

# EXPERIMENTAL DESIGN

An experimental design is a layout (structure, frame, outline) which is made to perform an experiment.

## Some Typical Applications of Experimental Design

Experimental design methods have found broad application in many disciplines. As noted previously, we may view experimentation as part of the scientific process and as one of the ways by which we learn about how systems or processes work. Generally, we learn through a series of activities in which we make conjectures about a process, perform experiments to generate data from the process, and then use the information from the experiment to establish new conjectures, which lead to new experiments, and so on.

Experimental design is a critically important tool in the scientific and engineering world for improving the product realization process. Critical components of these activities are in new manufacturing process design and development, and process management. The application of experimental design techniques early in process development can result in

1. Improved process yields
2. Reduced variability and closer conformance to nominal or target requirements
3. Reduced development time
4. Reduced overall costs.

Experimental design methods are also of fundamental importance in engineering design activities, where new products are developed and existing ones improved. Some applications of experimental design in engineering design include

1. Evaluation and comparison of basic design configurations
2. Evaluation of material alternatives
3. Selection of design parameters so that the product will work well under a wide variety of field conditions, that is, so that the product is robust
4. Determination of key product design parameters that impact product performance
5. Formulation of new products.

The use of experimental design in product realization can result in products that are easier to manufacture and that have enhanced field performance and reliability, lower product cost, and shorter product design and development time. Designed experiments also have extensive applications in marketing, market research, transactional and service operations and general business operations. We now present several examples that illustrate some of these ideas.

# Guidelines for Designing Experiments

To use the statistical approach in designing and analyzing an experiment, it is necessary for everyone involved in the experiment to have a clear idea in advance of exactly what is to be studied, how the data are to be collected, and at least a qualitative understanding of how these data are to be analyzed. An outline of the recommended procedure is shown in Table 1.1. We now give a brief discussion of this outline and elaborate on some of the key points. For more details, see Coleman and Montgomery (1993), and the references therein. The supplemental text material for this chapter is also useful.

✓ 1. *Recognition of and statement of the problem.* This may seem to be a rather obvious point, but in practice often neither it is simple to realize that a problem requiring experimentation exists, nor is it simple to develop a clear and generally accepted statement of this problem. It is necessary to develop all ideas about the objectives of the experiment. Usually, it is important to solicit input from all concerned parties: engineering, quality assurance, manufacturing, marketing, management, customer, and operating personnel (who usually have much insight and who are too often ignored). For this reason, a team approach to designing experiments is recommended.

(It is usually helpful to prepare a list of specific problems or questions that are to be addressed by the experiment. A clear statement of the problem often contributes substantially to better understanding of the phenomenon being studied and the final solution of the problem.) It is also important to keep the overall objective in mind; for example, is this a new process or system—in which case the initial objective is likely to be **characterization** or **factor screening**—or is it a mature or reasonably well-understood system that has been previously characterized—in which case the objective may be **optimization**? (There are many possible objectives of an experiment, including **confirmation** (Is the system performing the same way now that it did in the past?), **discovery** (What happens if we explore new materials, variables, operating conditions, etc.?), and **stability or robustness**.) (Under what conditions do the response variables of interest seriously degrade? Or, how can we reduce the variability in the response variable that arises from sources that we cannot directly control?). Obviously, the specific questions to be addressed in the experiment relate directly to the overall objectives. An important aspect of problem formulation is the recognition that one large comprehensive experiment is unlikely to answer the key questions satisfactorily. A single comprehensive experiment requires the experimenters to know the answers to a lot of questions, and if they are wrong, the results will be disappointing. This leads to wasting time, materials, and other resources and may result in never answering the original research questions satisfactorily. A

■ TABLE 1.1  
Guidelines for Designing an Experiment

1. Recognition of and statement of the problem	] Pre-experimental planning
2. Selection of the response variable"	
3. Choice of factors, levels, and ranges"	
4. Choice of experimental design	
5. Performing the experiment	
6. Statistical analysis of the data	
7. Conclusions and recommendations	

sequential approach employing a series of smaller experiments, each with a specific objective, such as factor screening, is a better strategy.

