MINI PAPER

BY MARILYN HWAN, SANTA CLARA, CA (REGION 6 COUNCILOR)

A HOLISTIC APPROACH TO THE DESIGN OF EXPERIMENTS - PART I

ABSTRACT

Part I of this paper introduces the design of experiments for the development and improvement of products and processes in a holistic manner. According to Webster, "holistic" means "emphasizing the organic or functional relation between parts and wholes (rather than atomistic)." Instead of presenting the design of experiments in isolation of other needed activities, this paper follows the holistic approach by first showing where experiments fit in the overall scheme of developing and improving processes or products. Then, to help novices execute experiments, a step-by-step guide for planning and analyzing experiments is given. Although this paper is written for novices, the author hopes it will also inspire seasoned users to constantly question whether existing methodology can be improved upon.

Part II (next issue) will illustrate the design of experiments with actual applications.

1.0 INTRODUCTION

Running "designed experiments" in industry has somehow maintained an air of mystery, even though "experiments" have been run for many centuries. This paper is written to help clarify some concepts of designed experiments through presenting an overall approach to show how the design of experiments fits into the overall scheme of developing or improving products and processes. This paper will cover:

- A template for developing and improving of products and processes
- Steps of designing experiments
- References for additional reading

Although the methods discussed in this paper can be applied to developing and improving many products and processes, it should be noted that they cannot be used to solve all problems. It should also be noted that there may be additional areas during improvement and development where experiments are extremely useful which are not mentioned in this paper.

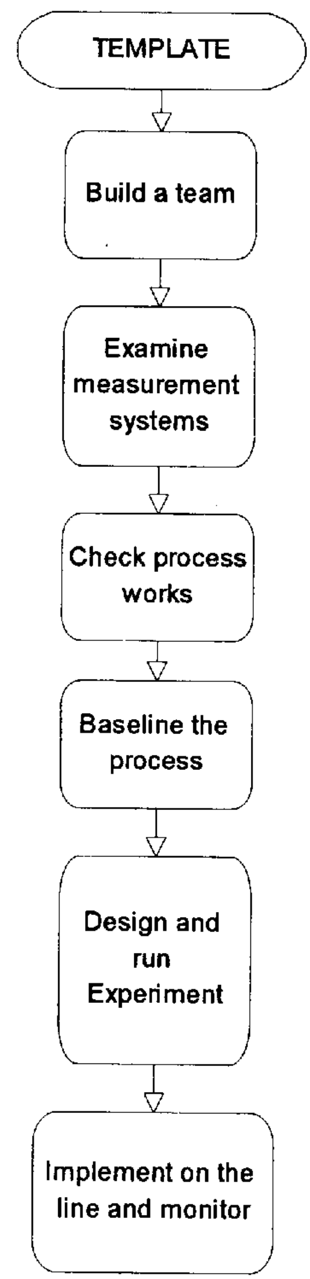
Please also note that the material in this paper applies to both process and product development. In order to simplify, "process" will be used throughout this paper.

2.0 A TEMPLATE

The purpose of this section is to show how experiments can fit in the overall scheme by listing steps which are often needed to develop or improve a process. Each step is really a topic of its own and is expanded in many books and papers in the quality and engineering literature. Since many development or improvement efforts can follow this pattern, it can be considered a "template" for developing and improving processes. (See Flowchart 1: TEMPLATE)

Running experiments in isolation of other engineering considerations can lead to disappointment and failure. For example, running an experiment and unknowingly using an inadequate measurement system to take the measurements, can lead to erroneous conclusions. Or, one might run an experiment to find good conditions for a process, yet find that the conditions are not good when actually placed on the manufacturing line. Following certain steps can help to minimize these kinds of problems. Although the necessary steps may vary between different applications, the following template gives a general idea of steps to be taken to develop or improve processes:

FLOWCHART 1: TEMPLATE



2.1 BUILD A TEAM

Build a team from persons who can contribute knowledge from different angles. For example, to develop a process, a team typically includes members from process engineering, equipment engineering, production, and quality engineering/ methodology. The team should brainstorm, plan, and execute the activities in this template together. Fishbone diagrams, Pareto charts, Affinity diagrams and other quality tools may prove useful at this stage.

