

Eight Keys To Successful DOE

DOE provides an efficient path to improvement for those who know how to use it.

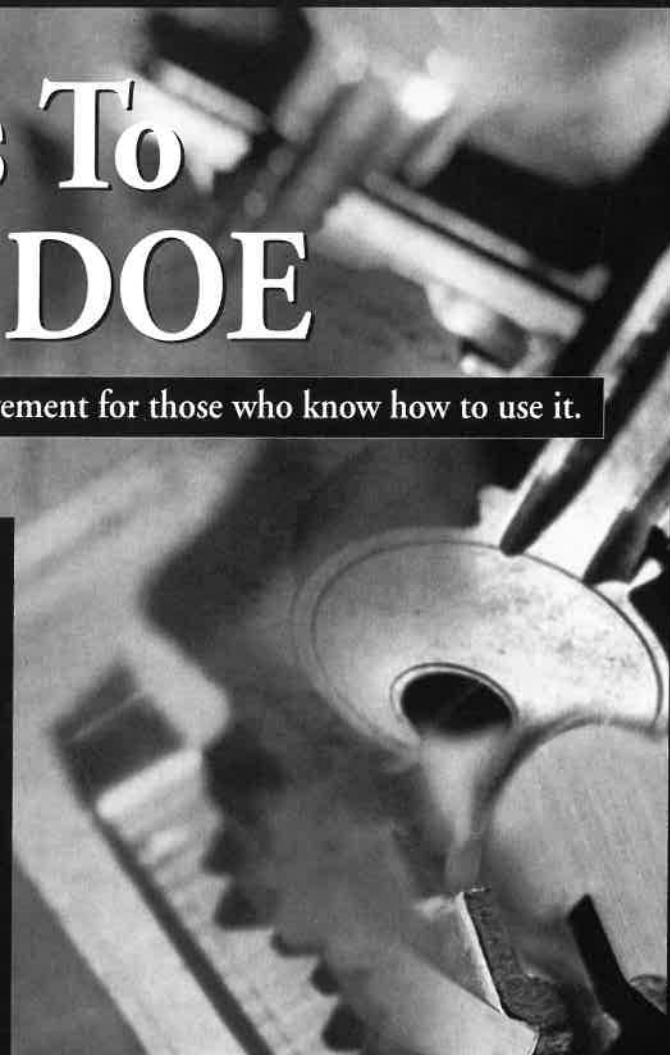
Quality managers who understand how to apply statistical tools for design of experiments (DOE) are better able to support use of DOE in their organizations. Ultimately, this can lead to breakthrough improvements in product quality and process efficiency.

DOE provides a cost-effective means for solving problems and developing new processes. The simplest, but most powerful, DOE tool is two-level factorial design, where each input variable is varied at high (+) and low (-) levels and the output observed for resultant changes. Statistics can then help determine which inputs have the greatest effect on outputs. For example, Figure 1 shows the results for a full two-level design on three factors affecting bearing life. Note the large increase at the rear upper right corner of the cube.

In this example, two factors, heat and cage, interact to produce an unexpected breakthrough in product quality. One-factor-at-a-time (OFAT) experimentation will never reveal such interactions. Two-level factorials, such as the one used in Figure 1, are much more efficient than OFAT because they make use of multivariate design. It's simply a matter of parallel processing (factorial design) vs. serial processing (OFAT). Furthermore, two-level factorials don't require you to run the full number of two-level combinations ($2^{\# \text{ of factors}}$), particularly when you get to five or more factors. By making use of fractional designs, the two-level approach can be extended to many factors without the cost of hundreds of runs. Therefore, these DOEs are ideal for screening many factors to identify the vital few that significantly affect your response.

Such improvements will obviously lead to increased market share and profit. So why don't more manufacturers use DOE? In some cases, it's simple ignorance, but even when companies provide proper training, experimenters resist DOE because it requires planning, discipline and the use of statistics. Fear of statistics is widespread, even among highly educated scientists and managers. Quality professionals can play a big role in helping their colleagues overcome their reluctance.

Using DOE successfully depends on understanding eight fundamental concepts. To illustrate these keys to success, we'll look at a typical example: reducing shrinkage of plastic parts from an injection molding process. The molding case will demonstrate the use of fractional two-level design.