2. **Selection of the response variable.** (In selecting the response variable, the experimenter should be certain that this variable really provides useful information about the process under study. Most often, the average or standard deviation (or both) of the measured characteristic will be the response variable.) Multiple responses are not unusual. Gauge capability (or measurement error) is also an important factor. If gauge capability is inadequate, only relatively large factor effects will be detected by the experiment or perhaps additional replication will be required. (In some situations where gauge capability is poor, the experimenter may decide to measure each experimental unit several times and use the average of the repeated measurements as the observed response.) It is usually critically important to identify issues related to defining the responses of interest and how they are to be measured *before* conducting the experiment. Sometimes designed experiments are employed to study and improve the performance of measurement systems. For an example, see Chapter 12.
- ✓ 3. **Choice of factors, levels, and range.** (As noted in Table 1.1, steps 2 and 3 are often done simultaneously or in the reverse order.) (When considering the factors that may influence the performance of a process or system, the experimenter usually discovers that these factors can be classified as either **potential design factors** or **nuisance factors**. The potential design factors are those factors that the experimenter may wish to vary in the experiment. Often we find that there are a lot of potential design factors, and some further classification of them is helpful. Some useful classifications are **design factors**, **held-constant factors**, and **allowed-to-vary factors**. The design factors are the factors actually selected for study in the experiment. Held-constant factors are variables that may exert some effect on the response) but for purposes of the present experiment these factors are not of interest, so they will be held at a specific level. For example, in an etching experiment in the semiconductor industry, there may be an effect that is unique to the specific plasma etch tool used in the experiment. However, this factor would be very difficult to vary in an experiment, so the experimenter may decide to perform all experimental runs on one particular (ideally "typical") etcher. Thus, this factor has been held constant. As an example of allowed-to-vary factors, the experimental units or the "materials" to which the design factors are applied are usually nonhomogeneous, yet we often ignore this unit-to-unit variability and rely on randomization to balance out any material or experimental unit effect. We often assume that the effects of held-constant factors and allowed-to-vary factors are relatively small.

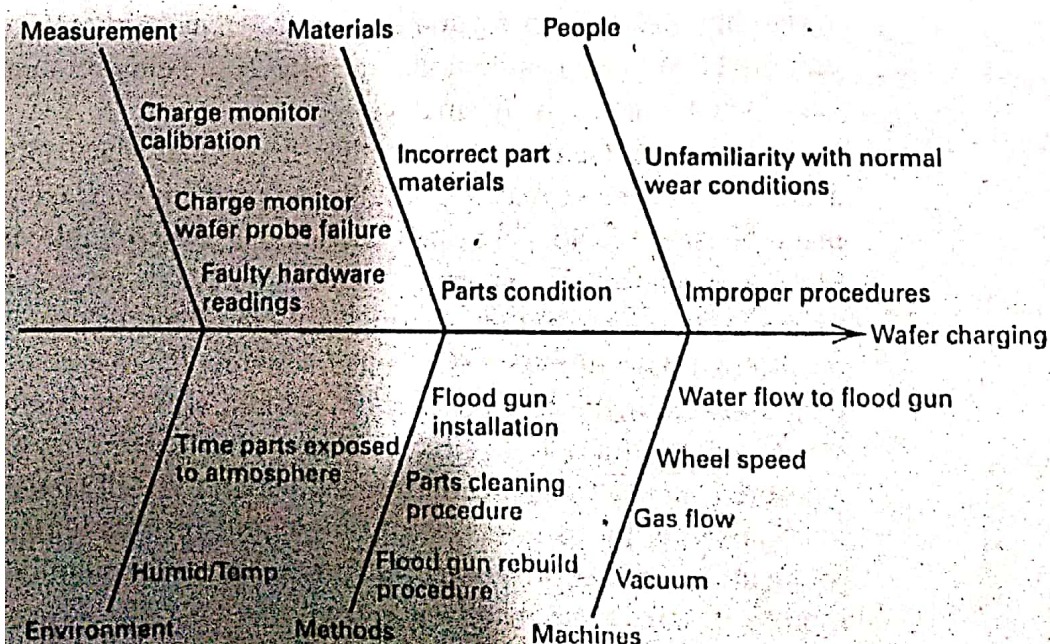
Nuisance factors, on the other hand, may have large effects that must be accounted for, yet we may not be interested in them in the context of the present experiment. Nuisance factors are often classified as **controllable**, **uncontrollable**, or **noise factors**. A controllable nuisance factor is one whose levels may be set by the experimenter. For example, the experimenter can select different batches of raw material or different days of the week when conducting the experiment. The blocking principle, discussed in the previous section, is often useful in dealing with controllable nuisance factors. If a nuisance factor is uncontrollable in the experiment, but it can be measured, an analysis procedure called the **analysis of covariance** can often be used to compensate for its effect. For example, the relative humidity in the process environment may affect process performance, and if the humidity cannot be controlled, it probably can be measured and treated as a covariate. When a factor that varies naturally and uncontrollably in the process can be controlled for purposes of an experiment, we often call it a **noise factor**. In such situations, our objective is usually to find

the settings of the controllable design factors that minimize the variability transmitted from the noise factors. This is sometimes called a process robustness study or a robust design problem. Blocking, analysis of covariance, and process robustness studies are discussed later in the text.