2.2 EXAMINE THE MEASUREMENT SYSTEMS

Check that each measurement system needed for the process is adequate. Use a Gauge Repeatability and Reproducibility (R&R) Study or an equivalent (or better) method, to study the amount of major sources of variation such as operator reproducibility and equipment repeatability. The Gauge R&R Study examines the variation of a measurement system, and compares this variation against the product or process tolerance. Even if the exact tolerance of the product or process is unknown, the study can still be used to obtain an estimate of the variation of the measurement system so that its capability for future more demanding processes can be estimated.

2.3 CHECK THAT THE PRODUCT OR PROCESS BASICALLY WORKS

If an old production system or a product is being studied, check that it functions properly. For example, this might include checking that there are no gas leaks, malfunctioning pumps, reversal of electrical connections, etc. If new production equipment is being studied, check that the equipment is properly assembled. This might be done by running the equipment and process according to the supplier's recommended settings, and checking the inputs and outputs of the system against expected values. A short capability study might be performed here. (See Reference 3.)

2.4 BASELINE THE PRODUCT OR PROCESS

Measure and record data on factors which could affect the process, but are not varied in the experiment. Certain elements of the system being studied may not be included in the experimental design, but could influence the results or subsequent reproducibility of the results. Equipment related base-lining might mean recording data such as the efficiency of pumps or age of a metal target. Product or process related data might mean recording the lot number of material used.

2.5 DESIGN, RUN AND ANALYZE EXPERIMENT(S) USING STATISTICAL METHODS

After checking that the system basically works and evaluating the capability of a system, experiments might be run to improve the quality of a process and/or to characterize the process. Although some choose to run experiments only if the process is not adequate, others may choose to run experiments regardless of the existing capability to improve the quality of the process and to understand the process. Understanding the effects of the factors involved in running a process can greatly help to develop troubleshooting guides which can be used on the manufacturing line.

Design the experiment(s) along with the method of analysis. Before any experiment is run, plan how the data will be analyzed. Design clear data collection sheets. Provide clear instructions as to how to run the experiment, sample, measure, and record the data. Check that the conditions chosen have no harmful effects elsewhere; - it would not be fruitful to solve one problem, only to produce others.

The design and analysis of experiments is expanded in Section 3 of this paper.

2.6 IMPLEMENT ON THE LINE AND MONITOR

Once the changes have been placed on line, key steps of the process should be monitored. Some tools for monitoring on line are Shewhart Control Charts, Feedforward, and Feedbackward On-Line Quality Control.

3.0 STEP-BY-STEP GUIDE TO RUNNING STATISTICALLY DESIGNED EXPERIMENTS

This section expands upon Section 2.5 by providing steps for designing and analyzing experiments. (See Flowchart 2: STEPS FOR RUNNING AN EXPERIMENT)

3.1 DETERMINE AND AGREE UPON THE PURPOSE OF THE EXPERIMENT

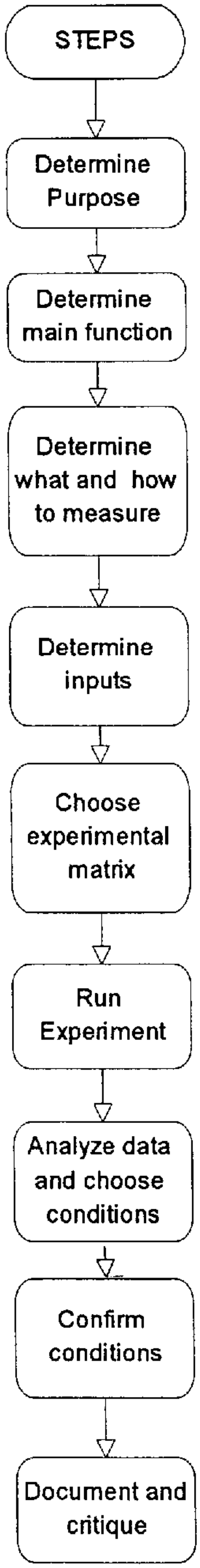
The methods described in this paper are used to develop new processes or improve existing processes. The specific purpose may be to minimize the variation and place a process on target, or, it may be to characterize a process by creating a model. Often, experiments are run to troubleshoot. It is important to define the purpose carefully, as the decision on how to design the matrix and analyze the data depends on what questions are being asked.

