Section II

Engineering Applications

PART A
Process Capability Studies

This part of the Handbook covers the theory and mechanics of the Process Capability Study. It starts with the selection of a problem to work on. The problem is translated into statistical terms. The problem is then solved statistically by following a definite set of procedures. Finally, the solution is translated back into the original terms.

The Process Capability Study is a basic technique for analyzing data. It can be used for any type of data obtained from a production process. It can be made by an engineer, a supervisor or anyone else having responsibility for the job. Primarily, however, the Process Capability Study is a research technique and, as such, it is particularly important in all fields of Engineering.

Process Capability Studies are also the foundation of all shop applications of quality control, and many studies are made jointly by Quality Control Teams.

An example of a Process Capability Study is given on pages 66–72.

A-1 THE SCIENTIFIC FOUNDATION OF A PROCESS CAPABILITY STUDY

A-1.1 Definitions and terms

Process

The term "process" refers to any system of causes; any combination of conditions which work together to produce a given result. While it often refers to the combination of men, materials, machines and methods used to manufacture a given product, it is also capable of taking on other meanings as explained on pages 3–4. The process to be studied may be as simple as the motion of a hand about the wrist. It may be as complex as the complete set of operations in the plant.

Process capability

The term "process capability" refers to the normal behavior of a process when operating in a state of statistical control; the predictable series of effects produced by a process when allowed to operate without interference from outside causes. In manufacturing terminology, process capability refers to the inherent ability of the process to turn out similar parts; the best distribution that can be maintained in statistical control for a sustained period of time under a given set of conditions.

Process capability may be expressed as percent defective or as a distribution. In the latter case it refers to a single distribution with an irreducible spread (where "irreducible" means not reducible economically).

The "capability" of a process is not the same thing as its "performance," since performance may include all sorts of unnecessary variables and undesirable disturbances in the cause system. Capability means the natural or undisturbed performance after extraneous influences are eliminated. This is determined by plotting data on a control chart.

Process capability study

A "process capability study" is a scientific systematic procedure for determining the capability of a process by means of control charts, and if necessary changing the process to obtain a better capability. This procedure is continued as long as may be necessary until the problem which prompted the study is solved. A process capability study is sometimes described as "an industrial investigation whereby demonstrably true answers are found for one limited question after another until enough answers are found to make further questions unnecessary." In any case the term "process capability study" implies the solution of problems.

A-1.2 Scope of this technique

The field of application for process capability studies is very wide. They can be applied to almost any problem in Management, Engineering, Manufacturing or Inspection. The problems in these areas can ordinarily be reduced to those involving quality, cost, the need for new knowledge or information, the establishment of standards or estimates, new development and research.

The following is a list of typical problems which can be solved through the proper application of process capability studies. The list is included here to indicate the broad scope of this technique. The items on this list are not intended to be mutually exclusive, but rather to show the many different forms in which basically similar problems may present themselves. All of these problems are related essentially to the nature and behavior of a *distribution*. They can be solved through process capability studies because these studies provide a method for analyzing and changing distributions.

(1) Quality

- Too many defects leaving Operating.
- Too many defects leaving Inspection.
- Product unstable or drifting.
- Wrong distribution (in case of distribution requirements).
- Bad piece parts or material coming into the assembly line.

(2) Cost

- Too much inspection.
- Too much adjusting.
- Too much repair or rework.
- Too much scrap.
- Excessive merchandise losses.
- Trouble in meeting schedules.
- Low yield.

(3) Information

The need to:

- Trace the causes of trouble.
- Find why things happen.
- Discover correlations.
- Find how the early characteristics affect the end product.
- Find which dimensions are important.
- Obtain new knowledge about materials, methods, testing equipment, type of product.
- Get reliable information from pilot runs.
- Find the capability of new tools, methods, machines.
- Compare designs, tools, assembly methods.
- Study the effect of engineering changes.
- Study the effect of going over to a new design.
- Obtain continuity on intermittent operations.
- Make sense out of engineering data.
- Interpret the results of engineering experiments.
- Find the degree of training of operators.
- Test for significant differences.
- Detect trends.
- Determine whether conditions are constant.
- Check on error of measurement.
- Find assignable causes.
- Keep from being deceived by statistical fluctuations.

(4) Standards

- Estimates to be used for engineering purposes:
  - Wage incentives.
  - Standard costs.
  - Normal amount of inspection.
  - Normal losses.
  - Normal yields.
  - Normal capacity.
  - Normal amount of sorting by Operating.
  - Overall capability.
  - Machine capability.
  - Natural tolerances.
- Specifications of all kinds.
- Reliability of test sets, gages, and other standards.
- Maintenance schedules.
- Engineering responsibility vs. shop responsibility, etc.

(5) New development

- New products.
- New methods.
- Cost reduction.
- Automation.
- Machine and tool design.
- Purchase of new types of machines, test sets, etc.
- Elimination of difficult or expensive operations.

A-1.3 The scientific steps in experiment

A process capability study follows the method of scientific research. This method may be said to consist of four steps, as follows:

(1) Experiment.
(2) Hypothesis.
(3) Test of hypothesis.
(4) Further experiment.

The *experiment* consists of collecting observations from the process at several different points in time.

The *hypothesis* is that the observations, if they came from a stable process, should exhibit only natural fluctuations.

The *hypothesis is tested* by making a control chart and testing its pattern for naturalness. If the pattern is unnatural some "assignable cause" was interfering with the normal process. The cause is identified by proper study of the pattern, and its effect on the problem is traced.

Finally, depending on what the patterns show, it may be necessary to change the process, collect more data, revise the specification, or in some other way perform a *further experiment.*

Figure 51 shows the four steps in a process capability study expressed in slightly different terms.

Repetition of these four steps

If the person making the study finds a complete solution at the end of Step 4, he concludes the study there. Frequently, however, he finds that the initial study is merely one step in arriving at a solution. In that case Step 4 becomes the first step in a second study.

An example of this is given on pages 66–72.

A-2 OBTAINING THE DATA

Some processes can be studied by obtaining data on the process directly: for example, variation in a test set, or changes in the heat treating temperature in an oven. Others can be studied by observing the effect of the process on the product: for example, the diameter of a piece part or the presence of defects in an assembly. In either case we begin a process capability study by obtaining a series of measurements or by accumulating percent defective data.

In any case where there is a choice between variables and attributes measurements, it is worth going to considerable trouble to devise some method of obtaining actual readings. This will make it possible to take advantage of the great sensitivity of the $\bar{X}$ and R chart. In some cases where it is not possible to obtain true variables measurements, it may be possible to get "semi-variables" measurements as explained in paragraph A-2.5.

A-2.1 Where and how to use variables control charts

An $\bar{X}$ and R chart requires less data for the same amount of information than any other control chart. This is the type of chart to use when it is difficult or expensive to take measurements, where the test is destructive, or where it is desired to get the maximum amount of in-

Scientific Experimentation	*Process Capability Study*
1. Experiment.	1. Collect data from the process.
2. Hypothesis.	2. Plot statistical patterns.
3. Test of hypothesis.	3. Interpret the patterns.
4. Further experiment.	4. Do what the patterns tell you until you reach the process capability.

Fig. 51. The four steps in a process capability study.

formation with the least amount of effort. In addition, $\bar{X}$ and R charts have two advantages which are not possessed by any others:

(1) *Different kinds of trouble show up in different ways on these charts.*

For example, a wrong machine setting shows up on the $\bar{X}$ chart, while a machine in need of repair shows up on the R chart. In a similar way, various causes of trouble can be distinguished in the case of assembly operations, chemical processing, and so on. The $\bar{X}$ and R chart is the best one to use for getting answers to questions like these: Why aren't we getting consistent results? What could be causing so much trouble at 450 cycles? What can be done to improve this process and make it behave better? In general, the newer the job or the more there is to learn about a given type of product, the more it will be necessary to use $\bar{X}$ and R charts.

(2) *$\bar{X}$ and R charts make it possible to study the process without regard to the specification.*

This is not true of a p-chart. A p-chart starts with the specification and simply records failures to meet it.

$\bar{X}$ and R charts start with the process itself and give an independent picture of what the process can do. Afterward the process may be compared with the specification or not, depending on the problem. For this reason, $\bar{X}$ and R charts can be used to *obtain changes in specifications* and bring about the establishment of more realistic limits. The more cases we have where it is suspected that the specifications may need to be changed, the more it will be necessary to use $\bar{X}$ and R charts.

$\bar{X}$ and R charts are at their best when used at early operations, close to the causes that may affect later results. Use $\bar{X}$ and R charts on individual characteristics, operators, machines, machine setters, shifts, sources of voltage supply, etc. These charts are much less effective when used at the end of the line on final tests.

A-2.2 Where and how to use attributes control charts

A p-chart requires larger samples than an $\bar{X}$ and R chart. It is less versatile and less sensitive than an $\bar{X}$ and R chart for the following reasons:

- The p-chart cannot tell whether trouble is caused by lack of control of the average value of a characteristic; or by the fact that it is located too close to a specification; or by an uncontrolled process spread; or by a spread that is controlled but is too wide for the specification.
- The p-chart cannot warn of shifts or trends in the process unless those trends have proceeded so far that they have actually resulted in defectives.

On the other hand, a p-chart often has the advantage of using records which are already available in the shop. It is generally necessary to obtain special data for an $\bar{X}$ and R chart.

One common use for p-charts is to study an entire assembly process by means of an overall chart. The p-chart can be made to cover all defects and all characteristics, combined in a single percentage. This kind of chart can be a valuable capability study in itself, and will also provide a good measure of the effectiveness of changes, corrections or improvements which have been made as a result of other studies.

When used alone, however, p-charts on the overall process are often difficult to interpret. The causes for unnatural patterns may be so deeply hidden that it is not possible to find them in the overall data. One of the standard ways of interpreting p-charts is to break them down into individual sources or individual defects. If interpretation is still difficult, use an $\bar{X}$ and R chart.

p-Charts may be used for:

(1) Characteristics on which it is difficult or impractical to obtain variables measurements.

(2) Studies of defects produced by machines or operators which are directly under the machine setter's or operator's control.

(3) Direct studies of the amount of dropouts, shrinkage or scrap at specific operations.

Collect p-chart data, where possible, on the work of individual operators or individual machines.

In this Handbook, unless otherwise stated, whatever is said about p-charts should be taken to apply also to np-charts, c-charts and u-charts.

A-2.3 Precautions in obtaining the measurements

In making a process capability study it is necessary to plan carefully for the proper collection of data. The following rules are based on the experience of many engineers.

(1) Take the data, if possible, in the same time-sequence in which the product is made.

(2) Arrange to take data on the product as made rather than after a screening or adjusting operation (unless the object of study is to be the screening or adjusting.) In the latter case you may wish to take data both before and after the screening or adjusting.

(3) Decide in advance on the proper technique for making the measurements.

(4) Decide how many measurements should be taken on each part and exactly where the measurements should be made.

(5) See that the proper identification is recorded in addition to the actual measurements. For example, time of day, number of machine, name of operator, number of test set, number of gage, etc. See paragraph A-2.9.

(6) Instruct the person taking the measurements to make a note of all known changes in the process during the period of study.

(7) If the data are to be taken on product and there is more than one source of product (for example, more than one machine, operator or test set) decide whether to cover all of these sources or only one or two.

(8) If the data are to be taken on processing conditions and there is more than one set of conditions, decide whether to cover all or only one or two.

A-2.4 Error of measurement

It is not possible to obtain the full advantages of the process capability technique unless the measurements are reliable to start with. This means that the measurements must be taken accurately, and at a point that is meaningful in its bearing on the problem being studied. It is very common in capability studies to find patterns that are seriously out of control when the first readings are plotted. Often this turns out to be largely the error or instability of the measurements rather than the actual condition of the product.

Remember that every observation on a piece of product is a composite of two different elements. One is the actual value of the characteristic; the other is the measurement of it. If the measurement is contributing more variability than the pieces of product, it will be difficult to detect some of the cause and effect relationships which may be important in solving the problem.

In making a capability study either obtain the measurements yourself or make sure that they are taken by someone who is properly instructed and in whom you have confidence. If there is doubt as to the adequacy of the method of measurement, it may be necessary to make a study of the measuring method itself before attempting to study the variations in the product.

It is possible to check (a) the accuracy of the measurements as compared to a fixed standard, and (b) their precision or reproducibility. For information on this see pages 84–91.

A-2.5 Semi-variables measurements

In cases where it would be desirable to have variables measurements but the characteristic is one which is ordinarily checked by attributes only, it is often possible to obtain "semi-variables" measurements by setting up a scale which is capable of showing degree, and ranking the units in accordance with this scale. For example, in a problem involving burrs it might be possible to use the following:

Size of Burr	*Artificial Number*
No burr	0
Small burr	1
Medium burr	2
Large burr	3
Very large burr	4

The numbers 0, 1, 2, 3 and 4 can be used to make $\bar{X}$ and R charts like any ordinary measurements.

Before attempting to assign numbers to the units, establish standards to make your judgment as consistent as possible. The "error of measurement" in semi-variables data can be reduced by obtaining separate observations or rankings from a number of independent observers. The error of measurement can be evaluated in the same way as any other error of measurement.

A-2.6 "Work sampling" measurements

It is possible to make process capability studies on such problems as the required frequency of machine settings or housekeeping activities in the shop, the ways in which a group of clerical workers or engineers spend their time, whether the shop is properly performing the specified chemical, washing, degreasing or heat treating operations, and similar situations where we do not ordinarily think of taking numerical measurements. To do this we make use of a special data-collecting technique which is known as "work sampling." This technique consists of the following:

(1) Prepare a check list of all the activities which it is desired to study. The list may include (a) desirable activities which we wish to encourage or on which we wish to set standards, and (*b*) undesirable activities which we wish to study for the purpose of reducing them or eliminating them from the process. It may also include activities which are essential but non-productive, such as waiting for work or changing the water in a wash tank. The key to the success of "work sampling" studies is the preparation of the check list.

(2) Have an observer go around at random intervals and record instantaneous observations of the activity being performed at that moment. The activity is recorded by making a tally mark on the check list. Randomness can be assured by having the observer draw a card from a shuffled deck of cards to determine the time at which he should make the observation, and a second card to determine the area or person to be observed. The number and frequency of the observations which must be taken depends on the nature of the process and the purpose of the process capability study.

The data obtained in this manner consist of a number of tally marks which constitute a "sample." The "sample size" is the total number of observations. Attributes data for *p*-charts or other purposes are obtained by taking the percentage of observations recorded for a given activity or group of activities as compared with the total number of observations in the sample.

Work sampling measurements can be used for process capability studies, shop control charts, or estimates based on samples. For further information on this method of obtaining data see References No. 1 and 22.

