Taguchi or DOE?
by John Cesarone, Ph.D., P.E.

Most engineers have heard of Design of Experiments (DOE) and of Taguchi Methods, but how many of us can really say we understand the difference between them? Or that we can correctly decide when to use which technique? In this article, we will examine the relative strengths and weaknesses of each approach, and develop some guidelines for selecting the best approach for solving our specific problems.

First of all, what are these techniques? In a general sense, they can both be thought of as techniques for optimizing some process which has controllable inputs and measurable outputs. In a manufacturing situation, the inputs might be settings in some production process, such as temperature of a heat treatment furnace or speeds and feeds on a milling machine. The outputs are generally quality or productivity oriented, such as process yield or units produced per hour by our production line. We might be trying to maximize some output (as in throughput or process yield) or minimize some other output (as in failure rate or scrap).

In a design situation, the inputs might be design decisions, and the outputs would then be performance oriented metrics. For example, inputs might be the number of supports in a structural design, the type of material to be used, or a qualitative decision such as a drum clutch vs. a face clutch. In these cases, outputs to be measured might be load carried, torque transferred, etc. In either type of analysis, production scenario or design situation, we are making decisions on how to do something that will affect what we get as an output.

What do DOE and Taguchi have in common? Besides the inputs and outputs described above, they both deal with multiple inputs. That is, we might have two, three, five, or a dozen or more input decisions to make, all affecting some measurable output. It would be nice if we could experiment with these inputs one at a time, optimizing our output for each input in turn, until we've selected ideal values for all input parameters. Unfortunately, this doesn't usually work, because the inputs generally interact with each other to some extent. For example, imagine that you are running a carburization process. You cannot set furnace temperature to optimize yield while ignoring carbon potential, and then experiment with carbon potentials after fixing the temperature; they interact with each other and affect the levels of output together. What DOE and Taguchi primarily have in common, then, is that they deal with multiple inputs and how they interact with each other.

How do DOE and Taguchi differ? We will get into this soon, but the primary difference lies in how they handle the interactions between inputs. When you remember that DOE was invented by scientists for scientists, and Taguchi methods were invented by engineers for engineers, the differences begin to make sense. Let's look at them each in turn.

DESIGN OF EXPERIMENTS: HOW IT WORKS

The main thing to know about DOE is that it was developed primarily within the scary world of statistics. Okay, come back out from under your desk; we won't dwell on that part. Just remember that the theory behind the technique comes from the classical world of pure math. Using it, however, requires only that small amount of math that you probably remember from your college days.
DOE theory starts with the assumption that all inputs might be interacting with all other inputs. This is a powerful statement. The technique makes no assumptions about some inputs being independent, and therefore can handle any interactions that might be lurking somewhere in your process. When you have no idea what interactions you need to be worrying about, DOE might be the choice for you. Of course this power comes at a price, and that price is lots of experimental runs and lots of calculations.

One of the first applications of DOE was in ancient agricultural sciences. Early farming experimenters were starting to understand things like irrigation, fertilization, crop rotation, etc. These are multiple inputs. You can also see how they interact with each other: What is the best irrigation method? Well, that might depend on your fertilization approach. What is the best fertilization process? Well, that could easily depend on your irrigation techniques, and what crop you grew in that field last year. Since any and all inputs could interact with all other inputs, a technique was needed which would model all of these inputs, and how they all relate to each other. Thus DOE was born.

Another peculiarity of these early agricultural experiments was that they wanted to get it right the first time. One experimental run took an entire growing season; a whole year! You absolutely did not want to collect a year's worth of data, stroke your beard a few times, and then do Phase 2 the next year before you had an answer; your village could starve to death before you were finished. This leads to another characteristic of the DOE approach: not only are all interactions studied, but they are all studied at the same time in one big round of tests.

These considerations can lead us to an assessment of the strengths and weaknesses of the DOE approach. The strength is that we can investigate all possible interactions between inputs at the same time; we don't need any innate knowledge of how the process works. The weakness is that we have no way to make use of any a priori process knowledge that we might happen to have; there is no way to make the experiment more efficient by thinking about how the inputs really do interact. If you think about it, the strength and the weakness are really the same thing!

Other areas, besides agriculture, where DOE makes perfect sense are any complex sciences with many highly coupled inputs where practitioners have little innate understanding of the fundamental processes involved. Biology, virology, and meteorology come to mind. In all of these fields, DOE is a powerful and logical method to optimize process and predict outcomes.

