Design of Experiment: Why?

And why not?

By Katie Daisey

Q I mostly understand the process of running a design of experiment (DOE), but why should I do one at all?

A First, for our readers not familiar with it, a DOE is a procedure in which data is obtained using sound statistical techniques from specific combinations of controlled variables. In turn, the data is used to build a mathematical model that predicts the outcome of all possible combinations of those variables (within a reasonable range). This is often contrasted with the standard iterative or "Aristotelian" process, although nothing prevents an experimenter from running sequential DOEs for more specific information.

Most DOE experiments are full or partial factorial experiments. In a factorial DOE, the experimenter first determines what variables (or factors) they would like to study. They then determine how many different levels, such as high/medium/low within each factor to look at. Then they do an experiment for all of the possible combinations of factors (full-factorial) or some purposefully defined set of combinations (partial-factorial) randomly. Finally, they give all that data to a statistician, who comes back with a magical model that tells the experimenter which factors matter, the size of their effects, and about what levels to use to get the best results.

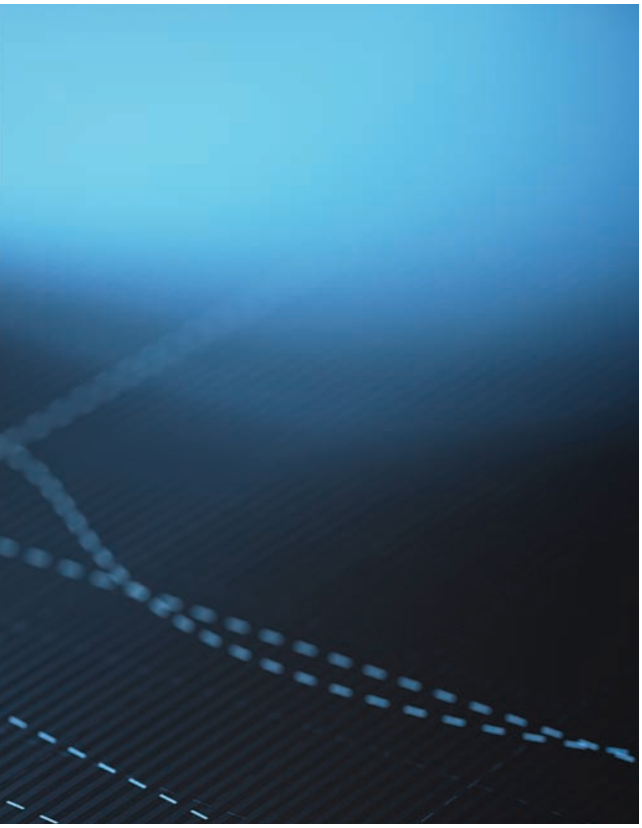
WHY NOT DOE?

Before outlining why a DOE should be used, let's consider in which situations a DOE is not necessary.

The first situation is when an experimenter is very familiar with the process and understands well the underlying reasons why and how each factor affects the final outcome. Instead of using a DOE to create a mathematical model to understand the experimental space, the experimenter, usually through years of experience, creates an internal model.

The second situation is when experiments are quick, cheap, and easy to run (or there is a large amount of historical data). Instead of using a DOE, it's easier to just run enough experiments to completely cover the experimental space. (A careful experimenter would first check any historical data to ensure that no possible combinations were missed, that there is no correlation with a hidden variable, and that it is otherwise suitable.)

The third situation is when experiments are dangerous and only small changes can be safely made. The experimenter can only make a series of small steps to optimize the process. Actually, there is a type of DOE specifically for this situation: a simplex optimization. The experimenter runs a small set of experiments in the "normal operating conditions" region and uses a mathematical formula to calculate where the next experiment should be run only a small step outside of "normal." The result of this experiment is added into the formula to calculate the next small step, and so forth.



WHY DOE?

One of the main reasons that experimenters choose, perhaps unwisely, not to use a DOE, is the perceived increased number of experiments to be done compared to their normal work. However, a DOE mathematically provides the most information with the fewest experiments. Experiments that seem unsuccessful are necessary to provide explanations for the entire process.

The perception that a DOE requires much more preparatory work than traditional experiments is often because this necessary work is done separately or is unfortunately seen as optional.

For instance, a statistician might ask an experimenter about the error in their measurement system as this information is critical to determining the signal to noise ratio, how large the error on the predictions will be, and how many replicates need to be completed. If a gage repeatability and reproducibility study or other measurement system analysis has never been done (or it has just been a while), it may seem like this is additional work that is only required for a DOE rather than routine analyses. Or, a statistician might want to know which variables are most likely to interact with each other, and whether the resulting output will be linear or polynomial. These are questions that require some insight into the process, a process diagram, and maybe some preliminary experiments.

One of the advantages of a DOE is the ability to randomize the experiments, which is not possible in traditional experiments. While great care and expense can be taken to control as many environmental variables as feasible, it is rarely possible to control everything. In non-randomized experiments, it is not possible to separate changes due to environmental variables from changes due to independent variables.

The second key advantage of a DOE is the understanding of the entirety of the system being studied. It is rare that an experimenter is only concerned about optimizing a single target variable. More often, there are multiple target variables being optimized and usually complicated decisions as to how the targets should be prioritized. There may be multiple acceptable combinations of the target variables, or none, and only exploring the entire space at an appropriate resolution will find that information. The mathematical model can also be reused in the future if there is a change in the desired output of the target variables. Furthermore, while not impossible with traditional experimentation, it is much easier to find and quantify how interactions between variables affect the outcome. Optimizing one independent variable at a time, as one would do intuitively, makes it extremely difficult to handle interactions. Finally, information across the entire study space can help inform the switch from an empirical, mathematical model to a mechanistic, causal model.

The third advantage of a DOE, which is often underutilized, is the ability to analyze the variance of a process across the space. It is important to know whether process variance is constant. A well-designed DOE with replicates may allow for statistically significant separation of variance, where a naive treatment may leave an experimenter unsure. More importantly, examining the variance of the system provides detail into the sensitivity of the process — not only how the different controlled factors affect the output, but how the uncontrolled factors may have a larger effect with one set of set points versus a different set. Note that this is a separate procedure from using a DOE to determine robustness, where possible failure mechanisms are chosen to be the controlled factors.

While not all situations require a DOE, time savings, information about a broader range of conditions, and insight into uncertainty are major advantages to consider. (Learn more about DOE in the works noted below.) ■

REFERENCES

1. Montgomery, D.C., *Design and Analysis of Experiments,* Wiley, Hoboken, New Jersey, 2019.
2. Cox, David, and Reid, Nancy, *The Theory of the Design of Experiments,* Chapman & Hall/CRC, Boca Raton, Florida, 2000.
3. NIST/SEMATECH, *Engineering Statistics Handbook,* April 2012, doi.org/10.18434/M32189.
4. Bezerra, M.A., et al., “Simplex Optimization: A Tutorial Approach and Recent Applications in Analytical Chemistry,” *Microchemical Journal,* Vol. 124, 2016, pp. 45–54, doi.org/10.1016/j.microc.2015.07.023.

Katie Daisey is a scientist (chemometrician) for Arkema Inc. A member of ASTM International since 2018, she is recording secretary for the committee on quality and statistics (E11).

John Carson, Ph.D., senior statistician for Neptune and Co., is the Data Points column coordinator. He is a member of the committee on quality and statistics (E11), as well as the committees on petroleum products, liquid fuels, and lubricants (D02), air quality (D22), and environmental assessment, risk management, and corrective action (E50).