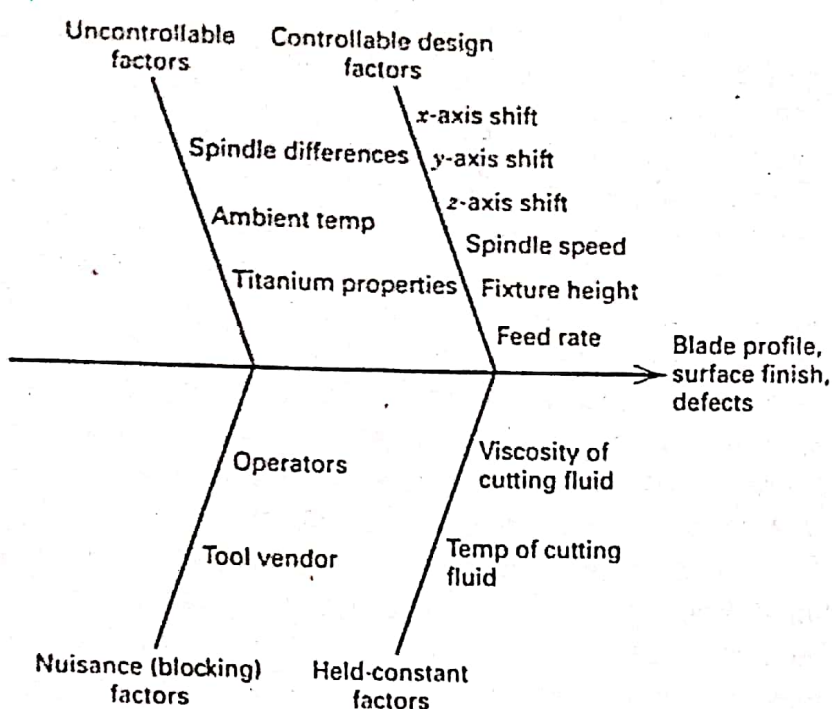
Once the experimenter has selected the design factors, he or she must choose the ranges over which these factors will be varied and the specific levels at which runs will be made. Thought must also be given to how these factors are to be controlled at the desired values and how they are to be measured. For instance, in the flow solder experiment, the engineer has defined 12 variables that may affect the occurrence of solder defects. The experimenter will also have to decide on a region of interest for each variable (that is, the range over which each factor will be varied) and on how many levels of each variable to use. **Process knowledge** is required to do this. This process knowledge is usually a combination of practical experience and theoretical understanding. It is important to investigate all factors that may be of importance and to be not overly influenced by past experience, particularly when we are in the early stages of experimentation or when the process is not very mature.

When the objective of the experiment is **factor screening** or **process characterization**, it is usually best to keep the number of factor levels low. Generally, two levels work very well in factor screening studies. Choosing the region of interest is also important. In factor screening, the region of interest should be relatively large—that is, the range over which the factors are varied should be broad. As we learn more about which variables are important and which levels produce the best results, the region of interest will usually become narrower.

The **cause-and-effect diagram** can be a useful technique for organizing some of the information generated in pre-experimental planning. Figure 1.10 is the cause-and-effect diagram constructed while planning an experiment to resolve problems with wafer charging (a charge accumulation on the wafers) encountered in an etching tool used in semiconductor manufacturing. The cause-and-effect diagram is also known as a **fishbone diagram** because the “effect” of interest or the response variable is drawn along the spine of the diagram and the potential causes or design factors are



■ **FIGURE 1.10** A cause-and-effect diagram for the etching process



■ FIGURE 1.11 A cause-and-effect diagram for the CNC machine experiment

organized in a series of ribs. The cause-and-effect diagram uses the traditional causes of measurement, materials, people, environment, methods, and machines to organize the information and potential design factors. Notice that some of the individual causes will probably lead directly to a design factor that will be included in the experiment (such as wheel speed, gas flow, and vacuum), while others represent potential areas that will need further study to turn them into design factors (such as operators following improper procedures), and still others will probably lead to either factors that will be held constant during the experiment or blocked (such as temperature and relative humidity). Figure 1.11 is a cause-and-effect diagram for an experiment to study the effect of several factors on the turbine blades produced on a computer-numerical-controlled (CNC) machine. This experiment has three response variables: blade profile, blade surface finish, and surface finish defects in the finished blade. The causes are organized into groups of controllable factors from which the design factors for the experiment may be selected, uncontrollable factors whose effects will probably be balanced out by randomization, nuisance factors that may be blocked, and factors that may be held constant when the experiment is conducted. It is not unusual for experimenters to construct several different cause-and-effect diagrams to assist and guide them during preexperimental planning. For more information on the CNC machine experiment and further discussion of graphical methods that are useful in preexperimental planning, see the supplemental text material for this chapter.