A-2.7 Amount of data required for a process capability study

$\bar{X}$ and R charts

The process capability study should cover at least three different periods in time. A suitable amount of data to start with would be as follows:

First period	50 measurements
Second period	25 measurements
Third period	25 measurements

This is a total of 100 measurements on the process.

p-Charts

Again we want the study to cover three or more periods in time. A suitable amount of data would consist of 20 to 25 samples for each one of these periods. Each sample should represent about 50 or 100 units checked.

Charts for individual measurements with control limits based on the moving range

In this type of study we are ordinarily limited to very little data. It is permissible to use as few as 10 consecutive numbers, provided these numbers cover a representative period of time. For example, if the chart is to cover merchandise losses or other accounting figures, the study might cover a period of approximately one year.

Special cases

While the rules given above are safe rules to follow, do not hesitate to use larger or smaller amounts of data if this is necessary or convenient. If more data are readily available, by all means use them. On the other hand, if very little data can be taken, it is still possible to obtain answers to many problems by using these data in a capability study.

A-2.8 Selection of samples

Samples should be selected so as to minimize all sources of variation other than the factor being studied. This can usually be accomplished by having each sub-group consist of consecutive units as produced. See page 151 for further information on suitable methods of selecting samples.

A-2.9 Identification of data

When collecting measurements, carefully identify the different periods in time. Also identify any other known changes in the source of data, or any surrounding conditions or elements in the process which might be able to affect the results. For example: Up to a certain point the data came from one location but after that the job was moved to another; at a certain point the design was changed; there was a new supervisor; the job was put on sampling inspection; the engineer decided to use a different furnace.

The more ways in which the data are identified, the more it will be possible to learn from the process capability study. This is one of the first principles in control chart analysis and is the source of much of its power. This applies not only to process capability studies but also to designed experiments.

A-2.10 Point at which the data should be collected

Data may be collected at any point where the problem is thought to exist. However, many troubles which first become apparent in the end product have roots which go back to the early operations. It is often possible to save time by studying the early operations in the first place.

A-3 ANALYZING THE DATA

A-3.1 Scales for plotting

Scales should be chosen in such a way as to get a readable width for the control limits. The control limits should preferably be not less than 1 inch and not more than 2 inches apart. For $\bar{X}$ and R charts, try to keep the width of the control band (that is, the distance between the centerline and either control limit) approximately equal for both $\bar{X}$ and R. For p-charts which are to be compared, use the same scale on all charts.

A-3.2 Calculating limits

When the data included in a capability study cover more than one set of conditions, some question may arise as to whether the centerline and control limits should be calculated from all the data or from one portion of it only. It is an important principle in capability studies that in any case where there are adequate amounts of data it does not matter which portion of it is used for calculations. A different selection of data may result in different control limits but should not result in different conclusions. This can be seen in Figure 52 on page 52. The data plotted here might represent an $\bar{X}$ chart, a p-chart or a chart for individual measurements.

In this example the pattern without control limits is shown in Figure 52(a). If control limits are calculated from the May data only, the chart will look like Figure 52(b). This chart shows that May and June are different, May being higher. If the control limits are based on the month of June only, the chart will look like Figure 52(c). This is a different centerline and set of control limits but it tells the same story: May is higher than June. If the control limits are based on May and June combined, this results in a third set of limits but the story is still the same. This will be true, in general, as long as there are adequate amounts of data on which to base the calculations.

When the data are limited, as in a designed experiment, it is best to use all the data in calculating the control limits. The data are considered to be "limited" if there are less than 20 points when the chart is plotted. The smaller

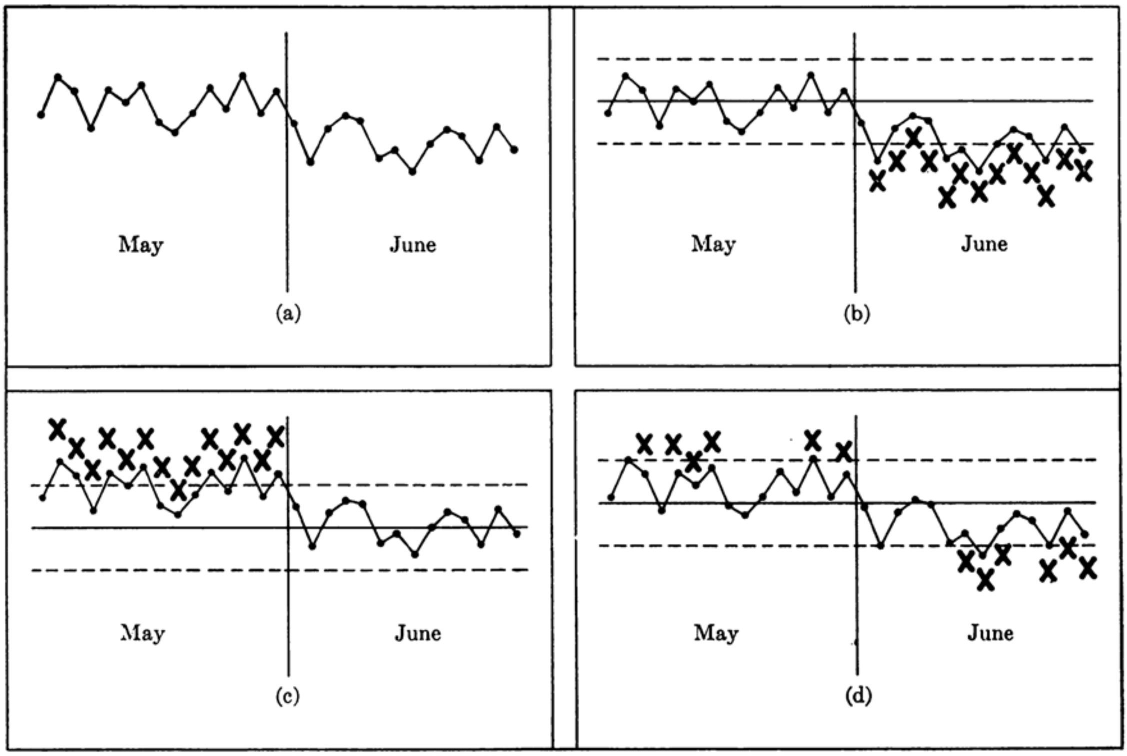

Fig. 52. Different levels of performance in May as compared with June.

the number of measurements, the less certain we are about the accurate location of the limits.

Addition of further data

The rules given above cover the initial calculation of control limits for a capability study. When more data are to be added, extend the original control limits across the page and plot the additional data against the original limits.

If you find on checking the patterns that the additional data are significantly different from the original data, it is proper to calculate new control limits, if desired, for the additional data. The new limits will make it possible to test the later set of data to see whether it shows control with respect to its own limits.

Do not, however, calculate new limits for additional data unless you have previously found a significant difference. Do not consider there is a significant difference unless there is reason to mark x's on the chart. If new limits are calculated without statistical evidence that the sets of data were different, this may easily lead to false conclusions. Separate control limits may make a set of data appear different when in fact it has come from the same system of causes.

A-3.3 What to do about freaks

Engineers frequently wonder whether they ought to throw away portions of the data that appear to be "wild" or to have been caused by a "freak." In general, it is never wise to throw away any data in the calculations or plotting unless there is definite knowledge that something has affected that one group of readings and no others. Ordinarily it is difficult to obtain such positive knowledge. If supposed "freaks" are eliminated in the calculations or plotting, this may destroy the very information which might lead to solving the problem.

One exception to the above rule is a case encountered occasionally where the "wild" readings are very numerous. This is sometimes found in the early studies on new products or designs under development. When wild readings are numerous they should not properly be considered as freaks, but rather as a separate distribution resulting from a different system of causes. Such freaks should be separated from the rest of the data in order to make the patterns interpretable. They should not, however, be eliminated from the study. They should only be plotted on a separate chart.

It need hardly be pointed out that in any case where it is necessary to separate the freaks in this manner, one of the immediate objects of the capability study should be to identify the extraneous system of causes so that the freaks can be eliminated. See pages 162–164 for further help in dealing with freaks.

Repeated "freaks"

If freaks tend to occur in a capability study repeatedly, the repetition means that they are not freaks but rather the result of some cause operating regularly in the process. The principal rule in dealing with such freaks is: Don't underestimate their importance and don't eliminate them from the data.

A-3.4 Plotting the control charts

In all cases, follow the standard methods of plotting control charts when making a process capability study. If necessary review the material on pages 12–23.

Do not overlook the fact that the data can be plotted in more than one way. Plotting a chart of individual measurements or a grouped frequency distribution may help in interpreting an $\bar{X}$ and R chart. Changing the scale at the bottom of the chart (for example, to show operations rather than time) may help in interpreting a *p*-chart or *c*-chart.

A-3.5 Studying the patterns

Preliminary analysis of a process capability study consists of two steps as follows:

(1) Carefully record, on the chart, all known pertinent facts about the data as provided in paragraphs A-2.3 and A-2.9. These may be needed later for separating the data according to source. Also make notes on any surrounding conditions or elements which might be capable of affecting the results. Repair of a fixture, re-calibration of a test set, or the fact that maintenance work was performed may be vital information in securing a correct interpretation.

(2) Go over the patterns carefully and mark x's where they belong. If instability or other changes are indicated, determine the type of change. (For example: cycle, trend, gradual shift, sudden shift, erratic fluctuation, freak, interaction.) Mark the type of pattern on the chart near the corresponding x's. Note also any background information you may have which is related to the type of pattern. For example: "Trend—probably tool wear."

Be sure to include a check for stratification, stable mixture, systematic variables, tendency for one chart to be correlated with another. Note on the chart any background information related to these patterns. For example: "Stratification—may be due to samples being associated with bifurcated springs."

If the pattern is natural, make the calculations indicated in paragraph A-3.6. If the pattern is not natural, follow the directions in paragraph A-3.7.

Regardless of whether the patterns are natural or unnatural, a properly conducted process capability study should give valuable information. The chart will be used in different ways, however, depending on the naturalness or unnaturalness of the pattern.

A-3.6 Drawing conclusions from a natural pattern

The primary significance of a natural pattern is that it indicates a process which is in statistical control. Such a process is stable and undisturbed by extraneous causes. It tends to repeat itself day after day and is consequently predictable. It is possible to determine the underlying characteristics of such a process by making calculations based on the pattern on the control chart.

Among the calculations which may be based on a natural pattern are the following. Do not attempt to make any of these calculations unless the control chart is in control:

(1) Estimate the center of the process distribution as $\bar{\bar{X}}$.

(2) Estimate the shape of the process distribution by making a frequency plot of the process capability data.

(3) Estimate the spread of the process distribution by making the calculations shown on page 56.

(4) Compare the process distribution with the specification or other standards as shown on pages 119–122.

(5) Estimate the percentages outside of limits, if desired, by making the calculations shown on pages 58–59.

(6) Use the centerlines on the $\bar{X}$ chart and R chart, if desired, to calculate the effect of overlapping tolerances.

(7) Use the centerlines if desired to set up economic limits for shop control charts.

(8) Use the centerlines if desired to establish standards for budgets, forecasts, wage incentive allowances, etc.

In the case of a p-chart, estimate the capability as explained on page 59, and use the centerline as indicated in (7) and (8) above.

Note particularly that a natural pattern is essential when we wish to.

- Determine capability.
- Compare with standards.
- Generalize.
- Predict.

In the absence of a natural pattern we can obtain other information, but we cannot do the four things listed above.

For further information on the meaning of a natural pattern see pages 170–171.

A-3.7 Drawing conclusions from an unnatural pattern

An unnatural pattern indicates a process which is disturbed or out of control. Such a process may be erratic and unpredictable. It may or may not tend to repeat. We cannot determine the underlying characteristics of such a process by making calculations from the out-of-control data. We cannot use the out-of-control data to generalize or predict. We can, however, obtain other useful information in the manner shown below.

(1) In a process capability study the primary significance of an unnatural pattern is *the fact that important causes, capable of exerting a large effect on the process, are present in such a form that they are susceptible to analysis and study.* While natural patterns are used mainly for setting standards and making estimates, *unnatural patterns are used mainly for gaining new knowledge about the process.*

(2) Unnatural patterns provide information about processing variables, process changes, cause and effect relationships, cost reduction possibilities, and potentialities for improvement. The information contained in an unnatural pattern may be far more important, for engineering purposes, than the information contained in a natural pattern.

Method of analysis

Unnatural patterns may be divided into two general types:

(a) Relatively simple.
(b) Relatively complex.

Among the relatively simple types are the following:

Cycles.
Trends.
Sudden or gradual changes in level.
Certain types of systematic variation.

Among the relatively complex patterns are the following:

Mixtures of all kinds, including both stable and unstable mixtures.
Freaks.
Grouping or bunching.
Stratification.
Tendency of one chart to follow another.
Interaction.
Instability.

The simple unnatural patterns can ordinarily be interpreted by the application of technical knowledge or shop experience. The relatively complex patterns must usually be reduced to one of the simple forms before they can be interpreted.

A-3.8 Simplification of complex patterns

General

The basic approach, in simplifying complex patterns, is to separate the data according to various sources. In many studies there is an obvious basis for performing this separation. For example, there may be several machines, shifts, operators, tools, chucks, sources of supply, fixtures, heads, methods of assembly, positions at different spots in an oven. In such cases it is easy to separate the data and plot a separate chart for each source. See Method A below.

In other cases the method of separating the data is not so simple. For example, a machine may behave like more than one machine if it is in a poor state of repair. A fixture may behave like more than one fixture if it is poorly anchored or has excessive play. In the same way, an operator who is careless or inadequately trained may behave like more than one operator. Piece parts which are not uniform, or which come from a mixed lot sent in by a supplier, may behave like two kinds of piece parts. In such cases it requires more ingenuity to separate the data. See Method B below.

The following methods of separating data are used in engineering studies:

Method A. Simple breakdown.
Method B. Elimination of variables.
Method C. Rearrangements of data.
Method D. Designed experiments.

All of these methods have the same basic objective. They attempt to separate the data into significant categories in such a way that the patterns will become simple enough to interpret.

Method A: Simple breakdown

This method is used where some of the possible sources of complexity are known or at least suspected. The engineer proceeds as follows:

(1) Separate the data according to known sources, major components, etc. Plot separate charts for each source. The source or component whose pattern is least stable, or whose pattern is most similar to the original complex pattern, is the one most likely to contain the important causes. Disregard the rest of the data and concentrate on this portion.

(2) Take the data for this least stable portion and break it down further. Plot separate patterns for each of the new sources or subcomponents. Again, the source or component with the most significant pattern is the one most likely to contain the important causes.

(3) With each separation, the patterns become simpler or stand out more prominently. Continue this process until the patterns become natural or until they consist of (a) simple shifts in level or (b) simple trends. At this point it is possible to make calculations as shown in paragraph A-3.10.

Suggestions on possible sources or "production paths" will be found on pages 166–167, 168, 180 and 219. It will also be helpful to study the material on control chart theory on pages 149–151, and the explanation of the R chart on pages 154–156. If the patterns are still complex when separated by source, follow the directions under Method B.

