I. Introduction

The typical approach to experimental design is either the one factor at a time (OFAT) or the design of experiments (DOE). Clearly, the DOE approach is generally recognized as a superior choice from a time savings perspective since in most experiments there is generally more than one perceived factor at play. But the more pronounced reason why a DOE is more powerful then the OFAT approach is due to the amount of knowledge that is attainable by the DOE. The true primary purpose of the DOE is to learn about the interactions that are occurring between the factors. This factor interaction cannot be determined through an OFAT experiment since by design only one factor is at play at any one time. If the DOE is designed properly, one can learn that potential interactions can be more important and have a greater impact on the output then any one single perceived major factor. Too many times, a DOE outcome or study is prejudiced by the factors that are chosen. In an effort to minimize the number of factors, factors are chosen purely on experience or prejudice on what is expected to impact the studied output. In this point, one is focusing on the factor as the primary cause or impact on the outcome so the true value or strength of the DOE is not fully utilized. Prior to the execution of a DOE, one should perform a statistical analysis of the factors without prejudice to determine what are in fact the key factors that should be a part of the DOE, by doing this, an initial determination of interaction between those factors may be realized and thus assist in designing a more meaningful DOE.

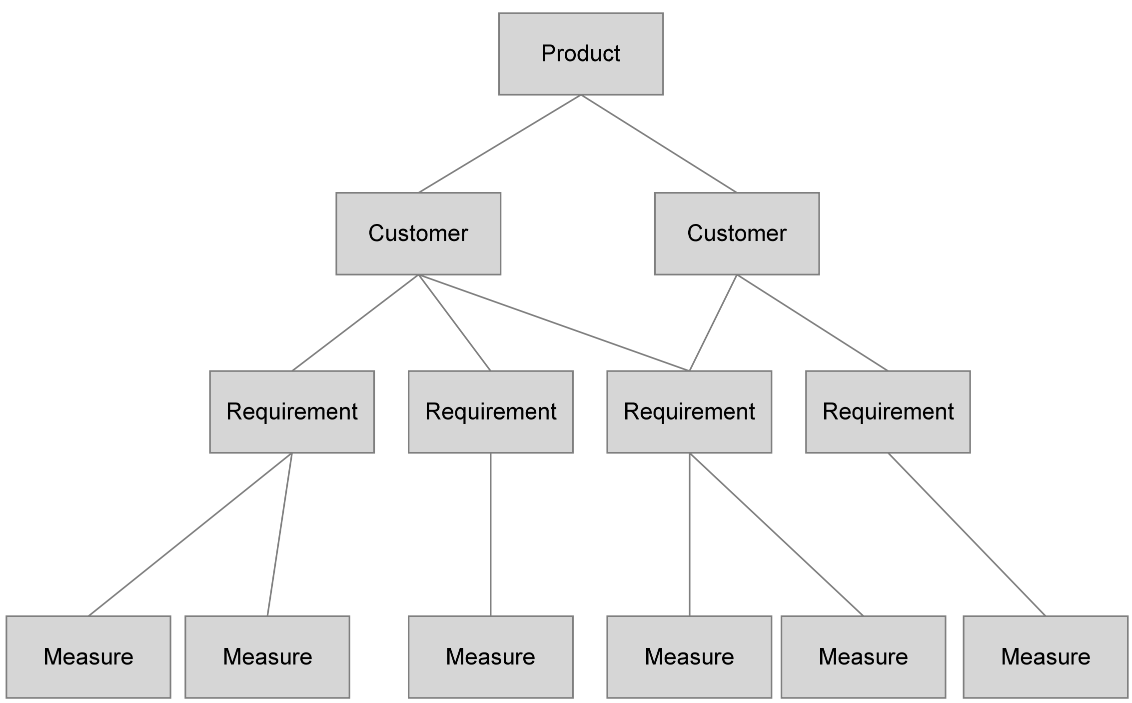
This statistical analysis of the factors is referred to as the study of the components of variation (COV). The COV can be thought of as a ‘funnel’ where potential factors pass through, the output of the funnel are the determined factors that should be studied in the ensuing DOE. It could be considered by some in the six-sigma community that the execution of a DOE without the prior execution of a COV as an improper approach to problem solving, a waste of resources, time and money. In the design of a COV, factors are taken through a wider study range in an effort to understand the statistical significance of factors and most importantly their potential interactions. Understanding these factor interactions with respect to their impact on the variability within a studied process, or unit of a process is to lead to a better understanding on what needs to be studied and fine tuned in order to impact or develop a control within the greatest gain for the study assignment. All processes and activities have variability (be them industrial processes, laboratory experiments or business processes), some variability are controllable and some are not and some are not worth expending the time and cost to control. It is incumbent on the analyst to determine which types of variability they have and what the sources of that variability are and ultimately how to control them if necessary.

II. Components of Variation Study Design

How is a COV study designed? What does it look like? As was stated above, the objective of a COV is to determine what causes have the largest impact on the variability in a system and typically these causes are due to interactions between more than one factor. It is not surprising to see an outcome of a study such as this that the major or significant impact is due to the interaction of two seemingly minor factors rather than one perceived significant factor.

The structure of a COV is to first understand the overall system or process flow, then to design a process or system 'tree'. The branches of this tree are the sub-units of the system and as one drills down the process flow the subsequent branches are developed for the ensuing study. The branches of the study tree, one layer to the next can be either independent of the previous layer or they can be dependent. Independent branches can be placed or studied interchangeably where dependent branches have to be specified with the dependent group preceding them in the design. A tree designed with independent sub-branches is said to be an uncrossed design and a dependent tree (with fixed variables) is said to be crossed.

In the figure below: one can see a very simple basic tree structure for a COV study:



This structure can be applied to any study of a process or system and can be more complex depending on the number of potential variables. The study can be designed to be completely crossed or can be a mix of both crossed and uncrossed. Typically, if the study is to be mixed, the tree is designed such that the lower sub-branches are set with a crossed structure, meaning that the output or study is directly dependent on the level above them and that they are not interchangeable.

III. Conclusion

There are many advantages to the execution of a COV not the least of which is saving time and money knowing that the subsequent DOE is designed and executed with the variables that have been statistically selected as having a significant impact on the studied output. Factor selection is the single most critical aspect of properly designing a DOE such that the true understanding and value are attained from the cost and time expended for those experiments. Understanding which variables and ultimately being able to quantify the impact, singly or as an interaction may make all the difference in understanding and controlling the performance and variability in a system.

Alpaslan (Alp) Yaman, Ph.D.

Certified Black Belt

Principal Consultant

Biotech, Pharma & Device Consulting, LLC

www.biopharmadvice.com