DOE – Recognizing common design problems

When there are insufficient statistical considerations incorporated in the design of an experiment, we can encounter inconclusive statistical analysis of the data obtained or, worse yet, get misleading conclusions. Below discussion centers on some potential problems that can occur when statistical methodology is not used to design scientific or engineering experiments.

1. Masking factor effects

The worst nightmare that a researcher can have is after investing substantial project funds and a great deal of time and effort only to find that the research hypotheses are not supported by the experimental results obtained. Many times the lack of statistical confirmation is the result of the inherent variability of the test results which mask the factor effects under study. The importance of necessity to consider the experimental error variation in the statistical design of an experiment cannot be over emphasized.

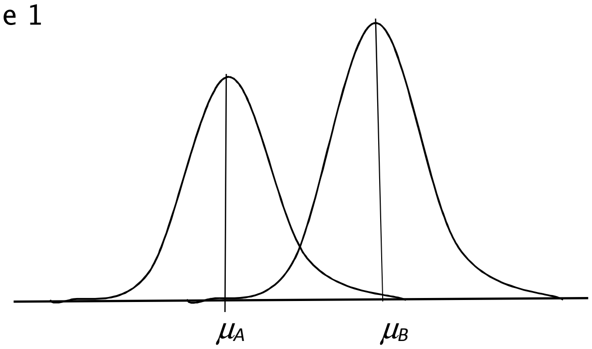
Consider for example the color measurements listed in the table below:

	Color Measurement		
Participant	Week 1	Week 2	Week 3
A	12.1	14.2	13.9
B	19.1	17.6	16.2
C	33.8	34.7	33.2
D	33.0	31.7	30.3
E	35.8	37.7	35.6
F	42.0	38.4	41.5
G	36.8	35.2	35.7

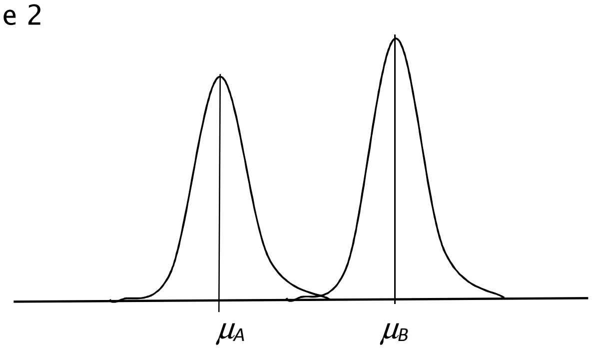
In this study of skin color measurements not only is there variation among the participants, but there is also variation for each participant over the three weeks of the study. Experiments that are intended to study factor effects (such as the suntan products) on skin color must be designed so that the variation in subjects and across time does not mask the effects of the experimental factors.

In another scenario where we take a look at the relationship between the detectability of factor effects and the variability of responses, we have the following schematic representation of two situations where A and B are two levels of a factor. In both cases, we have similar average response at each factor level but the variability of the response changes from case to case.

Case 1



Case 2



In case 1, the variability of the test results is so great that we would never be sure if the factor effect is indeed measuring a true difference in the population means μ corresponding to the two factor levels. There is a possibility that the variation of the responses is at about a common mean. So, we cannot be convinced that there is sufficient evidence to say the two population means are different because of the variability in the responses.

In case 2, we see better response precisions (or less variation of the responses) than in case 1. In here, there is a strong evidence, due to the small variation in the response relative to the large differences in the averages that the factor effect is indeed measuring a substantial difference between the population means corresponding to the two factor levels.

Hence we can conclude that the difference in the means due to the factor levels has masked by the variability of the responses in case 1 but not in case 2. The implication of this example for the statistical design of experiments is that the variation in case 1 must be reduced or compensated for (e.g. by blocking or a large experiment size) to ensure that the difference in the means is detectable due to the factor effect, instead.

2. Uncontrolled factors

In any experiment, we can encounter many factors on the response which are not controllable in actual operations but may be better controlled during experimentation, such as humidity, weather conditions, electronic response, ambient air pollutant interferences, etc. It is no doubt that few researchers would intentionally ignore factors that are known to exert important influences on a response, but there are many subtle ways in which failure to carefully consider all factors of importance can compromise the conclusions drawn from experimental results. The uncontrollable variation of certain factors can lead to the *confounding* of their effects on the response.

We must therefore need to construct designs in which factor of interest are systematically varied accordingly and to consider the likely magnitude of the inherent variation in the test results during planning the number of test runs.

3. Erroneous application of DOE efficiency

Time and cost efficiencies are always important objectives in experimental work. We always desire to run economical experiments but without proper statistical planning, we may encounter wasteful experimentation or even not achieving the project's goals.

This commonly happens when several factors are being investigated in an experiment. When guided only by intuition without much scientific reasoning, many different types of designs could be proposed, each of which might lead to flawed conclusions. Some would choose to hold factors constant that could have important influences on the response. Others would allow many unnecessary changes of factors that are inexpensive to vary and few changes of critical factors that are costly to vary.

Efficiency is achieved in statistically designed experiments because each observation generally provide information on all the factors of interest in the experiment. One may opt for a single factor at a time through optimization as a function of several factors, in order to economically conduct experiments. It is believed that the one-factor-at-a time testing is close to the minimum number of runs in investigating several factors simultaneously, and one can readily assess the factor effects as the experiment progresses because only a single factor is being studied at any stage. It is used merely to assess the importance of the factors in influencing the response.

However, such type of testing has a serious drawback because many times the goal of a study is not only to optimize the response due to a factor but also to model its behavior over the experimental resign (or so-called factor space). It may miss the early opportunity to identify any optimal combination of factor levels and hence has achieved no economy of effort relative to testing all possible combinations of the factor level.

Hence, we need to look for more economically efficient statistical experimental designs that do permit the fitting of curved response contours due to factor effects, the investigation of joint factor effects, and estimation of experimental error variation. *Factorial experiments* and their corresponding ANOVA computations are valuable designs when simultaneous conclusions about two or more factors are required.