Method B: Elimination of variables

This method is used where there is no prior basis for separating the data. An example of such a case is given on pages 66–71. The engineer proceeds as follows:

(1) The original pattern is used to discover some variable (usually a single variable) which needs to be eliminated. This requires a knowledge of control chart patterns and the ability to interpret these patterns as outlined in Engineering Part F. Study the explanations given on pages 66–71.

(2) As soon as the first variable is discovered, do what is necessary to eliminate this variable. Then collect more data and make, in effect, a second process capability study. The patterns in the second study will be simpler because of the removal of one of the large variables. See page 67.

(3) Continue this process until the patterns become natural or until they consist of (a) simple shifts in level or (b) simple trends. Then make the calculations shown in paragraph A-3.10.

This method of simplifying patterns can be used in any situation. It may be the only method possible if there are complicated variables or many engineering unknowns.

Methods C and D: Rearrangements of data and formal designed experiments

If the data used in the study have been

identified in enough different ways, it may be possible to simplify the pattern by merely rearranging the data. An example of this is given on pages 72–73. Correlation studies, scatter diagrams and "trend arrangements" of the measurements, and formal designed experiments, are all methods of classifying and rearranging the data.

These methods simplify the pattern by arranging the data in various ways, thus making it possible to identity certain causes. For correlation and similar studies see pages 143–148. For designed experiments see pages 75–117.

A-3.9 Checking to determine whether you have found the real cause

If the causes affecting the pattern have been properly identified, there should be an obvious correspondence between the presence or absence of the cause and subsequent changes in pattern. It should be possible to put the cause in or take it out at will and make the pattern behave correspondingly. In addition, there must be some logical engineering reason for believing that such a condition might be the cause. Be careful not to assume that one condition causes another merely because it precedes the other in time.

A-3.10 Calculations from a pattern showing only very simple shifts and trends

If the cause of the shift or trend has been conclusively identified, calculations may be made from simple patterns in much the same manner as in paragraph A-3.6. In the case of a shift, make the calculations separately for each distinct level. In the case of a trend, make the calculations for one or more levels along the slope of the trend line.

A-4 MAKING AN ESTIMATE OF THE PROCESS CAPABILITY

The capability of a process may be expressed numerically in two different ways:

(a) as a distribution having a certain center, shape and spread; or

(b) as a percentage outside of some specified limit.

In the first case the capability is estimated from an $\bar{X}$ and R chart; in the second, from a p-chart.

By a simple calculation, information from the $\bar{X}$ and R chart can also be translated into percentages. It is not possible to work in the opposite direction and get distribution information from a p-chart.

A-4.1 Estimating center, shape and spread from an $\bar{X}$ and R chart

At the conclusion of the process capability study, you have obtained a set of controlled patterns. From these patterns it is possible to find the distribution that the process is capable of producing.

- **Center.** The center of the distribution will be the centerline on the $\bar{X}$ chart ($\bar{\bar{X}}$).
- **Shape.** The shape can be judged for most practical purposes by making a frequency distribution of the data which produced the controlled patterns. Several hundred measurements may be necessary to give reliable evidence about the shape.
- **Spread.** The spread can be calculated by using the factors given on page 131. Proceed as follows:
  (a) For a normal* distribution, estimate the spread of individuals as $\pm 3\bar{R}/d_2$. Be sure to use the d_2 factor (from page 131) which corresponds to the sample size used to obtain R.
  (b) For a non-normal** distribution, estimate σ as $\bar{R}/d_2$. The distribution may spread more than 3 σ on one side and less than 3 σ on the other side. The total spread may be more or less than 6 σ.

Note that the estimate of spread, as well as shape, will be affected if the distribution is skewed.

A-4.2 Permanent and non-permanent skewness

In many processes the degree of skewness is subject to change even when the center and

* See pages 131–134.
** See pages 134–136.

spread of the distribution are reasonably constant. If you wish to base specifications or changes in specifications on the fact that the distribution is skewed, be sure that there is a good engineering reason to account for the skewness in the first place and also for arguing that it is a permanent feature of the process.

Some of the causes which tend to introduce skewness and make it permanent are:

(1) Parts coming up against a positive stop.

(2) Measurements reaching the physical limits of the material, as in strength of welds.

(3) Manual control of operations such as grinding, where the tendency is to grind just inside the limits in order to minimize work.

(4) Characteristics such as eccentricity, warpage, unbalance, runout, etc. where the natural limit is zero.

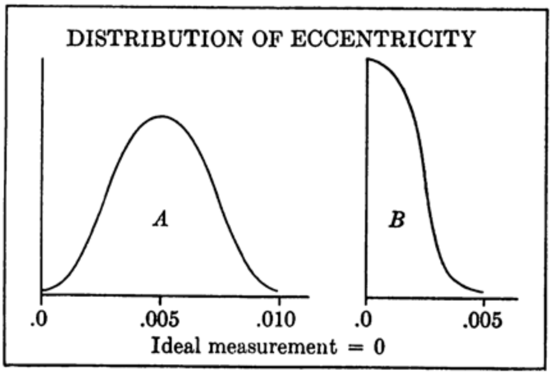

Fig. 53. *B* is a better distribution than *A*.

In many processes a skewed distribution is desirable. See Figure 53. Engineers should not jump to the conclusion that properly run processes, or processes in good control, will have "normal" distributions.

There are also certain mathematical reasons for skewness, an example of which is the following:

> Suppose we are making a product consisting of square pieces. We measure the sides of a large number of squares and find that their distribution is symmetrical.
>
> If the *sides* of the squares form a symmetrical distribution, it is obvious that the *areas* of the squares cannot form a symmetrical distribution. If we measured the areas (or volume or weights or anything related to the sides by a square or cubic relationship), the distribution would evidently be skewed. See Figure 54.

The above causes tend to introduce skewness as a permanent characteristic of the distribution. In addition, the following causes may introduce temporary skewness:

(1) Some skewness is created artificially by a sorting or screening operation. Such distributions are said to be "truncated." The truncation generally takes place at a specified maximum or minimum limit as in Figure 55.

(2) When skewness is associated with out-of-control patterns on a control chart, it is likely to be the result of a mixture of two

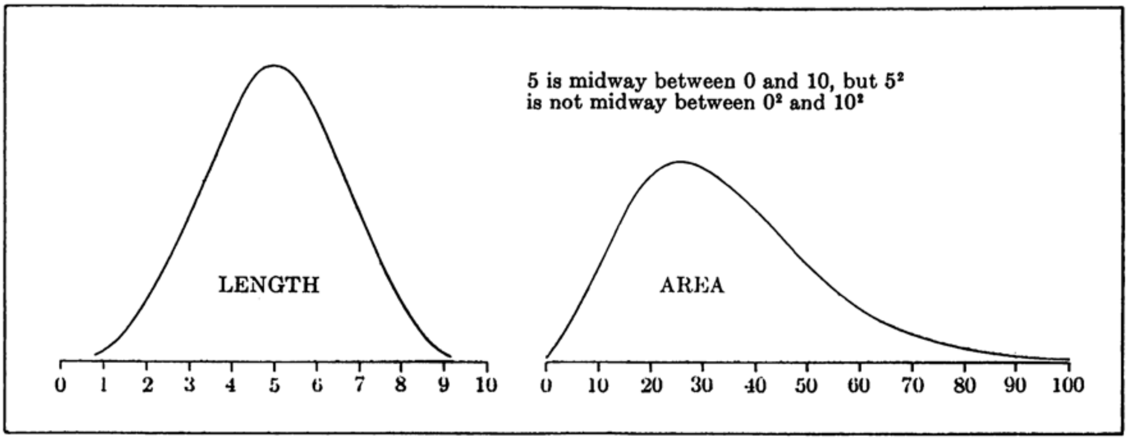

Fig. 54. If the distribution of length is symmetrical, the distribution of area cannot be.

or more distributions. This type of skewness tends to be nonpermanent and unstable. See Figure 56.

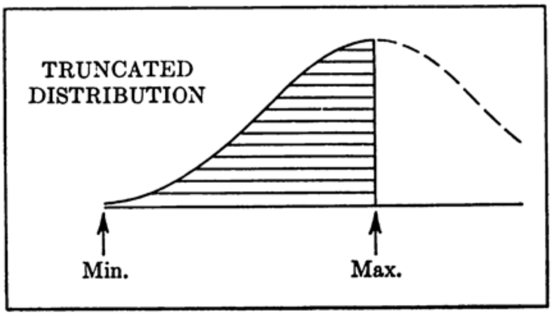

Fig. 55. Nonpermanent skewness due to screening.

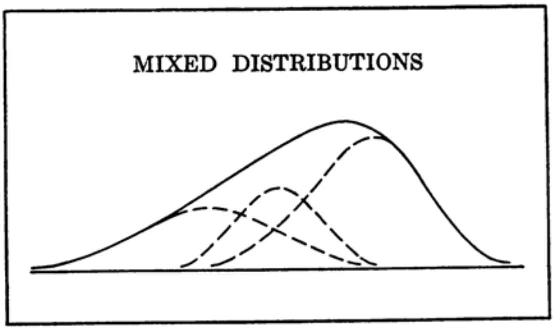

Fig. 56. Nonpermanent skewness due to mixture.

Occasionally skewness may be the result of "freaks" which cause a long tail to spread out on one side. In such cases, before deciding whether to treat the distribution as a skewed distribution, you must decide whether to consider the freaks as a regular part of the process.

A-4.3 Estimating percentages from an $\bar{\bar{X}}$ and R chart

The percentage outside of a given limit can be estimated from $\bar{X}$ and R charts that show control. This requires a knowledge of the distribution center, shape and spread. If you have reason to believe the distribution is reasonably normal (symmetrical, not too flat on top, and not too peaked), the percentage can be estimated with reasonable accuracy as follows:

(1) Estimate "sigma prime" from the centerline of the R chart.

$$\sigma' = \frac{\bar{R}}{d_2}$$

Values of d_2 are given on page 131.

(2) Calculate "t" according to one of the following formulas, depending on whether you are interested in a maximum or minimum limit.

$$t = \frac{\bar{\bar{X}} - \text{Max.}}{\sigma'}$$

$$t = \frac{\text{Min.} - \bar{\bar{X}}}{\sigma'}$$

(3) Read the percentage outside of limits from Table I on page 133.

Example

In a process capability study on thickness of a core plate, $\bar{R}$ for samples of five was found to be .0030 and $\bar{\bar{X}}$ was .7512. Both patterns were in control. The engineer felt that the distribution was "reasonably normal." He calculated sigma as follows:

$$\sigma' = \frac{\bar{R}}{d_2} = \frac{.0030}{2.326} = .0013$$

The specification for core plate thickness was .750 ± .003. Consider the two specification limits separately as follows.

Taking first the upper specification limit (.7530):

$$t = \frac{\bar{\bar{X}} - \text{Max.}}{\sigma'} = \frac{.7512 - .7530}{.0013} = \frac{-.0018}{.0013} = -1.38$$

Looking up −1.38 in the table on page 133 (under "Percentage Outside of Max."), we find the percentage is 8.4%.

Now considering the lower specification limit (.7470):

$$t = \frac{\text{Min.} - \bar{\bar{X}}}{\sigma'} = \frac{.7470 - .7512}{.0013} = \frac{-.0042}{.0013} = -3.23$$

Looking up −3.23 in the table on page 133 (under "Percentage Outside of Min."), we find the percentage is 0.1%.

To find the *total* percentage outside of specification, add the percentages outside of the upper and lower specification limits.

$$8.4\% + 0.1\% = 8.5\% \text{ total}$$

Before deciding that a distribution is "reasonably normal," compare the plotted frequency distribution with the one on page 132 by eye. See also page 133 on the use of probability paper and various other tests for normality. The shape of the distribution is more important in estimating percentages than in most other applications.

Non-normal distributions

If the distribution has a moderate skew, follow the procedure outlined above; but in Step 3 read the percentages from Table II or Table III on pages 135 and 136, depending on the direction of the skew.

Example

Suppose the following have been calculated from data (and the data are in control on an $\bar{X}$ and R chart):

$$\bar{\bar{X}} = .12225$$
$$\sigma' = .00045$$

A frequency plot of these data (Figure 57) indicates that we should use Table III (negative skew, $k = -1$ approx.).

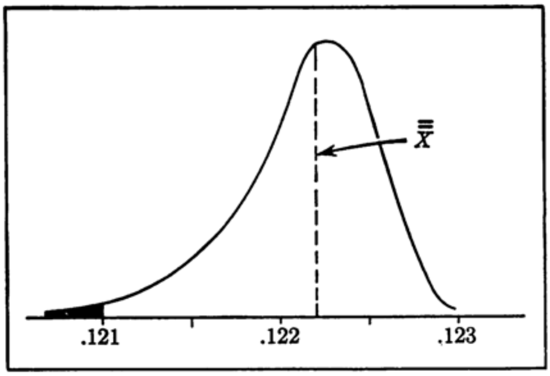

Fig. 57. Frequency distribution of some data which show control. A smooth curve has been drawn by eye around the tallied data.

Problem:

Determine the percentage outside $\bar{\bar{X}} \pm .001$.

(1) For percent outside $\bar{\bar{X}} + .001$ (.12325):

$$t = \frac{.12225 - .12325}{.00045} = -2.22$$

Find -2.22 in Table III (under "Percentage Outside of Max.") and read 0%.

(2) For percent outside $\bar{\bar{X}} - .001$ (.12125):

$$t = \frac{.12125 - .12225}{.00045} = -2.22$$

Find -2.22 in Table III (under "Percentage Outside of Min.") and read 3.6%.

Total percentage outside of $\bar{\bar{X}} \pm .001 =$ 3.6%.

If the distribution has a very pronounced skew, the best guide is an estimate based on the frequency distribution of observed results. In doubtful cases, it may be advisable to run a p-chart in the shop in conjunction with your $\bar{X}$ and R chart, in order to determine whether the percentage in the tail is controlled.

A-4.4 Estimating a percentage from a p-chart

At the conclusion of the process capability study, you have obtained a long series of points that show control. The process capability in terms of percentage (that is, the percentage outside of whatever limit was used in determining what would be called "defective") is merely the centerline on the p-chart. This estimate must always be considered as tentative unless the p-chart represents one source of data only.

If the p-chart does not represent one source of data, the sources should be studied separately before attempting to estimate the overall capability. Break the data down according to source, type of defect, or other obvious contributors to the composite pattern. Percentages should be estimated separately for the various contributors and later combined into an overall estimate.

A-4.5 Estimating capability from a pattern that is not in control

The following discussion applies to either $\bar{X}$ and R charts or p-charts.

Occasionally it is necessary to estimate the capability of a process prior to the time when it has been possible to bring the pattern into control. In such cases, the estimates can be only rough and tentative since the average of uncontrolled data cannot be taken to be the true capability.