Mary Orlonca

1. Set good objectives

Before you can design an experiment, you must define its objective. The focus of the study may be to screen out the factors that aren't critical to the process, or it may be to optimize a few critical factors. A well-defined objective leads the experimenter to the correct DOE.

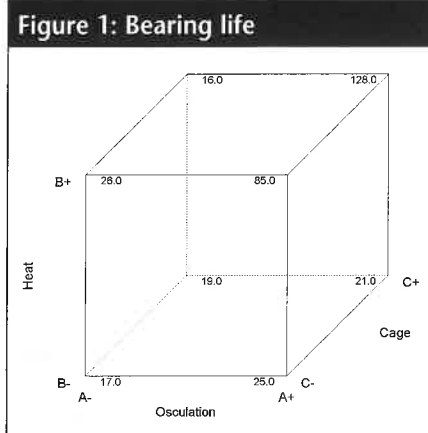
In the initial stage of process development or troubleshooting, the appropriate design choice is a fractional two-level factorial. This DOE screens a large number of factors in a minimal number of runs. However, if the process is already close to optimum conditions, then a response surface design may be most appropriate. It will explore a few factors over many levels.

If you don't identify the objectives of a study, you may pay the consequences—trying to study too many or too few factors, not measuring the correct responses, or arriving at conclusions that are already known.

Vague objectives lead to lost time and money, as well as frustration, for all involved. Identifying the objective upfront builds a common understanding of the project and expectations for the outcome.

In our case study of the injection molder, management wants to reduce variation in parts shrinkage. If the shrinkage can be stabilized, then mold dimensions can be adjusted so the parts can be made consistently.

Factor Name	Units	Low Level	High Level
A: Mold Temperature	Degrees F	130	180
B: Holding Pressure	PSIG	1200	1500
C: Booster Pressure	PSIG	1500	1800
D: Moisture	Percent	0.05	0.15
E: Screw Speed	Inches/sec	1.5	4.0
F: Cycle Time	Seconds	25	30
G: Gate Size	Mils	30	50



The factors and levels to be studied are as shown in Table 1.

The experimenters have chosen a two-level factorial design with 32 runs. A full set of combinations would require 128 runs (2^7), so this represents a 1/4th fraction.

2. Measure responses quantitatively

Many DOEs fail because their responses can't be measured quantitatively. A classic example is found with visual inspections for quality. Traditionally, process operators or inspectors use a qualitative system to determine whether a product passes or fails. At best, they may have boundary samples of minimally acceptable product. Although this system may be OK for production, it isn't precise enough for a good DOE. Pass/fail data can be used in DOE, but to do so is very inefficient. For example, if your process typically produces a 0.1 percent defect rate, you would expect to find five out of 5,000 parts defective. In order to execute a simple designed experiment that investigated three factors in eight experimental runs on such a process, you would need to utilize a minimum of 40,000 parts ($8 \times 5,000$). This would assure getting enough defects to judge improvement, but at an exorbitant cost.

For the purposes of experimentation,

a rating scale works well. Even crude scaling, from 1 to 5, is far better than using the simple pass/fail method. Define the scale by providing benchmarks in the form of defective units or pictures. Train three to five people to use the scale. During the experiment, each trained inspector should rate each unit. Some inspectors may tend to rate somewhat high or low, but this bias can be removed in the analysis via blocking (see Key 5, "Block out known sources of variation"). For a good DOE, the testing method must consistently produce reliable results.

In the case study on injection molding, the experimenters will measure percent shrinkage at a critical dimension on the part. This is a quantitative measurement requiring a great deal of precision. The short-term variation in parts can be dampened by measuring several and in-putting the average. Other responses could be included, such as counts of blemishes and ratings of other imperfections in surface quality.

3. Replicate to dampen uncontrollable variation (noise)

The more times you replicate a given set of conditions, the more precisely you can estimate the response. Replication improves the chance of detecting a statistically significant effect (the signal) in the midst of natural process variation (the noise). In some processes, the noise drowns out the signal. Before you do a DOE, it helps to assess the signal-to-noise ratio. Then you can determine how many runs will be required for the DOE. You first must decide how much of a signal you want to be able to detect. Then you must estimate the noise. This can be determined from control charts, process capability studies, analysis of variance (ANOVA) from prior DOEs or a best guess based on experience.

