

Confidence - Or So You Think

By Robert A. Dovich, ASQ, Fellow, CQE, CRE, Quality Manager, Field Fastener Supply Company

Every once in a while, just when one thinks that all bases are covered, It is time to sit back and remember the saying: "Confidence: The feeling you have before you understand the situation." I bring that point up because I recently had an experiment go "down the tube" when I was sure I had all bases covered. So as a reminder to myself, as well as to those of you performing Design of Experiments (DOEs), I will provide a step-by-step outline for performing experiments. As my guide, I used *Design of Experiments*, by Anderson and McLean, published by Marcel Dekker, N.Y.C., 1974.

The first step is simple: Be aware that there is a problem. In many situations, the problem is chronic in nature, and personnel have become accustomed to the existing level of performance and have accepted it as status quo. These chronic problems lend themselves to DOE techniques in many cases, especially when the process is inherently stable at these levels. While not necessary, stable processes are easier to experiment with than unstable processes.

Next, try to define the problem. This is generally done in one or more brainstorming sessions with persons who are knowledgeable about the process under consideration. Someone once said that in almost every brainstorming situation, someone at the table has the information that will lead to the solution of the problem. Unfortunately, we usually don't know who that is, and everyone is usually confident that she or he has the answer.

Brainstorming sessions may lead to dozens of perceived factors that have an effect on the outcome. It is the task of the session leader to ensure that all possible factors are discussed, and help the group focus on the main potential causes of the problem. The more factors used in an experiment, the more complex an experiment becomes. I have read of studies containing over 140 factors, though this is rare. In many cases, experiments contain from 2 to 16 factors. It is sometimes enlightening to find out how much, or how little, the process experts agree on these factors.

The third step is to get a consensus on what the potential causes are and the levels to be tested. For example, in a plating experiment, what are the working levels of a given chemical concentration? If milling Inconel parts is the process, what are the lowest and highest speeds and feed rates of the cutting tool? In grinding wheels, what are the limits for grain size or bonding agents to be used in the experiment? The levels here are to be realistic, not necessarily historic. The goal, or end result, of this experiment is an optimum process, and all practical levels should be given consideration.

How will the outcome be interpreted? If the experiment is performed on a single piece of equipment, it may be unreasonable to extrapolate the results across all similar equipment in the company. To broaden the scope and applicability; of the experiment and the results, it may be necessary to perform the same operation on several pieces of equipment. It is not uncommon for various pieces of equipment to respond differently even if they are the same model, produced at nearly the same time and maintained under similar conditions. If you intend to make a statement about a large cross section of equipment, it makes sense to draw samples from all, or part, of this inference space.

The next steps: (1) design how the experimental units will have the treatment (factor levels) assigned to them; and (2) analyze the effect this randomization will have on the statistical model. The term "design" in Design of Experiments is defining the restrictions we place on the randomization of test runs. For example, if we wish to test the difference in surface finish produced by five drill types, and each lot of raw material is only sufficient to drill five holes, we would not want to drill five holes in Lot 1 with drill No. 1, five holes in Lot 2 with drill No.2, five holes in Lot 3 with drill No.3, and so on. Once the experiment was completed, we would not know if any differences in surface finish were due to differences in drill types or material lots. In this case, we would say that the effects of the drill are confounded with the effects of the different lots of raw material. We would not be able to make an intelligent statement about the differences in performance of the drills, which was the goal of the experiment. Then the comment comes from someone, "I knew that this DOE stuff would never work."

A better way to design the experiment would be to drill one hole per drill in each lot in a random manner and analyze the results. In this case, the differences in raw-material lots are minimized. The point here is to recognize the restriction on randomization so the proper model can be used. We may find out in some experiments that the model resulting from restrictions in randomization will not be suitable for the task at hand and changes may need to be made. This analysis, of course, should take place prior to conducting the experiment.

One of the more neglected aspects of performing a designed experiment is data collection. This is unfortunate, as the most efficiently designed experiment may produce meaningless results if the data is improperly collected (I know this from experience). Planning for this may be extensive and meetings of the original brainstorming team may be required. It may be necessary to

design sophisticated data-collection sheets that describe the order of performing the treatment runs, the setup for the run, and a place to record the output of the experiment, be it surface finish, tensile strength, purity, etc. The methods must be clear to everyone involved.

If the experiment is going to be run over several shifts, it may be a good idea to perform a trial data collection experiment that includes everyone that will be involved with the full-blown experiment. It could be an interesting experience. Sometimes it is discovered that the experiment, as designed, is simply not workable, and a redesign is necessary.

Another interesting case is one in which the operator believes that she or he knows more than the statistical person designing the experiment. An experiment that tests three factors at two levels each requires (when it is a full factorial) a total of eight experimental runs. With a minus sign indicating the low level of a factor (such as low cutter rpm), an experiment may be set up as follows, with the sequence of performing the experiment shown in the "Run Order" column.

The first test showed Factor A at the low level, Factor B at the low level and Factor C at the high level. The second test indicated Factor A at the high level, Factor B at the low level, and Factor C at the low level. The run order was randomly generated to block out the effects of extraneous, or unknown and uncontrolled, factors.

Factor	A	B	C	Run Order
	-	-	-	5
	+	-	-	2
	-	+	-	6
	+	+	-	8
	-	-	+	1
	+	-	+	7
	-	+	+	4
	+	+	+	3

The operator looks at the order of tests and can't believe the ignorance of that person from the office. Factor C takes time to change and stabilize, and time will be lost if the experiment is performed as designed. The operator, therefore, leaves Factor C at the high level and runs test Test 7 next, followed by Tests 4 and 3. The operator then changes Factor C to the low level and runs Tests 5, 2, 6 and 8. To avoid political problems, the operator keeps this to himself and reports the results. Unless the experimenter is aware of this, the design will be analyzed as if it were randomly run and the results could be disastrous. For this reason, the data-collection phase must be well planned and explained to all participants, along with a trial run, if necessary.

Once the data is collected, it is time for the statistical person to really get involved and perform the analysis. The job of the statistician is to explain the results in a manner that is understandable to all the decision makers. If the statistician is unable to make the results clear, the correct decision may not be made due to misinterpretation or ignorance of the meaning of the terminology. Graphical representation will be important when reporting the results.

Based on the report, a decision must be made. Have any of the factor effects showed significance? What is the gain made by implementing the change? If changes are made, the gains should result in some return on investment. If possible, the change should be made in a pilot study to prove out the results found in the experiment. Use caution in implementing the change. If it is felt that more gains can be made, it may be possible to continue the study using the method of Evolutionary Operation (EVOP). EVOP is a method of continuing production and making small changes to the process in a scientific manner to continue to increase yield.

If the results show that none of the factors under study affected the output, have we failed? No. If the factors tested do not show significance (assuming nothing is masking effects), we may be able to save money by running these factors at their most economical levels while searching for those factors that do affect process output. The key to experimental design is to understand the model and the restrictions placed on randomization, and plan and follow-up. Be involved in every step or the experimental results may be meaningless.