However, estimates based on the early patterns in a capability study, even when out of control, will be better than estimates arrived at without such studies. While out-of-control patterns will not show the true capability, they permit us to make a more intelligent guess as to where it lies. In addition, they show us how far we are at present from reaching the desired state of control.

To obtain the best estimate from uncontrolled patterns, proceed as follows:

(1) If the pattern shows a trend, determine the cause of the trend and decide which portion of it represents the way in which the process will be run in the future. Estimate the capability in the manner described in paragraph A-4.1, basing your estimates on the selected portion of the pattern only. See Figure 58.

(2) If the pattern is interrupted by periodic lack of control, this can sometimes be recognized as indicating the presence of two or more separate patterns in the data. It should be possible to run the process at any one of the indicated levels provided we are able to identify the causes and bring the process, at some later time, into a state of control. See Figure 59; also Figure 207 on page 194.

Wherever it is possible to pick out such probable levels by eye, this provides a reasonable basis for estimating capability. Use engineering judgment in deciding which points are likely to indicate separate patterns.

(3) If the pattern is erratic in such a manner that it is not possible to pick out the separate patterns by eye, it may be that the best available estimate of capability will be the center of the out-of-control pattern. If this estimate is used for want of a better one, keep in mind that it is a very uncertain estimate. See Figure 60.

When estimates are based on out-of-control patterns, always explain the basis for the estimate and show the pattern of the data from which the estimate came. Also remember that

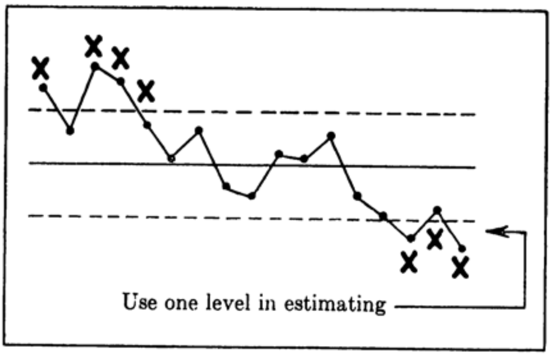

Fig. 58. Estimating from a trend.

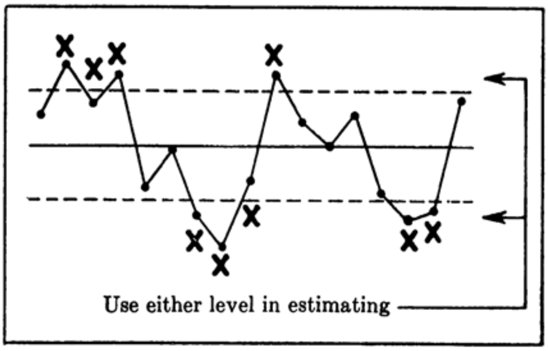

Fig. 59. Estimating from an interrupted pattern.

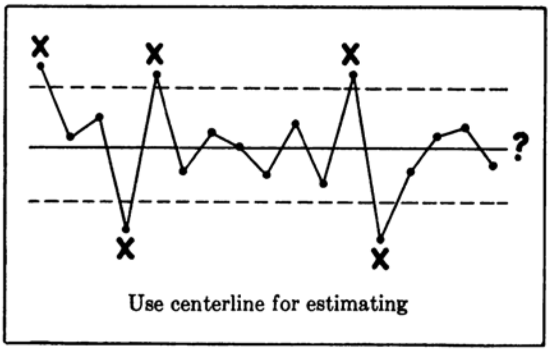

Fig. 60. Estimating from an erratic pattern.

no estimate from uncontrolled data is reliable. Reliability comes from knowledge that the data show control.

A-4.6 Wrong ways of estimating the process capability

There are two common practices in estimating capability which have no valid statistical foundation. Engineers should avoid using either of them in arriving at estimates or setting standards. The things to avoid are the following:

I. Do not attempt to use a distribution without a control chart.

A distribution will not give a reliable estimate of the process capability unless the data making up the distribution came from a controlled process. This can only be determined by plotting the data on a control chart. To use a distribution correctly in determining capability, proceed as follows:

> Plot the data on a control chart (in the order of production, if possible), and determine whether there were any significant changes during the time when the data were accumulated. If there were significant changes, the distribution should not be used to determine capability.

II. Do not attempt to use the average of past data without a control chart.

Even a long and intimate knowledge of the job cannot show whether the average of past data is anything like the real process capability.

To use the past data from a job correctly in estimating capability, proceed as follows:

> First check the data statistically by means of a control chart to determine whether it shows significant changes. If significant changes are found, the data should not be used without modification to determine capability.

These two points have important implications in engineering planning and in process control. Unless the true capability is determined, acceptance of past performance could merely authorize and allow a continuation of questionable conditions and practices.

A-4.7 Meaning of "short-term" and "long-term" capability

The "short-term" capability of a process refers to its normal behavior at any given instant of time. The "long-term" capability of a process includes the normal effect of toolwear, minor variations from batch to batch of material, and similar small and expected variations. Process control charts are set up, wherever possible, on the "long-term" capability of the process. See page 64.

The short-term capability of the process includes, among other things, the concept of its "natural tolerance." In the case of symmetrical distributions, the natural tolerance is usually taken to be a spread of ±3 sigma from the center. In the case of unsymmetrical distributions the spread in either direction may be more or less than 3 sigma, and the use of the term "natural tolerance" is likely to cause confusion. The engineer will probably find it safer to use the term "natural spread" for distributions of all kinds, and to define it always as plus so much and minus so much from the distribution center. For example, "The natural spread of the process is estimated to be ±.003." "The natural spread of the process is estimated to be +.004, −.002." When used in this manner, the term "natural spread" always refers to the short-term capability.

A-5 USING THE INFORMATION FROM A PROCESS CAPABILITY STUDY

The first step in using the information from a process capability study is to see whether the capability, as revealed by the study, is what we want it to be.

A process may be in control but at an entirely wrong level. It might be in control and still be 50% outside of specifications. It might be in control and well inside of specifications, and still be at a point that would cause the shop unnecessary trouble in adjusting or assembly. It might be in control but have such a wide spread that it would be virtually impossible to reach a high yield. In addition to this, serious measurement problems may have been encountered, or it may be necessary to get more data or different data in order to obtain a permanent solution to the problem. It is sometimes found that a specification needs to be added, reworded, narrowed, widened, modified or removed.

In any event there will be action of some sort which is required as a result of the study. The required action may be:

a. Action on the process.
b. Action on the data.
c. Action on the specification.

To determine what action is necessary, first do the following:

- Make sure that the chart is in statistical control.
- Check the level of the $\bar{X}$ chart or p-chart to see whether it is in a desirable place.
- In a case of an $\bar{X}$ and R study, check the centerline on the R chart also to make sure that the process does not have too wide a spread. See page 56 for method of calculating the "spread of individuals" in the process. It is good practice to sketch the probable spread of individuals on the $\bar{X}$ chart as shown on page 31.

Then on the basis of the above information take the following steps.

Action on the process

(1) Decide what to do about the known assignable causes. It is possible to eliminate them or leave them in the process. Some of the causes you might not wish to eliminate are:

- Normal toolwear.
- Reasonable variations in the machine settings.
- Unavoidable variations in batches of material.
- Ordinary differences between operators, etc.

If you leave these causes in the process, it will generally require a shop control chart to keep them within bounds. If you wish to keep them out of the process, it may require a control chart also.

(2) Decide whether it would pay to set up a better process having a narrower capability. It may be possible to realize savings by holding the process closer than the present specifications so as to cut down adjusting difficulties, etc. If you think it would pay to set up a narrower process, it may be necessary to run a designed experiment or make further capability studies.

(3) Find whether the present process can be relocated in such a way as to get better results or higher yields. Perhaps a change in material, the re-design of a tool or fixture, a change in the operator's work pattern can move the process up or down. If the process can be relocated to advantage, will the new location be permanent? If not, and if the location will be under the control of the shop, you will need a shop control chart.

(4) Find whether it would be possible to get cost reduction by using a more economical process. No capability study should be considered complete until this possibility has been investigated.

(5) If the process at the end of the study is in control with a satisfactory level and spread, and if none of the above actions are necessary, take steps to see that this desirable condition is made permanent. This in itself may require a shop control chart.

Note that many of the actions that may be required on a process tend to result in installing shop charts. See Paragraph A-6 for the method of setting up shop charts to complete the action in a process capability study.

Action on the data

This may be of two different kinds:

(1) Error of measurement. There may be indications that the measurements are not reproducible. In that case the engineer may need to make an error of measurement study before collecting further data.

(2) Inadequate amount of data. The patterns on the chart may still be inconclusive. If so, more data are needed to solve the original problem.

Action on the specification

This may take at least six different forms.

(1) Attempt to widen specifications which are found to be narrower than the process capability.

(2) Attempt to narrow specifications, if by so doing you can obtain economic advantages or a reduction in complaints. Investigate the value of using tighter tolerances on piece parts or components in order to reduce trouble in later assemblies.

(3) Take off specifications found to be unnecessary.

(4) Add specifications found to be necessary or desirable.

(5) Shift the nominal of specifications found to be in the wrong place.

(6) Re-word specifications found to be in need of modification or clarification.

Do not be discouraged if you find that one process capability study merely leads to another study and that one to a third study, etc. The problems you are tackling may have existed for many years. It may not be possible to resolve such a tangled situation overnight. Some process capability studies may extend for a period of many months. See the example on pages 66–71.

The determining factor is whether the study continues to reveal information which leads to a reduction in cost or improvement in quality or both.

A-6 TRANSLATING A PROCESS CAPABILITY STUDY INTO A SHOP CONTROL CHART

Having obtained certain information from a process capability study, the engineer knows what distribution or range of distributions can be maintained economically. He now wishes to set up a process control in the shop which will enable the shop to maintain the desired distribution and obtain the desired benefits. Very often the control which is needed in the shop will be a standard shop control chart.

The engineer should keep in mind that while the control chart technique is used for both capability studies and shop charts, the two applications are entirely different. Some of the differences are shown in Figure 61. Study these differences before attempting to set up a shop chart.

Preparation for setting up the chart

Properly speaking, before a shop chart is set up, the engineer should have experimented with all elements of the process. He should have studied the effect of changes in material, methods, personnel and tools. He should also have weighed the economics of each change so that the shop standards arrived at will have real economic meaning. However, it is seldom possible for an engineer to carry a study to this point prior to the installation of the shop chart.

CHARACTERISTIC	PROCESS CAPABILITY STUDY	PROCESS CONTROL CHART
PURPOSE	To obtain information.	To maintain a predetermined distribution.
SAMPLES	Relatively few.	A running series.
ANALYSIS	Very careful analysis and interpretation.	Shop watches only the more obvious changes in pattern.
ACTION	Any change may be important, either good or bad.	The shop acts only on unwanted changes.
INFORMATION	Distribution shape is studied as well as average and spread.	Attention focused mainly on average and spread (or percent defective).
CENTERLINES	Centerlines are calculated from the data to reflect the distribution of the process being studied.	Centerlines are set to represent a balance between quality and cost. They show where we want the process to run.
RELATION TO SPECIFICATION	Relation to specification is carefully checked. The study may lead to a change in either the process or the specification.	Proper relationship to specification is allowed for when the control chart is set up.

Fig. 61. Difference between a process capability study and a process control chart.

In fact, if he did so, it would postpone the immediate benefits which the shop could obtain from the chart. Consequently the usual procedure in setting up a shop chart is the following:

(1) The engineer makes a short intensive study with enough experimentation to effect the biggest improvements immediately.

(2) He then installs a chart in the shop, and the study is continued as a routine application of charting.

Continued use in the shop results in steady improvement and in time determines the ultimate capability of the process. In the meantime the engineer is making other studies.

A-6.1 Engineering a shop chart into the job on the basis of a process capability study

The following procedure is used in translating process capability information into permanent shop form.

(1) Base the shop control chart on the "long-term" capability of the process, as shown by your present studies, rather than its "short-term" capability. (See page 61.)

(2) Decide whether it is necessary to maintain a single level as nearly as possible, or whether it would be satisfactory to let the distribution shift. If it is to be permitted to shift, determine whether the amount of shifting should be limited on the high side only, on the low side only, or on both sides. Set the centerlines in accordance with this decision.

(3) Calculate control limits for the shop charts using the $\bar{R}$ or $\bar{p}$, as the case may be, from the controlled patterns in your process capability study. The mechanical details of installing the chart should be handled jointly by the Quality Control Team. See pages 187–199 and 228–229.

(4) Work with the shop regularly and at frequent intervals in using the information which develops from the chart. The Shop Section of the Handbook contains a large amount of practical wisdom on this subject which has been accumulated through years of application of statistical quality control. This will be found useful to engineers as well as Operating people.

On an $\bar{X}$ and R chart, the presence of a specification limit on one side or both is frequently a factor in setting up the chart. Other economic considerations may also exert a strong influence on the engineer's decisions, as follows:

(1) Some processes are not capable, in the present state of engineering and manufacturing knowledge, of turning out product which is compatible with specified limits.

(2) Some specifications are not compatible with each other. If the process is run at a level which ensures meeting one set of limits, large quantities of product may be outside of another set of limits.

(3) Some processes have an optimum level which will minimize, say, later difficulties in assembly. In such cases the engineer may decide to run the process well inside of the present specifications. Or he may decide to run it on the high or low side of nominal.

(4) Some processes, such as soldering or impregnating, tend to cause shifts in the characteristics of certain distributions. It may be desirable to allow for such anticipated shifts in choosing the optimum levels for the prior operations.

(5) There are some unavoidable conditions in materials, piece parts etc. for which it is necessary to compensate at a subsequent point in the process.

(6) Some failures to meet specifications are economically less undesirable than others. It may be cheaper to repair units which fail on the low side than on the high; or we may prefer to have a number of spoolheads so tight that they fail to fit relay cores, rather than run the risk of having loose spoolheads in the field.

(7) Some processes have maximum stability and predictability when run at a particular level.

(8) Some characteristics need to be controlled although they have no specified limits.

(9) Specifications themselves sometimes need to be changed as a result of the information developed from shop control charts.

A-6.2 Typical example of the installation of a shop chart

Figure 62 shows a typical shop chart derived from a process capability study. A representative portion of data from the capability study is shown at the left-hand side of the chart. This is useful guidance for the shop. It is reproduced from a master chart along with the scales, headings and control limits, so as to be a permanent part of every chart. The lines on the right-hand side of the chart represent the engineer's economic decision as to where the process should run.

The right-hand side of the chart will be used for the shop's samples. Note that the engineer has provided two centerlines on the $\bar{X}$ chart, to allow the shop the greatest possible leeway in running the process. The shop will ignore any patterns which form between the two centerlines, but will apply the usual tests for unnatural patterns to any points which fall between one of the centerlines and its control limit. Note that only one control limit is shown for each centerline.