The team must agree upon the purpose of the experiment. This purpose should be written down; otherwise, each person on the team may still think the experiment is being run for a different purpose.

3.2 DEFINE THE MAIN FUNCTION OF THE PROCESS OR PRODUCT YOU ARE STUDYING

The definition of the main or ideal function of a process or product is important; in fact, it can actually be the key to developing a robust process. For example, the function of a semiconductor low temperature oxide process (LTO) process is to deposit a thin layer of oxide; the ideal function is that the amount deposited is proportional to the amount of deposition time. The function of an operational amplifier may be to amplify voltage; the ideal function may be that the output voltage is lin-

FLOWCHART 2: STEPS FOR RUNNING AN EXPERIMENT



9

ear with respect to the input voltage. The definition of the function helps to determine the input and outputs of the experiment.

3.3 DETERMINE THE CHARACTERISTICS TO BE MEASURED AND THE METHOD OF MEASUREMENT.

Some outputs are better than others for use as responses in experiments. It is usually better to use an output which is continuous rather than discrete (attribute). Examples of continuous data are width, thickness, and bond pull strength. Examples of attribute data are "pass or fail, " defective or not, "go or no-go." The measurement should be well-defined and easy to execute (if possible). This includes choosing the measurement and processing equipment to be used, what to measure, how to measure, where to measure, and where to document the data. Training to cover all these aspects should be required. If the variability or capability of the measurement method has not already been evaluated, it should be evaluated before the experiment is run.

3.4 DETERMINE THE INPUTS

Choose the input factors. For example, if the experiment is used for modeling an output of a process, only the inputs which we can control and affect only the average are used as input factors. If the experiment is used for minimizing variation in the face of factors which we cannot, or do not want to, control during manufacturing or in the field, then these (noise) factors may be deliberately included in our experiment.

Determine the range to be used for experimentation along with how many different number of settings or "levels" to be assigned to each variable. It is usually best to experiment with the largest range feasible, so that the variation inherent to the process does not mask the effect of the factors, especially if it is necessary to use small samples.

Determine the number of settings or levels to be used for each factor. If it is expected that the relationship between the input and output is linear, then two levels will be used. If a nonlinear function is expected, then three or more levels should be used to confirm and quantify curvature in the response.

Determine the number of samples (replications) for each run. This depends on the sensitivity or risk desired (statistical considerations), along with time and economic considerations.

3.5 CHOOSE AN EXPERIMENTAL MATRIX

An experimental matrix defines the combination of settings for each of the factors being considered in the experiment. There are many types of experimental matrices. The choice of experimental designs depends on many factors. For example, it depends on the purpose of the experiment, on the number and choice of input and outputs, and on the number of levels to be used for each factor.

To help illustrate some of the calculations which can be made, a simple experiment using 2 factors, each evaluated at two settings (a 2^2 full factorial), will be used. In general, factorial designs are the most efficient for experiments with more than one factor. Factorial means that for each complete trial or replication of an experiment, all possible combinations of the levels of each factor are investigated. Suppose we would like to run an experiment using the factors and settings in Table 1.

FACTOR	LOW SETTING	HIGH SETTING
Temperature	450	550
Time	250	275

Table 1: Factors in an Experiment

RUN #	TEMPERATURE	TIME
1	LOW - 450	LOW - 250
2	HIGH - 550	LOW - 250
3	LOW - 450	HIGH - 275
4	HIGH - 550	HIGH - 275

Table 2: A 2^2 Full Factorial design

RUN #	TEMPERATURE	TIME	YIELD %
1	LOW - 450	LOW - 250	30
2	HIGH - 550	LOW - 250	70
3	LOW - 450	HIGH - 275	80
4	HIGH - 550	HIGH - 275	30

Table 3: Results of an Experiment

A full factorial would consist of running all the combinations. The design matrix for this 2^2 full factorial design is shown in Table 2.