To be fair, not all DOE-based investigations look at all possible interactions. Those that do are called "full factorial" DOEs. "Fractional factorial" DOEs can eliminate some interactions, and therefore slim down the amount of work that needs to be done. But they are still based on the idea of full modeling, and then whittled down to improve efficiency. The savings are generally fairly meager, such as a factor of two or four, and there is still no way to inject understanding of the fundamental process into the mix.

The details of performing a DOE can be found in many textbooks, and we won't duplicate them here. But in a nutshell, in a full factorial, all-interaction DOE, tests are performed for all possible combinations of all inputs. If you have three inputs, each with two possible settings (known as levels), you would need to perform eight tests (that is, two raised to the third power). You can see how the number of tests can get really large really fast. Then, the outputs are averaged for all tests where a particular input was set to a particular level, and compared to the average output for all tests where that particular input was set to its other levels. This comparison gives an insight into the overall affect of that input on the output. Similar calculations can show the
affects of the interactions between the inputs. After a bit of number crunching, many useful chunks of knowledge can be derived on how the inputs interact and how they affect the process output.

**TAGUCHI METHODS: WHAT ARE THEY?**

In our section on DOE, we started by looking at its origins in the world of statistics and agricultural science. Likewise, to understand Taguchi methods, its helps to realize that it came from the world of design engineering. Taguchi methods start with the assumption that we are designing an engineering system, either a product to perform some intended function, or a production process to manufacture some product or item. Since we are knowledgeable enough to be designing the system in the first place, we generally will have some understanding of the fundamental processes inherent in that system. We have some idea of the theory behind an internal combustion engine, or behind a metal removal process or an injection molding process. This is a completely different situation than the agricultural or biological systems discussed above, where we have no idea what's going on behind the scenes. How do we use this knowledge to our advantage?

Basically, we use this knowledge to make our experiments more efficient. We can skip all that extra effort that might have gone into investigating interactions that we know do not exist. Without going into the details, it has been shown that this can decrease our level of effort by a factor of ten or twenty or more. Sometimes much more.

Another distinction in Taguchi methods is the recognition that there are variables that are under our control, and variables that are not under our control. In Taguchi terms, these are called Control Factors and Noise Factors, respectively. An early step in a Taguchi analysis is to use our understanding of the fundamental process to try to identify all of the important Control Factors and Noise Factors. In our heat treatment example above, Control Factors would include temperature, carbon potential, time in the furnace, part layout and orientation, etc. Noise Factors might include the inherent variability in the furnace temperature distribution, errors in timing, outside air temperatures and humidities, etc. We would want to ensure that all important Control Factors and Noise Factors are represented in our experiments.

In our full factorial DOE approach, our experimental plan was quite simple: we tested all possible combinations of input values. But what do we do in a Taguchi analysis? We will test a very small subset of all possible combinations, but which combinations do we use? The answer to this question is one of the most powerful parts of the Taguchi method. Again, without going into too much detail, a variety of experimental plans have been published for various numbers of Control Factors and levels. You don't have to design them for yourself; they are readily available in many textbooks. Some examples include:

<table>
<thead>
<tr>
<th>Taguchi Array</th>
<th>Number of Factors</th>
<th>Number of Levels</th>
<th>Number of Tests Required</th>
<th>Tests in an equivalent Full Factorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>L8</td>
<td>7</td>
<td>2</td>
<td>8</td>
<td>128</td>
</tr>
<tr>
<td>L12</td>
<td>11</td>
<td>2</td>
<td>12</td>
<td>2048</td>
</tr>
<tr>
<td>L16</td>
<td>15</td>
<td>2</td>
<td>15</td>
<td>32,768</td>
</tr>
<tr>
<td>L27</td>
<td>13</td>
<td>3</td>
<td>27</td>
<td>1,594,323</td>
</tr>
</tbody>
</table>
In this chart, the L4 array, for example, is an experimental layout designed for three control factors, each at two levels. This might be two different materials for some new product, two manufacturing methods, and two design variations. We might be interested in how these factors affect power consumption by our new product. A full factorial DOE of all combinations of these factors would require eight tests, two to the third power. The Taguchi L4 layout recommends a subset of four of these eight tests which allow the main effects to be determined. Similarly, the L27 layout allows thirteen factors, each at three levels, to be tested with a mere 27 tests, rather than the somewhat prohibitive 1,594,323 tests required by a full factorial set of tests. The "L", by the way, stands for Latin, since most of these designs are based on classical Latin Square layouts, appropriately pruned for efficiency.