If new control limits should be calculated later as a result of new studies, a portion of the new capability information will be shown at the left-hand side of the chart in place of the old information. This allows the shop people to see at all times what is expected of them, and to compare the current process with the engineering study on which the chart has been based.

Further information on shop charts will be found in the Shop Section on pages 187–229.

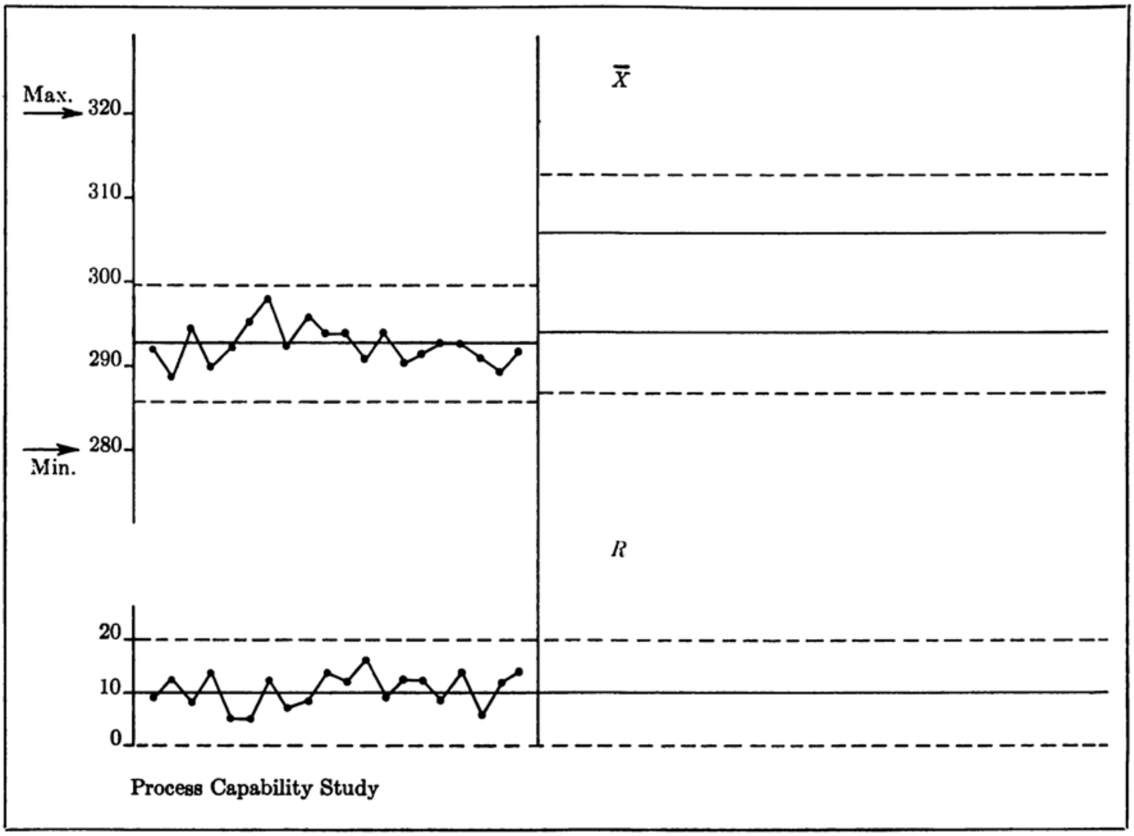

Fig. 62. Example of shop chart based on a Process Capability Study.

A-7 SIMPLE EXAMPLES OF PROCESS CAPABILITY STUDIES

The following are typical examples of process capability studies. They illustrate (a) the use of $\bar{X}$ and R charts and (b) the use of c-charts. In the first case the study is carried through to completion. In the second it is only begun. Both studies illustrate the simplification of complex patterns which was described on pages 54–56.

First process capability study: $\bar{X}$ and R charts

This problem involved an electrical characteristic on a certain type of switch. The switch was manufactured on a machine with 12 different heads. Performance was erratic, there seemed to be large differences among the heads, and a large percentage of product was being rejected because of failure to meet requirements. A process capability study was undertaken by the Quality Control Team.

The Team made the following decisions before collecting the original data:

Characteristic to be plotted: Operate value of the switch
Number of switches per sample: 5
Number of samples in the study: 50
Type of chart to be used: $\bar{X}$ and R chart
Sources to be studied: Individual source of product (Head No. 6)
Person to collect the data, and instructions to this person: Machine setter to take data and make a note of all known changes in the process during the period of study
Period of time to be covered, and amounts of data: 10–15 samples on each of several different days

Comments on these decisions

The Team properly made plans in advance for collecting suitable data. For the decision on type of chart, see pages 47–48. For the decisions on source to be studied and person to collect the data, see page 49. In addition to the instructions noted above, the machine setter should be told how many measurements to take on each switch and just how to take them.

For the period of time to be covered, see page 50. The period should be long enough to make sure that the data will fairly represent the cause system. For example, if there are monthly production cycles, the study should cover at least a month.

For the point at which the data should be collected, see pages 49 and 51.

Control chart

The control chart obtained from the initial data is shown in Figure 63. This chart has many points out of control. It is typical of the complex patterns often obtained at the beginning of a study.

Statistical analysis of Figure 63

This is a complex pattern. It is just as important for the Team to know what not to do with it as what to do with it. The following are examples of what not to do with this pattern:

(1) Do not give up, and conclude that this is not a suitable area for applying statistical quality control.

(2) Do not decide to ignore this pattern, collect more data and see if the trouble disappears.

(3) Do not waste time trying to find assignable causes for the out-of-control points in this pattern. It will be virtually impossible to find them with the pattern in its present form.

Do, however, recognize the following:

(1) To interpret a complex pattern properly, we must be able to interpret the R chart. Here the pattern on the R chart is "masked" or obscure. This means that it is inflated by the presence of hidden variables. It will be necessary to reduce or remove this inflation in order to interpret the R chart.

(2) To reduce the inflation we must eliminate at least one of the large process variables. (See page 55.) The pattern will almost always give a clue as to what this variable should be.

In the present case we see that the fluctuations on the $\bar{X}$ chart are much wider than the fluctuations on the R chart. This indicates

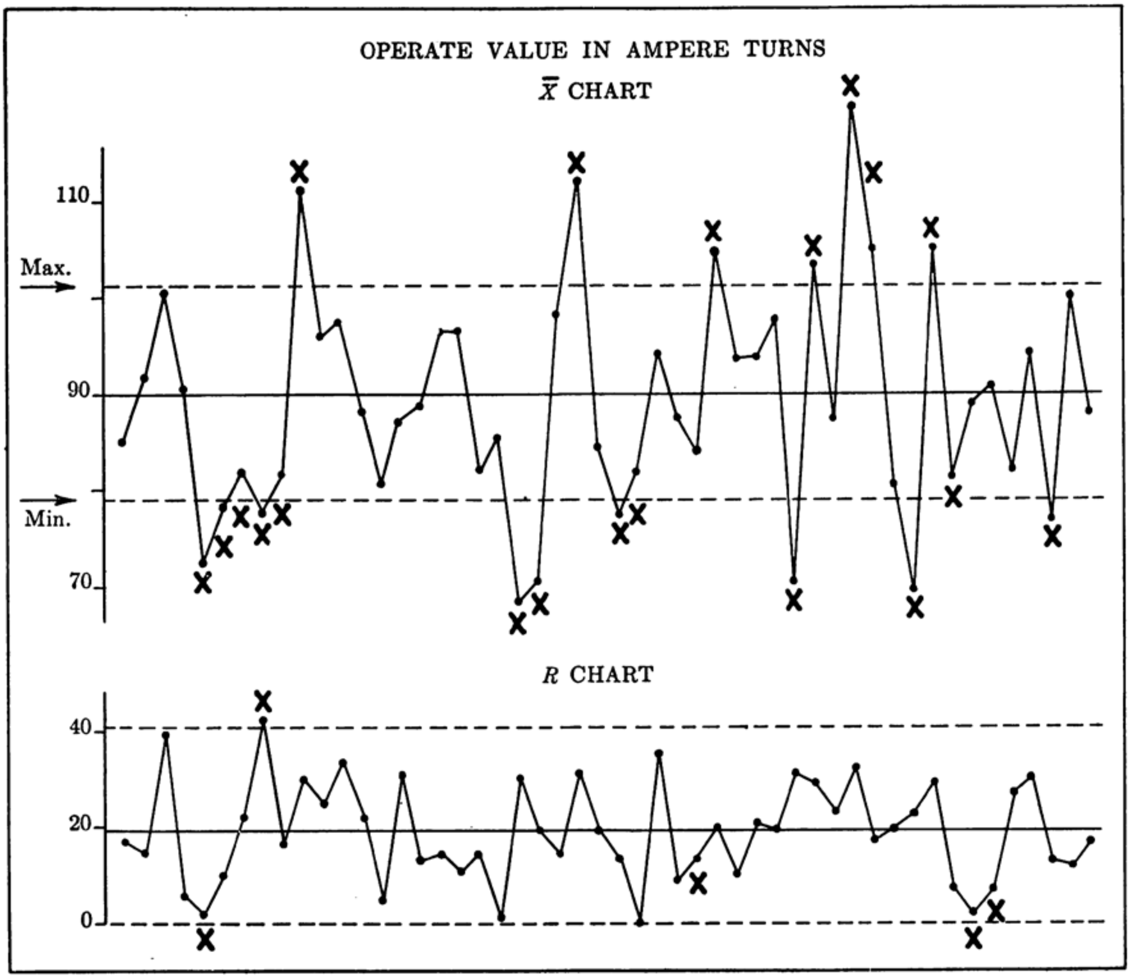

Fig. 63. First $\bar{X}$ and R chart in a process capability study.

over-adjustment. (See pages 153 and 203.) We therefore look for some obvious process adjustment and eliminate this before collecting further data.

Action taken as a result of this analysis

The Team found, on checking the machine setter's notes, that a meter had been adjusted from time to time to keep the process from "drifting." They instructed the machine setter not to make any adjustments of this meter while they were collecting the next data. The result of this action is shown in Figure 64 on the following page.

Statistical analysis of Figure 64

The chart on page 68 shows the effect of eliminating one adjustment. The $\bar{X}$ pattern is more stable, showing that the meter adjustment was one of the controlling variables. The R chart, however, shows just the opposite effect. There are 23 x's on this R chart where there were only 5 in Figure 63. To those who are not experienced in reading control charts, this pattern may look worse than the one in Figure 63.

Nevertheless this R chart contains the key to solving the problem. This can be seen by a more careful comparison of the R charts in Figures 63 and 64. The first pattern was obscure and could not be interpreted. The second is sharp and clear and can easily be interpreted. *The increased sharpness, which is reflected in the x's,* *results from the fact that we have eliminated one of the major variables.*

This is equivalent to filtering out noise in an electrical circuit. When some of the noise, or inflation, has been eliminated, this permits the signal, or hidden pattern, to come through more clearly.

Proceed as follows to interpret the R chart in Figure 64:

(1) This can be recognized at once as a pattern of "Instability." (See page 166.) Check the material in the Handbook on the subject of Instability (pages 166–167). We find that this pattern is characterized by unnaturally large, erratic fluctuations. There may be a single cause operating on the process erratically, or there may be a group of causes operating in conjunction with one another. We find that the first thing to do is check the process for *unstable mixtures.*

(2) Now check the material in the Handbook on Unstable Forms of Mixture (pages 179–180). We find that these are caused by having several distributions in the product at the same time. When the mixtures appear on the R chart, as in Figure 64, this indicates that the samples are coming at random from the various sources combined. In our own case we are not sampling deliberately from more than one source at a time. If mixtures exist in this process they are due to something that the people are not aware of.

Note the statement that unstable mixtures may show up as interactions, grouping, bunching or freaks. All of these are evident on the R chart in Figure 64.

(3) Now check the causes which are listed under "R Chart" on page 180. Among these we find the following:

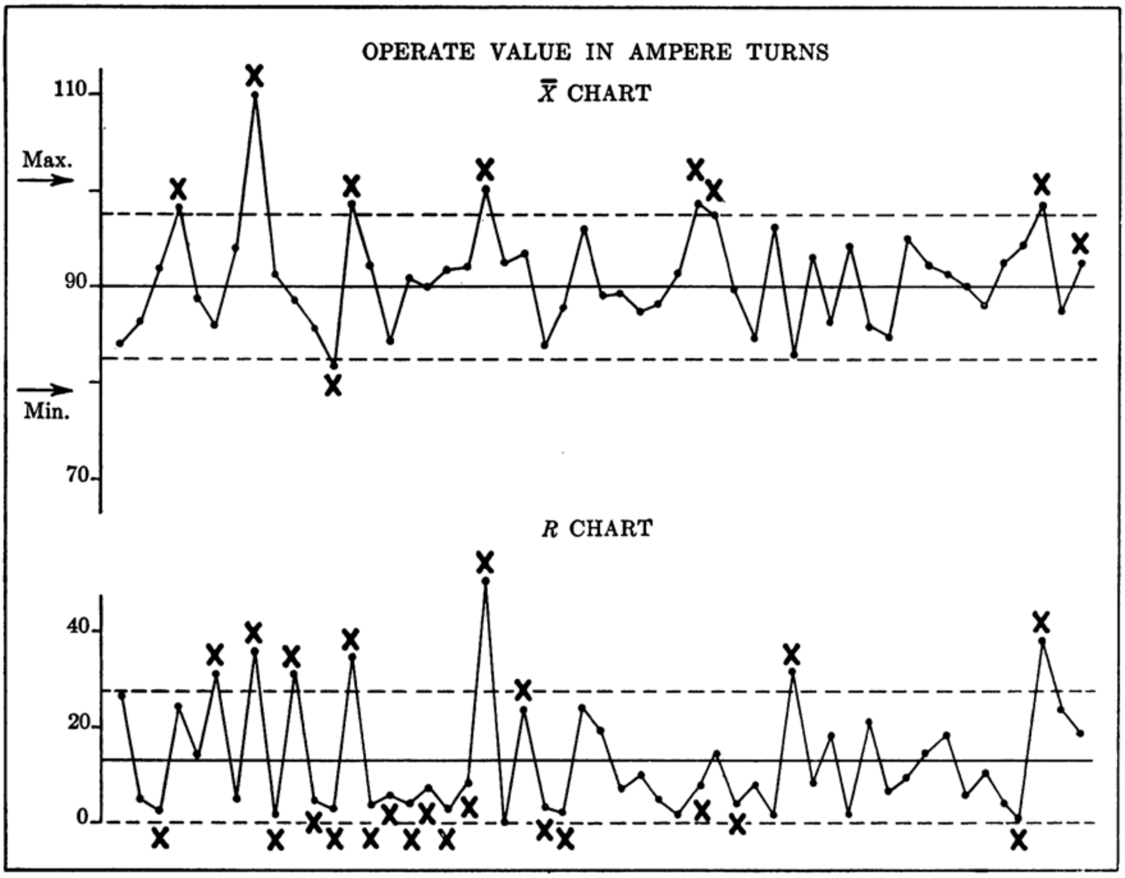

Fig. 64. Second $\bar{X}$ and R chart in a process capability study.