If an experiment has 4 factors, each at 2 levels, a full factorial would consist of $2^4 = 16$ runs, and may be feasible to run. But very often, we must deal with as many as 7 factors each at 3 levels. A full factorial would produce lifetime employment running $3^7 = 2187$ combinations. Therefore, special subsets of these full factorials or other, more efficient designs, are used in order to manage the experiment. Information on commonly used designs such as "Orthogonal Arrays"; "Fractional Factorials", Plackett-Burman or Box-Behnken designs, can be found in the references in Section 5 of this paper.

3.6 RUN THE EXPERIMENT

Whether to randomize the order in which the matrix settings is run depends on two main considerations: the cost of randomization and whether a time dependent factor (known or unknown) will disturb the result. For example, if the cost of randomization may be high because it takes a very long time to reach and stabilize a setting, such as temperature, the experiment may be run (in blocks) such that the number of temperature changes is minimized. On the other hand, randomization of the order may be needed if the drift of a piece of equipment, or the environment, may be large enough to strongly influence the results.

The team is responsible for seeing that the runs are made according to the matrix, and in the agreed-upon order.

3.7 ANALYZE THE DATA

The effects of the factors should be analyzed. Depending on the design and purpose of the experiment, the main effects and certain interactions (between factors) will be calculated, and often graphed. To illustrate the calculations, the main effects and interactions of a simple 2^2 factorial will be analyzed.

Suppose an experiment on Deposition Time and Temperature is run, and percent yield is used as an output. The product is acceptable if the thickness of material is within defined tolerances. The results might be as shown in Table 3.

If only the first three runs were observed, we could conclude:

- If temperature is held constant, increasing the time raises the yield.
- If the time is held constant, raising the temperature improves the yield.

We would expect that if both the temperature and pressure were high (run 4), that the yield would be very high. Instead, it is low. Such a situation indicates that there is an "interaction" between the two main factors.

After all four combinations are run, the effects can be calculated. The main effect indicates how changing the level of a factor affects the output, and is computed by taking the responses (here % yield) at the high level and subtracting the responses at the low level of each factor. The main effect of temperature is:

$[(70 + 30) - (30 + 80)] / 2 = - 5\%$

The main effect of time is:

$[(80 + 30)] - (30 + 70)] / 2 = 5\%$

[For simplicity, we have made no reference to the statistical significance of these "effects" which may involve such tools as an F-max test (2 factors), or Bartletts test (more than 2 factors) for homogeneity (sameness) of variances, a Two-Way Analysis of Variance to test for average (level) effects, and a Multiple Range Test (more than 2 factors) such as Newman-Keuls, to order and compare the range of paired mean effects.]

When the temperature is low, the effect of time is:

$80 - 30 = 50\%$

When the temperature is high, the effect of time is

$30 - 70 = - 40\%$

This means that the effect of time when temperature is low is opposite in direction from when the temperature is high. This inconsistency of the effect of time on "yield" when the temperature is changed is known as "interaction." This is graphically shown by the crossing lines in Figure 1.

If there are only a very small number of factors involved, then there is no problem in running full factorials. If, however, there are many factors, a lot of runs are needed to calculate the interactions. Choosing different responses can reduce the amount of interaction. For example, "energy-related" responses such as thickness tend to have a minimum of interaction. Suppose an energy-related output such as "thickness" were used. The graph of temperature and time versus thickness might look like Figure 2, which shows no interaction between time and temperature. If interactions do not have to be calculated, then less runs are needed in experiments. Taking advantage of using such considerations (as Taguchi teaches), can help us obtain more meaningful outputs and can greatly increase the efficiency of experimentation. Using ideas such as this to make experiments more efficient is discussed in References 4 and 5 in Section 5 of this paper.