Another major difference between DOE and Taguchi is the use of our Noise Factors. In a traditional DOE, each combination of inputs is tested once. In a Taguchi test, each combination which is tested (remember, we are only testing certain combinations) is tested several times. Each of these replications is different, however, in that we use different levels of our Noise Factors. For example, we might test a particular set of inputs at high temperature and high humidity, high temperature and low humidity, low temperature and high humidity, and low temperature and low humidity. Even though we don't really control temperature and humidity in a production setting, we want to vary it in our tests to inject maximum variability into our experimental procedures. This lets us determine not only which combinations of inputs give us maximum output, but which gives us the most repeatable output, which is often even more important. Taguchi calls this quality robustness, or insensitivity to noise.

Mathematical analysis of results is similar in a Taguchi test to the analysis of a DOE. We can still calculate average effect of each of our inputs (a.k.a., Control Factors), and the effects of some of our interactions. However, we can also calculate effect on robustness of each of our inputs. Taguchi recommended a signal to noise ratio to represent robustness, but a simple variance or standard deviation will work just as well and is more familiar to most engineers. Likewise, we can calculate average affects on output and on robustness for each of our Noise Factors. For details of these calculations, see any elementary textbook on Taguchi analysis.

One final difference from DOE remains. Taguchi recommends a final confirming experiment be performed. Remember that in our full factorial DOE, we tested all possible combinations of input values. After our analysis, we determined which combination gave us the best output, and that led to our decision of what to use in the future. Since we tested all combinations, the "winning combination" was necessarily one of our tests, and we had direct confirmation that it was, indeed, the best. However, in a Taguchi analysis, we only performed a small fraction of all possible input combinations. Our mathematical analysis lead us to select a certain combination of input decisions as "optimal", even though we might not have tested that particular combination. Indeed, if we performed an L16 test, we only did .046% (15 out of 32,768) of the possible input combinations; chances are the recommended combination was not one of our actual tests. So, just to make sure, it is always a good idea to do one more test, using the combination of inputs (that is, Control Factor levels) that our analysis recommended. If it does indeed perform as predicted, we know we have a good solution. Is it really, truly, the best possible combination? We'll never really know, but odds are it is pretty darn good, and our confirming experiment shows us just how good it is.

You can see why this confirming test makes more sense in engineering situations than in, say, agricultural situations. You don't want to run one final crop growing experiment, wasting another entire year. However, you can always hold back one engineering prototype from your
initial round of experiments, and run one more test using your recommended inputs, after your analysis is complete, just to make sure.

COMPARISON OF APPROACHES:

Just to recap, let's look at the major characteristics of our two methods side by side:

**Process Knowledge:**
- DOE assumes no understanding of the fundamental mechanisms governing the process we are investigating.
- Taguchi assumes we have a certain understanding of the process and the interactions that are likely to exist between inputs.

**Combinations of Inputs Tested:**
- DOE tests all combinations of input levels, or some symmetrical subset (such as one half or one fourth).
- Taguchi tests a small fraction of all possible combinations, but in a manner that allows us to calculate the affects of all inputs on the output.

**Noise Factors:**
- DOE traditionally ignores Noise Factors, although they could be added to the experimental plan if desired.
- Taguchi makes use of Noise Factors to test robustness of the system and find optimal inputs.

**Understanding of Variability:**
- DOE ignores variability in the process; it assumes a deterministic nature to the system, and finds combinations of input variables that maximize or minimize output, as the case may be.
- Taguchi assumes a stochastic nature to the system; it looks at both the levels of output and the variability of the output; it lets us select levels of input variables to maximize or minimize output or to minimize variability of output (i.e., maximize robustness).

**Confirming Experiment:**
- DOE requires no confirming experiment, since all combinations of inputs were tested (in a full factorial). In effect, the confirming experiment was taken care of in the original experimental plan.
- Taguchi recommends a confirming experiment just to make sure, since the winning set of inputs was probably not part of the original experimental plan.

WHICH APPROACH SHOULD YOU USE?

Hopefully, this article will have given you enough insights into the similarities and differences between these two powerful techniques for you to decide on a case by case basis which should be used, and when. In general, however, think DOE when you have no idea about the fundamental mechanisms governing your process, when you have no idea about interactions between your inputs, and when experiments take so long that you absolutely must get it right the first time. Think Taguchi when you have a fairly firm grasp of the underlying processes and are just trying to optimize an additional notch, when robustness or consistency of output is just as important as maximizing (or minimizing) your output, when you cannot afford to test all
possible combinations of inputs, and when you have the luxury of going back and doing a final confirming test later. Finally, remember that these two techniques are really only two ends of a spectrum of possible optimization methods, and you are free to adapt elements of each to your own specific needs and situations as circumstances warrant. As always, try to understand the principles, and then do what makes sense.

**ADDITIONAL READING:**