- Two or more materials, operators, etc.
- Too much play in a fixture.
- Holding or locking devices unreliable.
- Looseness of a chuck.
- Machine in need of repair.
- Fixtures or holders not holding the work in position.
- Lack of alignment.
- Etc.

Some of these possibilities can be eliminated at once, since they would not apply to this job. Any which cannot positively be eliminated should be carefully checked. Here we would look with particular care at any devices for holding the parts.

Action taken as a result of this analysis

The Team checked all the mechanical devices which had to do with positioning or holding the assemblies. They changed one of the fixtures and provided for magnetic alignment of certain parts. The result of this action is shown in Figure 65.

Statistical analysis of Figure 65

This chart shows the effect of the improvements in positioning and alignment. Much of the instability on the R chart has disappeared. The $\bar{X}$ chart also shows smaller fluctuations, and *these are now seen to be repeating themselves in more or less regular cycles.*

Cycles were present in the original pattern also (Figure 63) and again in the second pattern (Figure 64). But it would have been almost impossible to recognize them in the presence of larger, more erratic variables.

The causes for cycles are fairly easy to trace. (See page 162.) These were found to be associated with the *time allowed for cooling* before the assemblies were removed from a certain chuck.

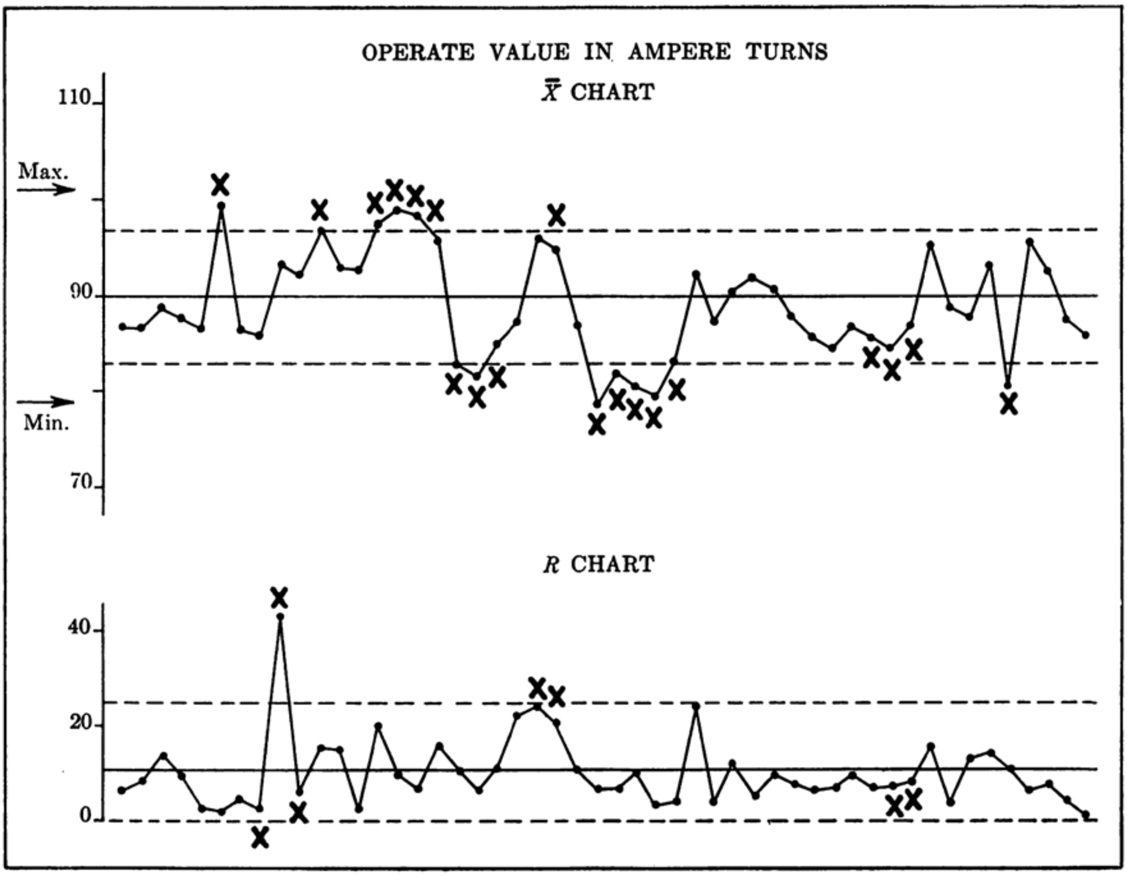

Fig. 65. Third $\bar{X}$ and R chart in a process capability study.

Many things can be read from the control chart as the patterns become simpler. In this case, for example, the cycles on the $\bar{X}$ chart tend to "follow" those on the R chart. In a very large number of instances the fluctuations follow each other point to point. This indicates skewness in the distribution of product. (See page 156.)

In a manually controlled process, one of the common causes of skewness is the operator's tendency to short-cut an operation. (See page 57.) This is consistent with the above cause of cycles.

"Freaks" on the R chart or $\bar{X}$ chart can also be checked. These were found to occur immediately before or after rest periods, lunch periods etc. This again is consistent with the cycles.

Action taken as a result of this analysis

The engineer installed an automatic timer to prevent the operator from removing the switches too soon. The result of this is shown in Figure 66.

Statistical analysis of Figure 66

This chart shows the effect of installing the automatic timer. Only an occasional point is now out of control. This indicates that most of the large assignable causes have been eliminated. On the other hand, any cause which is still outstanding will show up very plainly, since it now occurs singly rather than in combination with others. In the present example several of the high points were found to be caused by tightness or "binding."

Action taken as a result of this analysis

The engineer relocated the individual motors on each head (from a position in front to one

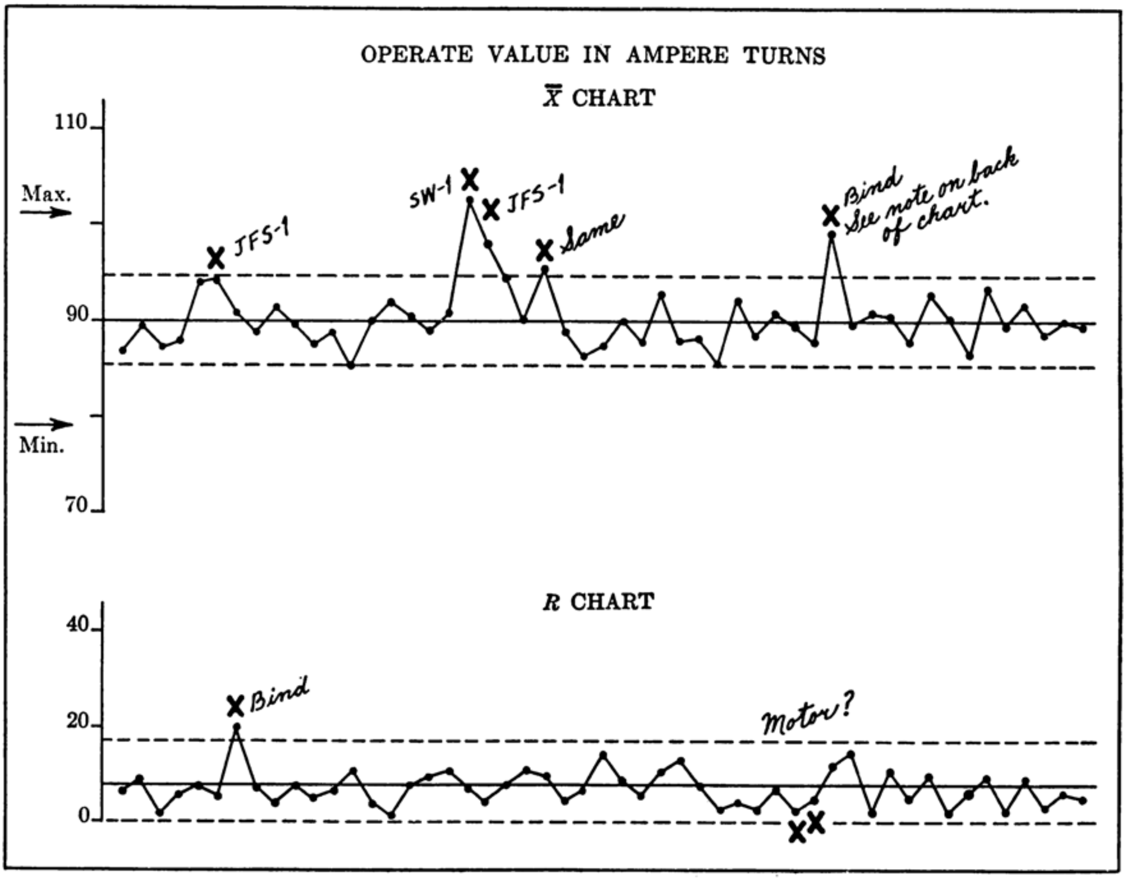

Fig. 66. Fourth $\bar{X}$ and R chart in a process capability study.

at the back of the head). This alleviated a binding condition in the sliding portions of the head. The result of this action is shown in Figure 67.

Statistical analysis of Figure 67

This is the fifth chart in the study. It shows little change from the fourth. The level on the R chart has dropped slightly, showing the effect of relocating the motors. Ordinary production variables now come and go, leaving easy traces on the chart. The shop is able to find causes when out-of-control points appear.

These things indicate to the engineer that the process is probably approaching its capability.

Capability of this process

The capability of this process is calculated as follows:

$\bar{\bar{X}} = 90$ amp. turns. Can be held consistently.

$\sigma = \bar{R}/d_2 = 6.5/2.326 = 2.8$ amp. turns

Spread of distribution $= \bar{\bar{X}} \pm 3\,\sigma = 90$ amp. turns $\pm$ 8.4.

Specification calls for 90 amp. turns $\pm$ 11.

The process is capable of meeting this specification easily and economically.

Notes on results

Original distribution (prior to process capability study) had spread of at least ± 25, and up to 40% of product might be outside of specification. Compare this with the capability calculated above.

Cost reduction was realized through (a) fewer defectives, (b) shop now making more assemblies per hour, and (c) reduced inspection.

Reliability of the product was also greatly improved.

Information on Head No. 6 was carried over

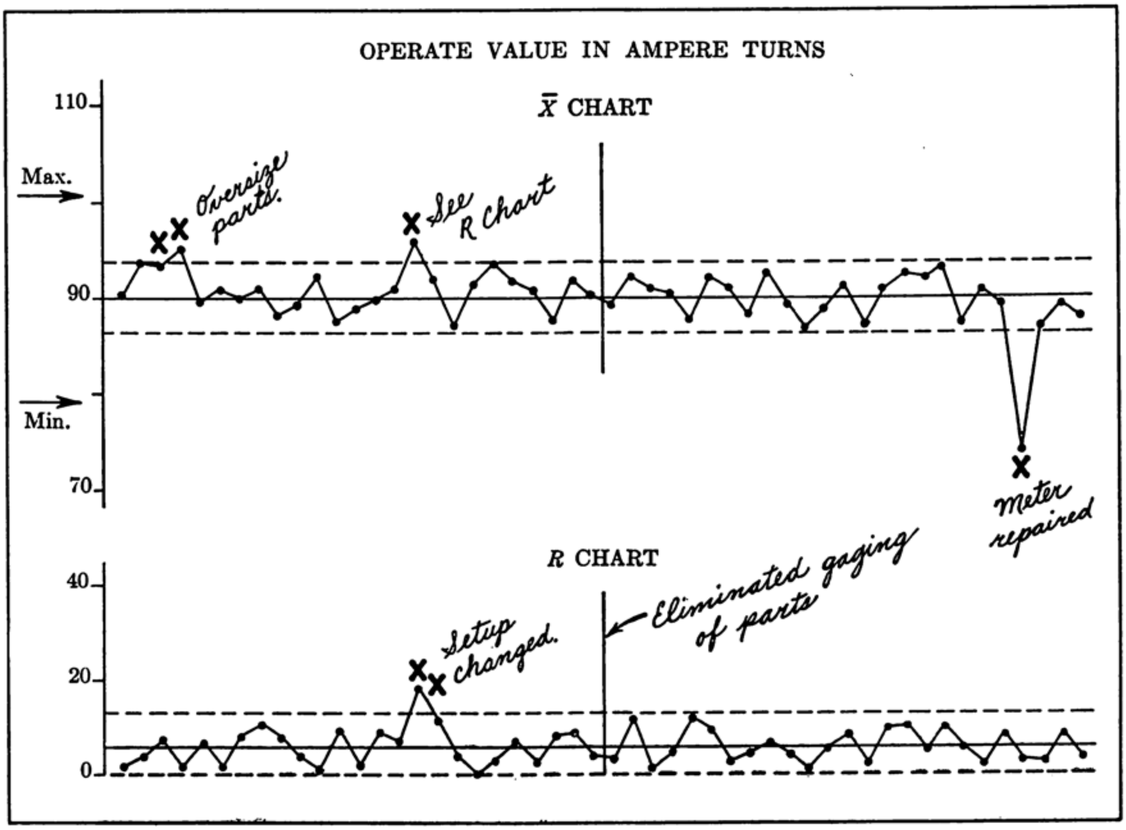

Fig. 67. Fifth $\bar{X}$ and R chart in a process capability study.

to the other heads. Similar shop charts provided for all heads.

Charts can now be used to evaluate the effect of process changes. For example, the shop wished to eliminate the 100% gaging of certain parts. When the chart showed no adverse effect (Figure 67) this change was made a permanent part of the process.

Conclusion

This was a relatively lengthy process capability study. In some cases the causes are simple enough to show up on the first chart.

A-7.2 Second process capability study: c-charts

The following is an example of the use of *c*-charts to study the work of operators. The original pattern is complex. (See Figure 68.) The control limits appear to be much too narrow for the fluctuations in the pattern, and the fluctuations are erratic. We recognize this at once as a pattern of "Instability." (See pages 166–167.)

As in all complex patterns, this must be simplified before it can be interpreted. (See page 55.) Simplification of this pattern would require the following steps:

(1) Complex patterns mean that the variable used as a basis for plotting the points in sequence is not the most significant variable. In Figure 68 the points are plotted by "operator." To simplify the pattern, select some other variable which is likely to be significant and show this on the bottom scale. For example, the defects might vary according to shift, or according to length of time on the job. Set up a scale which is arranged according to these possibilities.

(2) Re-plot the original data using the new bottom scale. Retain the same control limits that were used in the original chart. If the pattern tends to break up into simple shifts or trends, the variable used in the bottom scale is an important variable.