We reiterate how crucial it is to bring out all points of view and process information in steps 1 through 3. We refer to this as **pre-experimental planning**. Coleman and Montgomery (1993) provide worksheets that can be useful in pre-experimental planning. Also see the supplemental text material for more details and an example of using these worksheets. It is unlikely that one person has all the knowledge required to do this adequately in many situations. Therefore, we strongly argue for a team effort in planning the experiment. Most of your success will hinge on how well the pre-experimental planning is done.

✓ **A. Choice of experimental design.** If the above pre-experimental planning activities are done correctly, this step is relatively easy. Choice of design involves consideration of

sample size (number of replicates), selection of a suitable run order for the experimental trials, and determination of whether or not blocking or other randomization restrictions are involved. This book discusses some of the more important types of experimental designs, and it can ultimately be used as a catalog for selecting an appropriate experimental design for a wide variety of problems.

There are also several interactive statistical software packages that support this phase of experimental design. The experimenter can enter information about the number of factors, levels, and ranges, and these programs will either present a selection of designs for consideration or recommend a particular design. (We prefer to see several alternatives instead of relying on a computer recommendation in most cases.) These programs will usually also provide a worksheet (with the order of the runs randomized) for use in conducting the experiment.

In selecting the design, it is important to keep the experimental objectives in mind. In many engineering experiments, we already know at the outset that some of the factor levels will result in different values for the response. Consequently, we are interested in identifying *which* factors cause this difference and in estimating the *magnitude* of the response change. In other situations, we may be more interested in verifying uniformity. For example, two production conditions A and B may be compared, A being the standard and B being a more cost-effective alternative. The experimenter will then be interested in demonstrating that, say, there is no difference in yield between the two conditions.

✓5. *Performing the experiment.* (When running the experiment, it is vital to monitor the process carefully to ensure that everything is being done according to plan. Errors in experimental procedure at this stage will usually destroy experimental validity. Up-front planning is crucial to success. It is easy to underestimate the logistical and planning aspects of running a designed experiment in a complex manufacturing or research and development environment.

Coleman and Montgomery (1993) suggest that prior to conducting the experiment a few trial runs or pilot runs are often helpful. These runs provide information about consistency of experimental material, a check on the measurement system, a rough idea of experimental error, and a chance to practice the overall experimental technique. This also provides an opportunity to revisit the decisions made in steps 1-4, if necessary.

✓6. *Statistical analysis of the data.* (Statistical methods should be used to analyze the data so that results and conclusions are objective rather than judgmental in nature. If the experiment has been designed correctly and performed according to the design, the statistical methods required are not elaborate.) There are many excellent software packages designed to assist in data analysis, and many of the programs used in step 4 to select the design provide a seamless, direct interface to the statistical analysis. Often we find that simple graphical methods play an important role in data analysis and interpretation. Because many of the questions that the experimenter wants to answer can be cast into an hypothesis-testing framework, hypothesis testing and confidence interval estimation procedures are very useful in analyzing data from a designed experiment. It is also usually very helpful to present the results of many experiments in terms of an empirical model, that is, an equation derived from the data that express the relationship between the response and the important design factors. Residual analysis and model adequacy checking are also important analysis techniques. We will discuss these issues in detail later.

Remember that statistical methods cannot prove that a factor (or factors) has a particular effect. They only provide guidelines as to the reliability and validity of

results. When properly applied, statistical methods do not allow anything to be proved experimentally, but they do allow us to measure the likely error in a conclusion or to attach a level of confidence to a statement. The primary advantage of statistical methods is that they add objectivity to the decision-making process. Statistical techniques coupled with good engineering or process knowledge and common sense will usually lead to sound conclusions.

✓ 7. **Conclusions and recommendations.** Once the data have been analyzed, the experimenter must draw *practical conclusions* about the results and recommend a course of action. Graphical methods are often useful in this stage, particularly in presenting the results to others. **Follow-up runs and confirmation testing** should also be performed to validate the conclusions from the experiment.

Throughout this entire process, it is important to keep in mind that experimentation is an important part of the learning process, where we tentatively formulate hypotheses about a system, perform experiments to investigate these hypotheses, and on the basis of the results formulate new hypotheses, and so on. This suggests that experimentation is **iterative**. It is usually a major mistake to design a single, large, comprehensive experiment at the start of a study. A successful experiment requires knowledge of the important factors, the ranges over which these factors should be varied, the appropriate number of levels to use, and the proper units of measurement for these variables. Generally, we do not perfectly know the answers to these questions, but we learn about them as we go along. As an experimental program progresses, we often drop some input variables, add others, change the region of exploration for some factors, or add new response variables. Consequently, we usually experiment **sequentially**, and as a general rule, no more than about 25 percent of the available resources should be invested in the first experiment. This will ensure that sufficient resources are available to perform confirmation runs and ultimately accomplish the final objective of the experiment.