Trimming the FAT out of Experimental Methods

Mark J. Anderson, Stat-Ease, Inc.

In 1969 the renowned management guru Peter Drucker predicted that “to make knowledge work productive will be the great management task of this century.” Who knows more about their work than engineers and scientist? Yet, by traditional scientific methods, even the world’s foremost subject-matter experts generally do their experiments in a most unproductive way—one factor at a time (OFAT). This so-called “scientific method,” commonly attributed to Francis Bacon in the 17th century, but with origins as early as the Greeks in 1600 BC, reached its apogee in the work of Thomas Edison, who doggedly applied trial and error OFAT to invent the light bulb, among other things. His approach was “1% inspiration and 99% perspiration.” The Edisonian approach inspired A.N. Whitehead to say in 1926 that “The greatest invention of the nineteenth century was the invention of the method of invention.”

Ironically, around this time a British statistician, Ronald Fisher, while working in the field of agriculture developed a new form of experimentation called two-level factorial design. Mathematically these are characterized as “2^k”, where k represents that number of experimental factors. Fisher’s innovative 2^k design of experiments (DOE) overcame the year-long growing cycles by studying many factors in parallel fashion via sophisticated matrix-based test plans. Figure 1 contrasts a three-factor 2^3 design with an OFAT of equivalent replication, about which no one disagrees—the more the better.

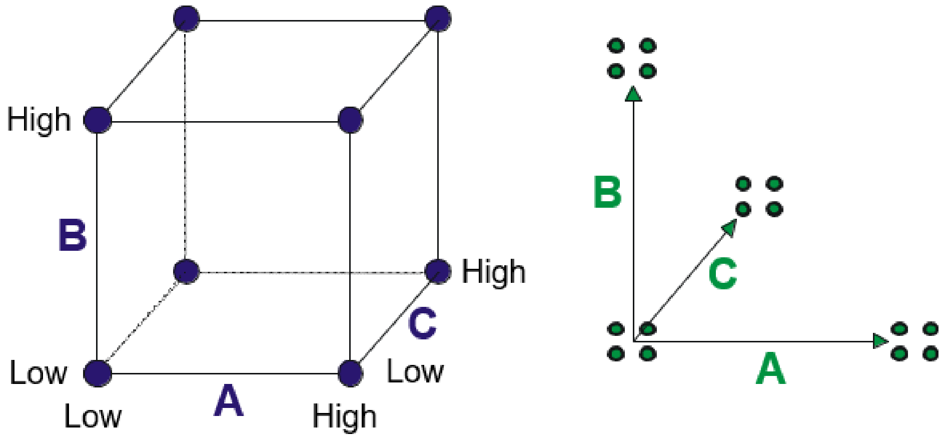

Figure 1: Comparison of two-level factorial designs (left) versus OFAT (right)

The factorial design offers four runs at the high level for A and the same for the low (right versus left faces of cubical experimental region). Similarly, factors B and C also benefit from having four runs at both high and low levels (top versus bottom of cube and back versus front; respectively). Therefore, the OFAT experimenter must provide four runs each at the high levels of each factor versus four at the base line (all low levels at origin) to provide similar power of replication for effect estimation. However, this necessitates a total of 16 runs for OFAT versus only 8 for the two-level factorial. Due to its far more efficient parallel processing 2^k DOE trumps the serial OFAT scheme, and its efficiency advantage only becomes more pronounced as the number of factors (k) increase.

Let's really put OFAT to the fire by making use of an even more powerful form of DOE called response surface methods (RSM). Assume that the OFAT experimenter arbitrarily starts with factor A and varies it systematically over nine levels from low to high (-2 to +2) while holding factor B at mid-level (0). The resulting response curve can be seen on Figure 2a. The experimenter sees that the response is maximized at an A-level near one (0.63 to be precise). The next step is then to vary B while holding factor A fixed at this "optimal" point. Figure 2b shows the results of the second OFAT experiment (note the change in scale for the response (y) axis).

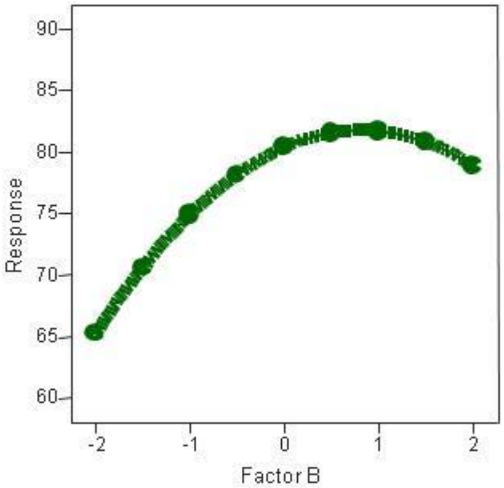

Figure 2a: OFAT on factor A

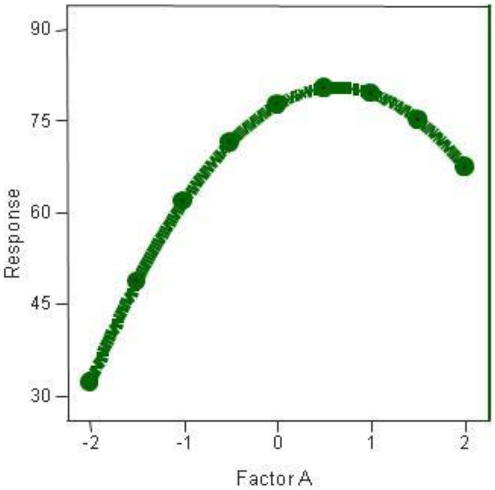

Figure 2b: OFAT factor B (A fixed at 0.63)

The response is now increased from 80 (previous maximum) to 82 by adjusting factor B to a level of 0.82 based on this second response plot. The OFAT experimenter proudly announces the optimum (A,B) combination of (0.63, 0.82), which produced a response above 80 in only 18 runs. However, as you can see on Figure 3 produced via RSM the real optimum is far higher (~94) than that attained via OFAT.

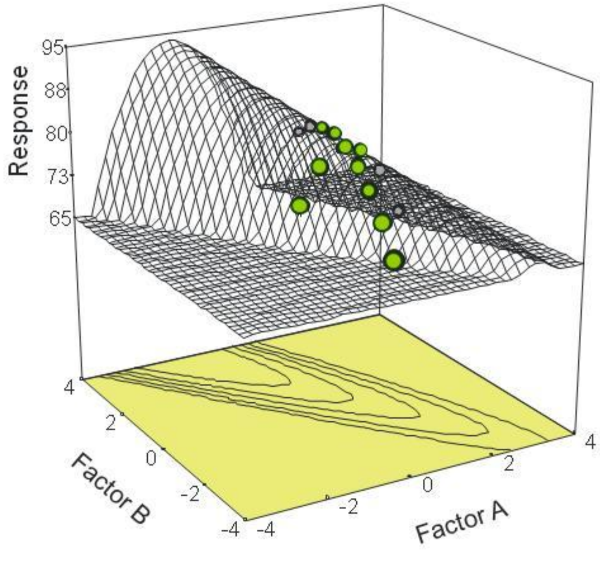

Figure 3: OFAT points shown on true surface

By letting go of OFAT in favor of multivariable testing via DOE/RSM, scientists and engineers can become more productive and effective in their experimentation, thus gaining greater knowledge more quickly—an imperative for the 21st century.