(3) If the pattern does not break up and become simpler, select some other variable for the bottom scale. If the pattern becomes simpler but you wish to simplify it still further, sub-divide the bottom scale according to some second significant variable.

Continue this simplification until the pattern consists of (a) simple shifts in level or (b) a simple trend. The following two charts are plotted from the data in Figure 68.

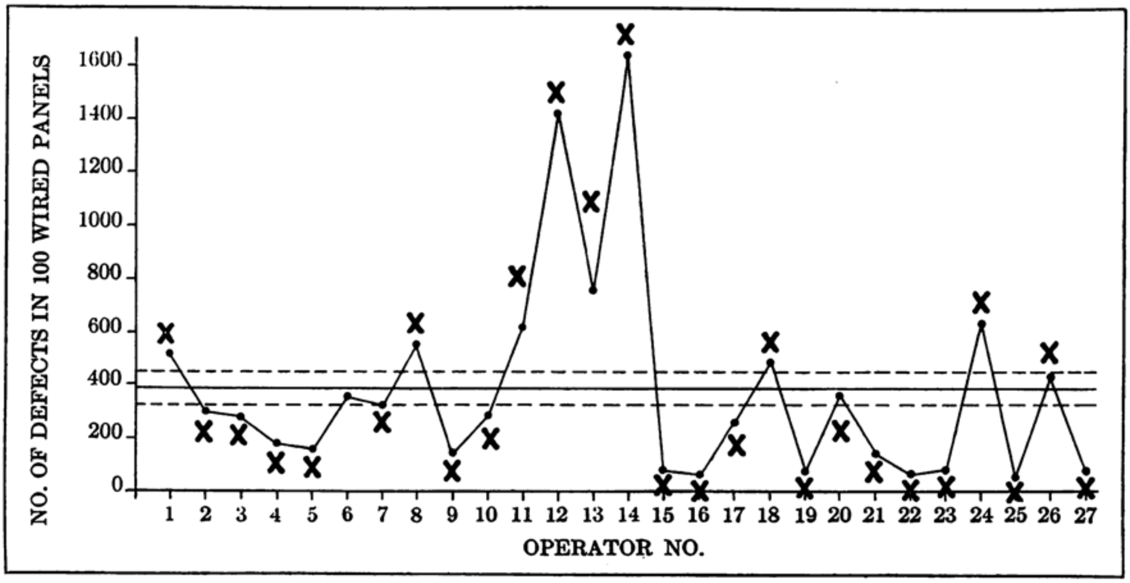

Fig. 68. Process capability study using a *c*-chart. Each point is the number of defects found in a group of 100 wired panels. Each point represents the work of a different operator.

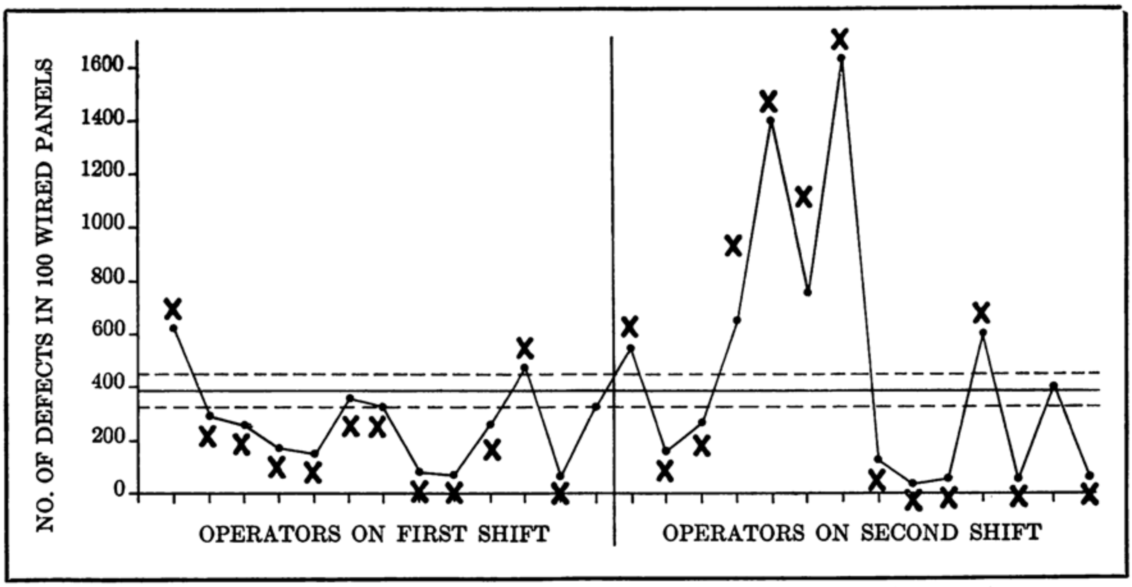

Fig. 69. First attempt to simplify complex pattern: data re-plotted according to shift.

The pattern remains complex. There is some other variable more important than shift.

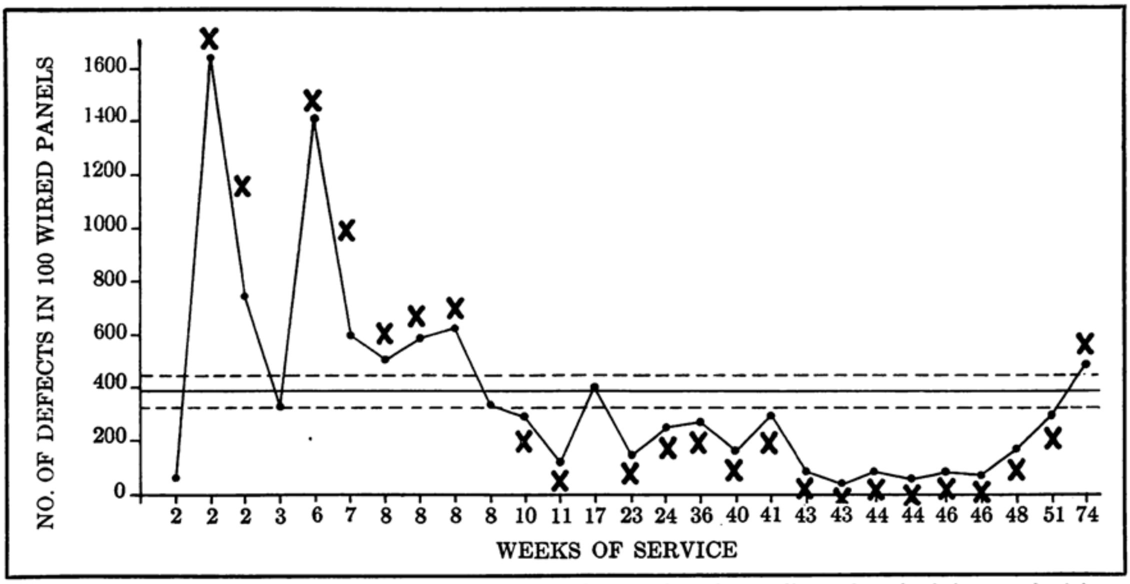

Fig. 70. Second attempt to simplify complex pattern: data re-plotted according to length of time on the job.

Now the pattern becomes simpler. Time on the job is an important clue to the cause of defects.

At this point it is possible to make tentative estimates of the capability of this process for operators with about 44 weeks of experience. The change in pattern for the operators with longest service (74 weeks) should also be investigated.

Note that, in Figure 70, the original control limits no longer appear to be unreasonably narrow. Complex patterns on a *p*-chart are simplified in much the same way as patterns on a *c*-chart.

A-8 PERFORMANCE STUDIES

"Performance studies" are temporary substitutes for process capability studies which are made by selecting data for a certain period, calculating control limits and determining whether the data show control. This permits the engineer to draw some of the conclusions he would draw from a formal capability study, and is sometimes useful until a more complete study can be carried out.

A series of performance studies over a period of time may be the practical equivalent of a process capability study, provided the necessary tracing and elimination of causes is carried out as the successive studies are made. Engineers frequently take advantage of this in setting up a program of shop control charts. The procedure is as follows:

(1) Obtain sufficient data to provide about twenty plotted points. Calculate control limits and determine whether the pattern is in control. Set up a temporary shop control chart in accordance with this information.

(2) Work with the shop to bring the chart into control or improve its pattern. When the pattern improves, select a period of data consisting of about twenty points and calculate a new set of control limits. Use this new set of limits as a second temporary shop control chart.

(3) Repeat this step from time to time in order to take advantage of progressive improvements.

Figure 71 shows how a performance study compares with a complete process capability study.

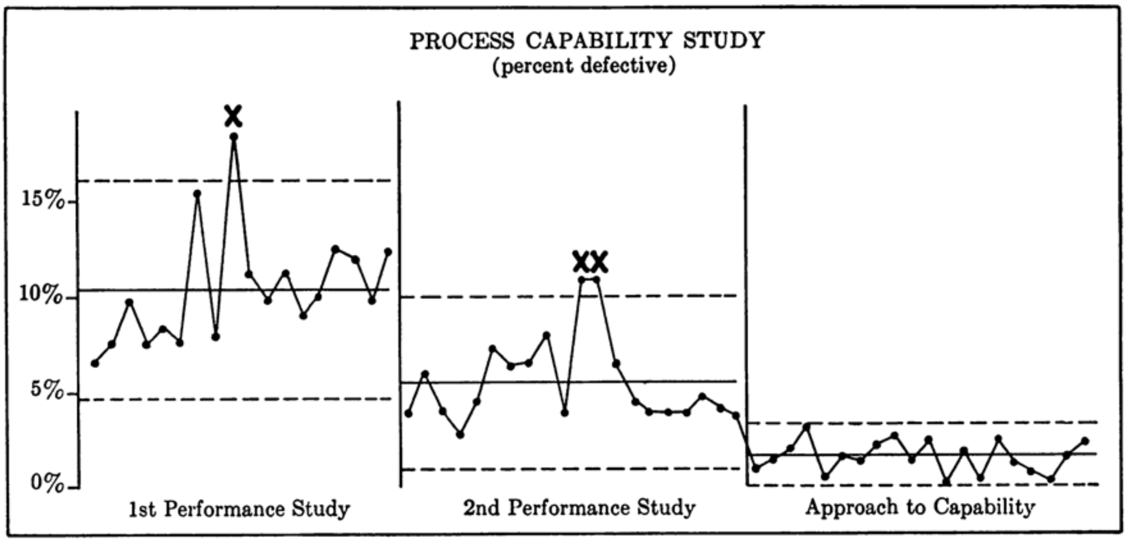

Fig. 71. A succession of performance studies over a period of time may be the equivalent of a process capability study.

PART B
Designed Experiments

This part of the Handbook covers some of the elementary principles involved in Design of Experiment. The engineer is presumed to be familiar with the making of process capability studies as covered in Part A. It is assumed that he has an acquaintance with the basic fact of variation, the fact that apparent differences in data may or may not be significant, and the need for applying statistical tests to data to discover whether a significant effect exists.

The following pages describe three types of designed experiment which engineers and shop people are called upon to handle more frequently than any others:

(1) Simple comparisons.
(2) Error of measurement.
(3) Four or five factor experiments to study main effects and interactions.

While reference will be made to statistical methods other than control charts, the principal emphasis in this section will be on the control chart analysis. Detailed information on other methods will be found in the references listed on pages 277–278.

This section will treat Design of Experiment as one of the steps in a process capability study. Particular emphasis will be placed on the tracing of causes in process capability studies which are hidden so deeply that they could not easily be discovered by other methods.

B-1 PLACE OF DESIGNED EXPERIMENTS IN A PROCESS CAPABILITY STUDY

B-1.1 Review of the theory of a Process Capability Study

The theoretical basis for a Process Capability Study was described on pages 35–36. The total variation in the process is separated by control charts into natural and unnatural portions. The unnatural portion is then studied for the purpose of identifying and eventually removing its causes. When these causes are removed the process is reduced to its true capability. It is clear that the crucial step in such a study is the *tracing and identification of causes*. If the engineer is unable to complete this step, he will be unable to get results from the capability study.

In a majority of cases, when the engineer looks at the plotted patterns in such a study, he finds there are obvious reasons for the pattern changes. All he has to do is draw on his knowledge of the job to tell what is causing the unnatural variation. In other cases he learns to identify and interpret the 15 different types of control chart pattern which are explained on pages 161–180. An example of this is given on pages 66–71.

If this is not sufficient the engineer may resort to the so-called "breakdown" techniques. He separates the data according to different sources or production paths. He uses scatter diagrams or trend arrangements to pick up correlations. A discussion of this is given on pages 54–56.

Finally, if all ordinary resources fail and the causes remain so deeply hidden that he is unable to find them, he may use a Designed Experiment to break up the variation into component parts and find the answer. The information obtained in the previous steps of analysis is a vital factor in properly designing the experiment. Omission of the previous steps often leads to failure in using the experimental techniques.

B-1.2 Comparison between Design of Experiment and Process Capability Study

A Process Capability Study can be thought of as a one-factor designed experiment. The one broad factor being studied is time (or production path, source of material, or any other factor which forms the basis for the plotting of points in sequence.) A Designed Experiment on the other hand may include a number of factors which are being studied all at once. The data are arranged and rearranged for study according to these factors. The feature which makes it a "designed experiment" is the fact that the data have been collected in such a manner as to make these re-arrangements possible.

From another point of view, Design of Experiment is a much more limited technique than a Process Capability Study. Process Capability Studies are concerned with the effect of any and all factors on the characteristic being studied. The rule that capability studies should cover a number of different points in time arises from the fact that it is necessary to allow sufficient opportunity for unanticipated and unintentional changes (or variables) to appear in the process. In a Designed Experiment we try to eliminate all factors except the few which have been selected for study.

Design of Experiment is a powerful technique, one which is very useful when employed in the appropriate circumstances and when surrounded with appropriate precautions. It is, however, a technique to be used with reserve and never under any circumstances substituted for the broader Process Capability Study.

B-1.3 Meaning of "experiment" and "experimental techniques"

Every Process Capability Study can be thought of as an experiment. The principal object of the study is to learn what a given process is able to do. By changing one or more elements in the process and observing the effect on the charts, it is possible to use the Process Capability Study as an experimental technique.

Ordinarily, however, when we speak of "experiments" on a process we mean something which is, at the same time, more formal and more limited than a capability study, and more concerned with research or with delving into the unknown. We sometimes conduct experiments before we have even set up a process in order to discover what the best process would be. By using the formal experimental techniques it is possible to study the effect of several variables simultaneously and also to study any inter-relationships or interactions between them. The techniques are useful for deliberately disturbing causes which are in balance, for breaking apart the effects of hidden variables in a going process, and for studying the possible effects of variables during development and design.

