little effect at a given temperature, but a much greater effect at a higher or lower temperature.

The only way to discover and understand interactions is by testing a factor at different levels and in various combinations with other factors — which is exactly what designed experiments do.

Let's say we were studying the effect of temperature and catalysts on the yield of a chemical reaction. We test two levels of temperature and two types of catalysts. Instead of going through lengthy statistical analysis, we use software programs to plot the results — two possible outcomes are shown to the right. (A good software package should offer flexible, comprehensive plotting capabilities that display your data in a way that enhances your ability to understand the results of your experiments.)

If the effect of temperature is independent of which catalyst is used (that is, if the temperature and catalysts factors do not interact), we could get a graph similar to the one to the upper right. From this plot it is easy to see that the effect of raising temperature is to increase the yield by ten units for both catalyst A and catalyst B. Also, we can tell that catalyst A is better than catalyst B no matter what the temperature setting.

It could be, however, that there is an interaction between temperature and type of catalyst. In that case, we could get something that looked more like lower right graph: raising the temperature increases the yield by 15 units when catalyst B is used, but only five units when catalyst A is used. In this latter case, if we tested the two catalysts in a traditional experiment, our conclusion about which catalyst is better would have depended on what temperature we used for our tests. (Note: In this graph, the crossing lines reveal the interaction: in other cases with interactions, the lines may diverge or converge depending on how the points were plotted.)
Use of DOX also requires something few companies have attained: communication and openness between people and between departments. To get the information needed to choose the direction and focus of a designed experiment, teams or individuals will have to communicate with people beyond their immediate area.

This is true on many levels. For instance, when designing a product, engineers from different disciplines may need to talk to each other about what factors they think require special attention; the engineers should also talk to people in purchasing to find out what kinds of materials will be available once the product goes into production; they should talk to customers to find out what elements in the design are critical for customer satisfaction.

(3) Technical
Until recently, carrying out designed experiments has been harder than using control charts or other statistical tools. DOX has required some sophistication, or at least access to someone familiar with the concepts.

These barriers are being overcome as more people get trained in DOX concepts and as software and graphical techniques become wider spread. (See sidebar 3.) People no longer need be statistical wizards to use DOX because it is easy to plan the experiments and understand the results without having to know the theory, just as you can drive a car without knowing how it works.

New software packages can also help in choosing the proper design for a given situation. Yet though choosing a proper design isn't the major undertaking it once was, we need to expose more people to the concepts of DOX.

(4) Practical
A primary roadblock preventing designed experiments from becoming more popular is the common misconception that they are lengthy, costly undertakings. In reality, it's rare that the cost of a DOX program is prohibitively expensive either in terms of time or resources. As we have indicated, there are special designs used in "crunch" situations, designs for getting a lot of general information quickly.

Though DOX can, in fact, add some time "up front" in design stages, it more than pays for itself in terms of better quality and productivity once the product reaches the production stage.

In on-line settings, the investment will depend on how well your process works and how long it takes to make a complete product or component. Sometimes a process is so precariously balanced that once it reaches a modicum of performance, managers wouldn't dream of letting anyone change knobs and settings. Yet the very nature of DOX requires just that. And because you must test different combinations of variables, there is a chance of making unusable or unsalable product. However, there are ways to minimize the amount of disruption to production schedules or the affect on the quality of the output.

George Box has led development of a type of designed experiment protocol known widely by its acronym, EVOP (for evolutionary operation). Instead of testing factors at levels placed far apart, the EVOP system changes factor settings incrementally starting from the usual levels. At each phase the results are checked to see if there is a clear indication of where to direct the process next. The researcher constantly looks for patterns in the results that point towards even better settings for the factors.

This system is similar to automatic process controllers, which react to normal deviations in the system, except with EVOP the deviations are deliberately caused in hopes of detecting promising directions for improvement. Since the process is improved in incremental steps, production can continue at near-normal levels. Rarely is unacceptable product made.

In general the size and cost of the disruption depends on the situation and the design chosen. Characteristics that tend to result in inexpensive, brief disruptions are:

- having a process with a quick turnaround time (meaning you could get the results of any trial in minutes, or an hour or two at most);
- testing a relatively small number of factors (say, three or four);
- having a process where the controls, settings, materials, and so forth are easy to adjust;
- having a process that equilibrates rapidly after settings are changed;
- not testing all the possible combinations of the factors being studied, but rather choosing a select few that will give you the most information;
- using the EVOP tactic, which only nudges the process off of known levels; or,
- testing only two levels of each factor (say high/low or large/small).

Alternately, factors leading to expensive and/or time-consuming experiments are:

- having a process with long lead, setup, or turnaround times;
- testing a large number of factors, especially if many of the possible combinations of the factors are tested;
- having a sensitive process that is difficult to adjust or that takes a long time to equilibrate after being adjusted; or
- wanting to test many levels of some factors.
**Where Do I Start?**

The first step is to learn more about design of experiments. If you have in-house statisticians, talk to them about how the costs and benefits may apply in your organization. If you lack in-house expertise, you should seriously consider training people in these concepts. There are some excellent training programs offered through a variety of institutions. Develop your internal resources as much as possible.

Be forewarned that DOX will “feel” different than other strategies tried in the past. You will have to put more thought and planning into experiments up front. You will have to actively seek the opinions of a broad range of people: no single person can determine what factors to test unless they have experience in all parts of the process. For instance, when changing the design of a product, marketing and sales people should be consulted along with customers to determine desirable features; production engineers and line operators should be asked the feasibility of making a product a certain way.

There are inevitably many doubts and reservations that accompany the first use of DOX. But once people learn more about the long-term benefits of DOX, these reservations disappear. The uses of this multifactor approach, as indicated in the examples above, include finding ways to make products that fit customer defined specifications from a process that previously couldn’t come close, making more reliable products (and ultimately decreasing warranty costs), reducing rework and scrap by finding optimal process conditions, testing equipment to see if it can meet your needs (and also finding the best way to use this equipment), and creating product designs that are easy to manufacture and easier for customers to use.

The new DOX enthusiasm has only caught on in a few companies in the United States, though it is widespread in Japan. In the early 1950s, the Japanese, who were desperate to make rapid advances, listened to Dr. Deming and have reaped the rewards. Companies in this country therefore face a major challenge. It will take much time and dedication just to catch up with our foreign competitors, and even more to regain a competitive lead. It will also take a lot more than just using design of experiments, but DOX must be a part of any strategy aimed at long-term success.

**Suggested Readings**

For more information about design of experiments and quality improvement see:


**Acknowledgments**

We have been privileged to learn about design of experiments from some of the best people in the field--George Box, William G. Hunter, and Brian L. Joiner-and would like to acknowledge their contributions.