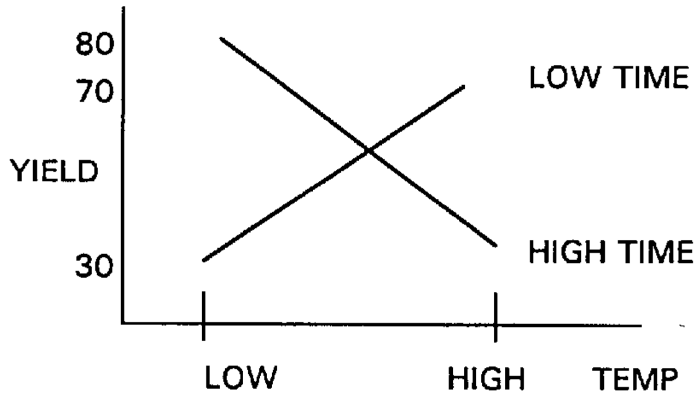


Figure 1: Graphical Illustration of Interaction Between Two Factors

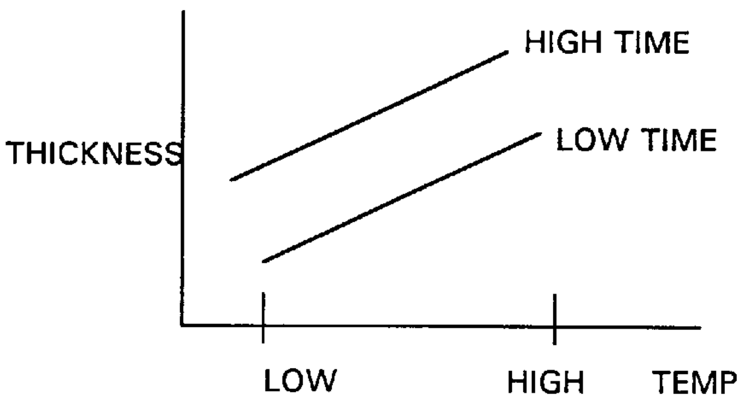


Figure 2: Graphical Illustration of Lack of Interaction Between Factors

3.8 CONFIRM THE "BEST CANDIDATE" CONDITIONS

Using a mathematical "model" derived from the experiment, estimate the output when particular settings are used. Such a model is often written algebraically (for example), as:

Thickness = $\beta_0 + \beta_1$ (time) + β_2 (temperature) + ε

where the coefficients of time and temperature are derived from their effects, β_0 is the intercept, and ε is an estimate of experimental error (plus noise).

Suppose that certain conditions are estimated to minimize variation and place the outputs on target. Using the "model," estimate the resulting variation and average value, or, choose several locations in the modeled region, and estimate the outputs. Run the chosen conditions next and compare the predicted responses against actual experimental values.

3.9 DOCUMENT EACH STEP OF THE EXPERIMENT

Develop a standard format for planning and documenting experiments. Doing so will not only record what was done this time, but will provide a basis for critiquing and improving the methodology.

3.10 CRITIQUE THE EXPERIMENT

Self-critique is a good tool for continuous improvement in experimental methodology. No matter how well the experiment was planned and executed,there might have been a better way to perform it. Also, what has been learned during an experiment will help us to plan the next experiment better. Some self-critiquing questions might be: Were the right questions asked during the experiment? Did the experiment answer the questions? Was it run in the most efficient manner? Was everyone trained sufficiently? Do the results make sense, and are they repeatable?

4.0 CONCLUSIONS AND CAVEATS

It is necessary to consider experiments in a holistic manner. Experimental design along with all the other steps are tools for development and improvement, and should not be used as "statistical magic." Sound engineering knowledge and questioning must accompany the use the steps described in this paper.

This paper is meant as a guide, and does not give details on how to execute each step. It is suggested that further reading or training, or consulting with an experienced experimental designer and/or statistician take place before running one's first designed experiment.

Part II of this paper (next issue), will focus on applications of experimental design.

5.0 REFERENCES

5.1 Box, Hunter, and Hunter , **Statistics for Experimenters**, (1978) John Wiley and Sons.

5.2 **Proceedings of the Ninth Symposium on Taguchi Methods (1991), Proceedings of the Tenth Symposium on Taguchi Methods (1992)**, The American Supplier Institute (ASI).

5.3 Mario Perez-Wilson, **Machine/Process Capability Studies (1989)**, Advanced System Consultants.

5.4 Genichi Taguchi, **System of Experimental Design (1976)**, Unipub, Kraus International Publications.

5.5 Madhav Phadke, **Quality Engineering Using Robust Design (1989)**, Prentice Hall.

5.6 Teruo Mori, **The New Experimental Design (1987)**, The American Supplier Institute Press.

5.7 John A. Cornell, **How to Apply Response Surface Methodology (1990)**, Volume 8 of the Basic References in Quality Control: Statistical Techniques.

Δ