The statisticians who developed two-

level factorial designs incorporated "hidden" replication within the test matrixes. The level of replication is a direct function of the size of the DOE. You can use the data in Table 2 to determine how many two-level factorial runs you need in order to provide a 90-percent probability of detecting the desired signal. If you can't afford to do the necessary runs, then you must see what can be done to decrease noise. For example, Table 2 shows a minimum of 64 runs for a signal-to-noise ratio of 1. However, if you could cut the noise in half, the signal-to-noise ratio would double (to 2), thus reducing your runs from 64 to 16. If you can't reduce noise, then you must accept an increase in the detectable signal (the minimum effect that will be revealed by the DOE).

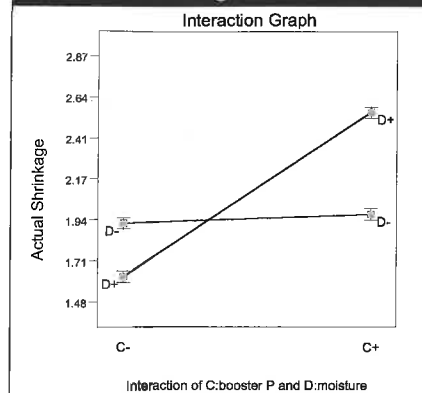
You can improve the power of the DOE by adding actual replicates where conditions are duplicated. You can't just get by with repeat samples or measurements. The entire process must be repeated from start to finish. If you do submit several samples from a given experimental run, enter the response as an average.

For our case study on injection molding, control charts reveal a standard deviation of 0.60. Management wants to detect an effect of magnitude 0.85. Therefore, the signal-to-noise ratio is approximately 1.4. The appropriate number of runs for this two-level factorial experiment is 32. We decide not to add further replicates due to time constraints, but several parts will be made from each run. The response for each run becomes the average shrinkage per part, thus dampening out variability in parts and the measurement itself.

4. Randomize the run order

The order in which you run the experiments should be randomized to avoid influence by uncontrolled variables such as tool wear, ambient temperature and changes in raw material. These changes, which often are time-related, can significantly influence the response. If you don't randomize the run order, the DOE may indicate factor effects that are really due to uncontrolled variables that just happened to change at the same time. For example, let's assume that you run an experiment to keep your copier from jamming so often during summer months.

Figure 2: Interaction plot for CD effect on shrinkage



During the day-long DOE, you first run all the low levels of a setting (factor "A"), and then you run the high levels. Meanwhile, the humidity increases by 50 percent, creating a significant change in the response. (The physical properties of paper are very dependent on humidity.) In the analysis stage, factor A then appears to be significant, but it's actually the change in humidity that caused the effect. Randomization would have prevented this confusion.

The injection molders randomized the run order within each of several machines used for their DOE. The between-machine differences were blocked.

5. Block out known sources of variation

Blocking screens out noise caused by known sources of variation, such as raw material batch, shift changes or machine differences. By dividing your experimental runs into homogeneous blocks, and then arithmetically removing the difference, you increase the sensitivity of your DOE.

Don't block anything that you want to study. For example, if you want to measure the difference between two raw material suppliers, include them as factors to study in your DOE.

In the injection molding case study, management would like the experimenters to include all the machines in the DOE. There are four lines in the factory, which may differ slightly. The experimenters divide the DOE into four blocks of eight runs per production line. By running all lines simultaneously, the experiment will get done four times faster. However, in this case, where the DOE already is fractionated, there is a cost associated with breaking it up into blocks: The interac-

Table 2: Number of Runs as a Function of Signal-to-Noise Ratio

Signal-to-Noise Ratio (Δ/σ)	Minimum Number of Runs
1.0	64
1.4	32
2.0	16
2.8	8

tion of mold temperature and holding pressure can't be estimated due to aliasing. Aliasing is an unfortunate side effect of fractional or blocked factorials.

6. Know which effects (if any) will be aliased

An alias indicates that you've changed two or more things at the same time in the same way. Even unsophisticated experimenters know better, but aliasing is nevertheless a critical and often overlooked feature of Plackett-Burman, Taguchi designs or standard fractional factorials.

For example, if you try to study three factors in only four runs—a half-fraction—the main effects become aliased with the two-factor interactions. If you're lucky, only the main effects will be active, but more likely there will be at least one interaction. The bearings case (Figure 1) can be manipulated to show how dangerous it can be to run such a low-resolution fraction. Table 3 shows the full-factorial test matrix. In this case, the interaction AB is very significant, so it's included in the matrix. (Note that this column is the product of columns A and B.) The half-fraction is represented by the shaded rows—the responses for the other runs have been struck out. Observe that in the highlighted area the pattern of the minuses (lows) and pluses (highs) for AB is identical to that of factor C. Put another way, $C = C + AB$, where the equal sign indicates aliasing. By going to the half-fraction, we don't know whether the effect is the result of C, AB or both.

Aliasing can be avoided by doing only full two-level factorials or high-resolution fractionals, which isn't practical. Plackett-Burman or Taguchi designs are often very low in resolution and therefore give very misleading results on specific effects. If you must deal with these nonstandard designs, always do a design evaluation to see what's aliased. Good DOE software

will give you the necessary details, even if runs are deleted or levels changed. Then, if any effects are significant, you will know whether to rely on the results or do further verification.

The injection molding study is a fractional factorial design with mediocre resolution: Several two-factor interactions are aliased. A design evaluation provides the specifics: $CE = CE + FG$, $CF = CF + EG$, $CG = CG + EF$. If you evaluate the effects matrix for CE vs. FG, you will see a perfect correlation. The plus symbol in the alias relationship tells you that the calculated effect could be due to CE plus FG. If these or any of the other aliased interactions are significant, further work will be needed.

7. Do a sequential series of experiments

Designed experiments should be executed in an iterative manner so that information learned in one experiment can be applied to the next. For example, rather than running a very large experiment with many factors and using up the majority of your resources, consider starting with a smaller experiment and then building upon the results. A typical series of experiments consists of a screening design (fractional factorial) to identify the significant factors, a full factorial or response surface design to fully characterize or model the effects, followed up with confirmation runs to verify your results. If you make a mistake in the selection of your factor ranges or responses in a very large experiment, it can be very costly. Plan for a series of sequential experiments so you can remain flexible. A good guideline is not to invest more than 25 percent of your budget in the first DOE.

8. Always confirm critical findings

After all the effort that goes into planning, running and analyzing a designed experiment, it's very exciting to get the results of your work. There is a tendency to eagerly grab the results, rush out to production and say, "We have the answer!" Before doing that, you need to take the time to do a confirmation run and verify the outcome. Good software packages will provide you with a prediction interval to compare the results within some degree of confidence. Remember, in statistics you never deal with absolutes—there is always uncertainty in your

Table 3: Test Matrix for Bearings Case (Darker Shaded Areas are Half-Fraction)

Factor A: Osculation	Factor B: Heat	Interaction AB	Factor C: Cage	Response: Life (Hours)
-1	-1	+1	-1	17
+1	-1	-1	-1	25
-1	+1	-1	-1	26
+1	+1	+1	-1	85
-1	-1	+1	+1	19
+1	-1	+1	+1	21
-1	+1	-1	+1	16
+1	+1	+1	+1	128

recommendations. Be sure to double-check your results.

In the injection molding case, the results of the experiment revealed a significant interaction between booster pressure and moisture (CD). The interaction is not aliased with any other two-factor interactions, so it's a clear result. As shown by the flat line on figure 2, shrinkage will be stabilized by

keeping moisture low (D-). This is known as a robust operating condition.

Contour graphs and 3-D projections help you visualize your response surface. To achieve robust operating conditions, look for flat areas in the response surface.

The 3-D surfaces are very impressive, but they're only as good as the data generated to

create the predictive model. The results still must be confirmed. If you want to generate more sophisticated surfaces, you should follow up with response surface methods for optimization. These designs require at least three levels of each factor, so you should restrict your study to the vital few factors that survive the screening phase.

Promoting DOE

Design of experiments is a very powerful tool that can be utilized in all manufacturing industries. Quality managers who encourage DOE use will greatly increase their chances for making breakthrough improvements in product quality and process efficiency.

About the authors

Mark J. Anderson and Shari L. Kraber are consultants at Stat-Ease, Inc. in Minneapolis. E-mail Anderson at mark@statease.com

qd

Reprinted from *Quality Digest*, July 1999. For subscription information, call (530) 893-4095, e-mail subscriptions@qualitydigest.com or visit www.qualitydigest.com

Stat-Ease, Inc., Hennepin Square, Suite 191, 2021 E. Hennepin Ave.
Minneapolis, MN 55413-2723, www.statease.com
p: 612.378.9449, f: 612.378.2152