B-1.4 Types of experiment

Experiments on industrial processes run all the way from informal and unofficial changes, introduced on a more or less hit-and-miss basis, to carefully planned formal experiments which may involve months of effort on the part of a group of people and which consist of an integrated set of plans using complicated mathematical and statistical designs. If we arrange some of these experiments in the order of increasing formality we might have a list something like this:

(1) Trial and error methods: introducing a change into the process and then watching to see whether an effect shows up in the results.

(2) Running "special lots," more or less carefully identified with respect to the conditions under which the special lots were made.

(3) Pilot runs, in which certain process elements are deliberately set up with the expectation of producing a desired effect. Results are then studied to see how close they come to what was anticipated.

(4) A planned experiment involving a simple comparison of two methods.

(5) A somewhat more complicated experiment involving more than one factor. For example, an error-of-measurement study where we wish to separate the effect of the measuring instrument from the effect of variables in the product.

(6) A more complicated experiment involving several factors, set up in such a way as to make it possible to study interactions between these factors.

(7) Experiments containing more factors and arranged in still more complicated designs.

(8) A comprehensive experimental plan embracing broad problems and including many experiments. Some "operations research" projects are of this type.

In using experiments of Types 1 and 2 (trial and error methods or running special lots) the informal conclusions reached should be checked on process control charts. Pilot runs (Type 3) should be analyzed with a standard process capability study. Experiments of Types 4, 5 and 6, which involve a limited number of factors and simple interactions, will be discussed in the present section.

This Handbook will not attempt to give information on the more complicated types of experiment such as No. 7 and No. 8.

B-1.5 Dangers and pitfalls in non-statistical experimentation

The ordinary type of experiment is subject to many pitfalls and dangers. Among the most serious of these are the following:

(1) Unless the experiment is carefully planned and the results studied statistically, conclusions may be wrong or misleading.

(2) Even if the answers are not actually wrong, non-statistical experiments are often inconclusive. This may cause the experimenters to fail to recognize a proper and productive course of action. It may also send them off experimenting along the wrong lines.

(3) In non-statistical experimenting, many of the observed effects tend to be mysterious or unexplainable. A given procedure may not yield the same results a second time. The results may be in conflict with job knowledge or shop experience.

(4) Time and effort may be wasted through studying the wrong variables or obtaining too much or too little data.

By planning his experiments statistically and analyzing them with control charts, the engineer is able to avoid many experimental problems and obtain solutions for others.

(1) He can save time and money.

(2) He can carry out the experiment with less interruption to the shop.

(3) He can drop out statistically the effects of unwanted variables.

(4) He can evaluate the results when experiments fail to repeat.

(5) It is easier to reconcile his new results with previous knowledge.

(6) He can plan scientifically how much data to collect and what variables to include.

B-2 EXPERIMENT I (Comparison of Two Methods)

B-2.1 Background

The simplest experiment which an engineer is called upon to conduct, and also the most common type of experiment performed in industry, involves the comparison of two machines or two methods. The background for an experiment of this type might be as follows:

> The engineer has designed Method 1 in the hope that it will be superior to Method 2. The variable in which he is interested is a certain electrical property. He wants the measurements to be high and as uniform as possible. The engineer sets up the two methods and obtains a certain amount of data for each.
>
> By a casual comparison of the two sets of data he is unable to tell conclusively whether Method 1 is better than Method 2. He decides to test the data statistically in order to find which method is better.
>
> His real purpose is to be able to make a further decision: that is, whether to change over to Method 1 or forget about Method 1 and try something else.

The measurements obtained for one such experiment were as follows:

	Method 1		Method 2	
Time ↓	8.6	−9.3	−3.9	4.6
	12.7	3.8	11.9	−3.0
	8.2	7.0	2.9	1.5
	10.0	−4.7	−4.0	2.3
	5.8	−8.2	0.7	−2.0
	5.5	0.8	−9.4	−8.4
	12.5	−8.7	0.4	8.6
	10.1	3.5	−7.6	5.6
	8.7	1.7	−1.9	−6.9
	9.1	−9.0	14.4	4.8

Fig. 72. Data for an experiment: comparison of two methods.

B-2.2 Methods of analysis

The analytical method recommended in this Handbook will be the control chart. However, the engineer should be acquainted, at least briefly, with various other techniques which are often used to analyze experimental data. Among these are:

(1) Observation of the data.

(2) Tests for normality.

(3) F-test (variance ratio test).

(4) Bartlett's test.

(5) Tests for constancy of the system of causes.

(6) *t*-Test.

(7) The analysis of variance.

B-2.3 Observation of the data

The following comparisons may be made without applying statistical tests:

(1) Visual comparison

Do the measurements under Method 2 appear to the eye to be different from Method 1?

(2) Average of the measurements

The measurements for Method 1 average +3.405. Those for Method 2 average +0.53. Should this be considered a significant difference in average?

(3) Observed range of measurements

The measurements for Method 1 range from −9.3 to +12.7. The measurements for Method 2 range from −9.4 to +14.4. Does this indicate a significant difference between methods?

(4) Simple observed proportions

There are 5 negative measurements in Method 1 and 9 in Method 2. Is this difference significant?

(5) Distributions

The following are the frequency distributions of the measurements in Method 1 and Method 2.

Measurements	Method 1	Method 2
+11.0 to +14.9	//	//
+ 7.0 to +10.9	~~////~~ ///	/
+ 3.0 to + 6.9	////	///
− 1.0 to + 2.9	//	~~////~~
− 5.0 to − 1.1	/	~~////~~
− 9.0 to − 5.1	//	///
−13.0 to − 9.1	/	/

Fig. 73. Distribution of measurements: Method 1 and Method 2.

The distributions may or may not be significantly different. We would hesitate to say, without a statistical test, that these two groups of measurements could not have come from the same population.

B-2.4 Analysis by formal statistical methods

The formal statistical methods require certain assumptions, the most common of which are the following:

(1) Normality of the distribution.

(2) Equivalence of the variances.

(3) Constancy of the cause system.

The following are some of the methods commonly used to check these assumptions:

(1) Tests for normality

For reasonably large amounts of data it is possible to use the "chi-square test" to test the normality of the observed data. (See Reference No. 13.) Another method would be to use normal probability paper on which the cumulative percentages could be plotted. (See Reference No. 13.) The present experiment does not include enough measurements to justify the use of either of these methods. In fact there is no satisfactory test for normality which involves so small an amount of data.

If either of the above tests were used in spite of the small quantity of data, they would not indicate any significant departure from normality, either in the case of Method 1 or Method 2, or both methods combined. Ordinarily, therefore, in a formal statistical analysis, we would assume that the data could be treated as having come from a normal population.

(2) F-test

This is a test for equivalence of the variances (also known as the variance ratio test). This test is used when there are only two variances to be compared. The calculations are shown in Figure 74.

To apply this test, look up the value of F for degrees of freedom 19 and 19 in a Variance Ratio Table (see Reference No. 5). The value of 1.29 is definitely not significant. Consequently, the variances may be considered equal.

(3) Bartlett's test

This is an alternative test for equivalence of the variances. It can be used for any number of variances. The calculations are given in Figure 75.

To apply this test, look up the value of χ^2 for $(k - 1)$ degrees of freedom in a Chi-square Table (see Reference No. 5). The value of .296 is definitely not significant. Consequently, the variances can be considered equal. This agrees with the variance ratio or F-test.

(4) Testing the assumption of constancy of the cause systems

There is no convenient way of doing this except by using control charts. In a classical analysis, we ordinarily assume that the cause systems did not change provided we feel that we have kept all conditions constant while collecting the data for Method 1 and Method 2.

F-Test

	Method 1		*Method 2*	
ΣX	68.1		10.6	
$[\Sigma X]^2$	4637.61		112.36	
Divided by 20 to obtain correction factor	231.88		5.62	
ΣX^2	73.96	86.49	15.21	21.16
	161.29	14.44	141.61	9.00
	67.24	49.00	8.41	2.25
	100.00	22.09	16.0	5.29
	33.64	67.24	.49	4.00
	30.25	.64	88.36	70.56
	156.25	75.69	.16	73.96
	102.01	12.25	57.76	31.36
	75.69	2.89	3.61	47.61
	82.81	81.00	207.36	23.04
	883.14	411.73	538.97	288.23
		883.14		538.97
		1294.87		827.20
Subtract correction factor		231.88		5.62
$\Sigma(X - \bar{X})^2$		1062.99		821.58
		÷ 19		÷ 19
σ^2		55.94		43.24
F		$\frac{55.94}{43.24} = 1.29$ at *df* (19, 19)		

Fig. 74. *F*-test for equivalence of the variances.

Bartlett's Test							
Method	$\Sigma(X)$	$\Sigma(X^2)$	$(\Sigma X)^2$	$(\Sigma X)^2 \div n$	$\Sigma(X - \bar{X})^2$	σ_i^2	$\log_e(\sigma_i^2)$
1	68.1	1294.87	4637.61	231.88	1062.99	55.94	4.025
2	10.6	827.20	112.36	5.62	821.58	43.24	3.767
						99.18	7.792

Mean variance $= S^2 = 49.590$

$\log_e(S^2) = 3.904$

$k = 2$ (number of methods)

$df = 19$

$$B = k(df)\log_e S^2 - (df)\,\Sigma(\log_e \sigma_i^2)$$
$$= 2(19)(3.904) - 19(7.792) = 148.352 - 148.048 = .304$$
$$C = 1 + \frac{k+1}{3k(df)} = 1 + \frac{3}{114} = \frac{117}{114} = 1.026$$
$$\frac{B}{C} = .296$$

Note: $\frac{B}{C} = \chi^2$

Fig. 75. Bartlett's test for equivalence of the variances.

(5) Conclusions reached by the foregoing methods

We conclude from the foregoing tests that the data may be treated as having come from a normal population, that the variances of the two methods may be considered equal and that the two cause systems may be considered constant. Under these assumptions, it is possible to test for a significant difference between the averages by using either a t-test or the analysis of variance.

(6) t-Test

The calculations for one form of t-test are shown in Figure 76.

Look up the value of "t" for 38 degrees of freedom in a "t" Table (see Reference No. 5). The value of 1.32 is definitely not significant. Consequently, the averages can be considered equal.

(7) Analysis of variance

This is an alternative test which can be used for any number of averages. It depends on finding the variance of the entire set of numbers, subtracting the variance due to the observed difference between averages, and using the remaining variance (or residual) to test

"t"-Test

Method 1

$$\bar{X}_1 = \frac{\Sigma X}{n} = \frac{68.1}{20} = 3.4$$
$$\sigma_1 = \sqrt{\frac{\Sigma X^2}{n} - \bar{X}^2} = \sqrt{\frac{1294.6}{20} - 11.6}$$
$$\sigma_1 = \sqrt{53.1} = 7.3$$

Method 2

$$\bar{X}_2 = \frac{\Sigma X}{n} = \frac{10.6}{20} = .53$$
$$\sigma_2 = \sqrt{\frac{\Sigma X^2}{n} - \bar{X}^2} = \sqrt{\frac{827.4}{20} - .28}$$
$$\sigma_2 = \sqrt{41.0} = 6.4$$

Let σ_D be the standard deviation of the difference between averages.

$$\sigma_D = \sqrt{\frac{(\sigma_1)^2}{n1} + \frac{(\sigma_2)^2}{n2}} = \sqrt{\frac{(7.3)^2}{20} + \frac{(6.4)^2}{20}} = \sqrt{2.65 + 2.05} = 2.17$$
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_D} = \frac{3.40 - .53}{2.17} = \frac{2.87}{2.17} = 1.32$$

Fig. 76. t-Test for significant difference between averages.

whether the observed difference between averages is significant.

This method can be extended to test a number of variables and interactions in a much more complicated experiment. (See pages 93–97.) The calculations for the present data are shown in Figure 77.

To apply this test, look up the value of F for degrees of freedom 1 and 38 in a Variance Ratio Table (see Reference No. 5). The value of 1.67 is definitely not significant. Consequently, the averages can be considered equal. This agrees with the t-test.

Practical result of using the above methods

Since none of the above tests have shown a significant difference between methods, the engineer would have to conclude that Method 1 is no different and therefore is no better than Method 2.

$n = 40$
T (grand total) $= 78.7$

Analysis of Variance

$C = \frac{T^2}{n} = 154.842$

Source	No. of ind. (i)	No. of Totals	Totals to be Squared						Grand Total
Individuals	1	40	8.6	12.7	8.2	10.0	5.8	5.5	
			12.5	10.1	8.7	9.1	−9.3	3.8	
			7.0	−4.7	8.2	0.8	−8.7	3.5	
			1.7	−9.0	−3.9	11.9	2.9	−4.0	
			0.7	−9.4	0.4	−7.6	−1.9	14.4	
			4.6	−3.0	1.5	2.3	−2.0	−8.4	
			8.6	5.6	−6.9	4.8			78.7
Methods	20	2	68.1	10.6					78.7
Residual	—	—	—	—	—	—	—	—	—

Source	Σ Squares	÷ i	Corrected Σ Squares*	*df*	Mean Square
Individuals	2122.07	2122.070	1967.228	39	
Methods	4749.97	237.498	82.656	1	82.656
Residual	—	—	1884.572	38	49.594
			*Obtained by subtracting C		

$$F = \frac{82.656}{49.594} = 1.67 \text{ at } (1,38) \text{ degrees of freedom}$$

Fig. 77. Analysis of variance to test for a significant difference between averages.

B-2.5 Control chart analysis

The control chart has certain advantages over other methods of analysis in the treatment of experimental data.

(1) It takes account of the order in which the measurements were made. In this it differs from the chi-square test, the F-test and Bartlett's test.

(2) It does not require assumptions of normality, equivalence of the variances or constancy of the cause systems. In this it differs from the t-test and the analysis of variance. The control chart sets up *hypotheses* of normality, equivalence and constancy, but the data in the experiment are able to make us reject any of these hypotheses.

The control chart will therefore, under certain circumstances, give different results from other methods.

For a control chart analysis of the foregoing

Method 1		8.6	8.2	5.8	12.5	8.7	−9.3	7.0	−8.2	−8.7	1.7
		12.7	10.0	5.5	10.1	9.1	3.8	−4.7	0.8	3.5	−9.0
	$\bar{X}$	10.65	9.10	5.65	11.3	8.90	−2.75	1.15	−3.70	−2.60	−3.65
	R	4.1	1.8	.3	2.4	.4	13.1	11.7	9.0	12.2	10.7
Method 2		−3.9	2.9	0.7	0.4	−1.9	4.6	1.5	−2.0	8.6	−6.9
		11.9	−4.0	−9.4	−7.6	14.4	−3.0	2.3	−8.4	5.6	4.8
	$\bar{X}$	4.00	− .55	−4.35	−3.60	6.25	.80	1.90	−5.20	7.10	−1.05
	R	15.8	6.9	10.1	8.0	16.3	7.6	.8	6.4	3.0	11.7

Fig. 78. Calculations for an $\bar{X}$ and R chart. Method 1 and Method 2.

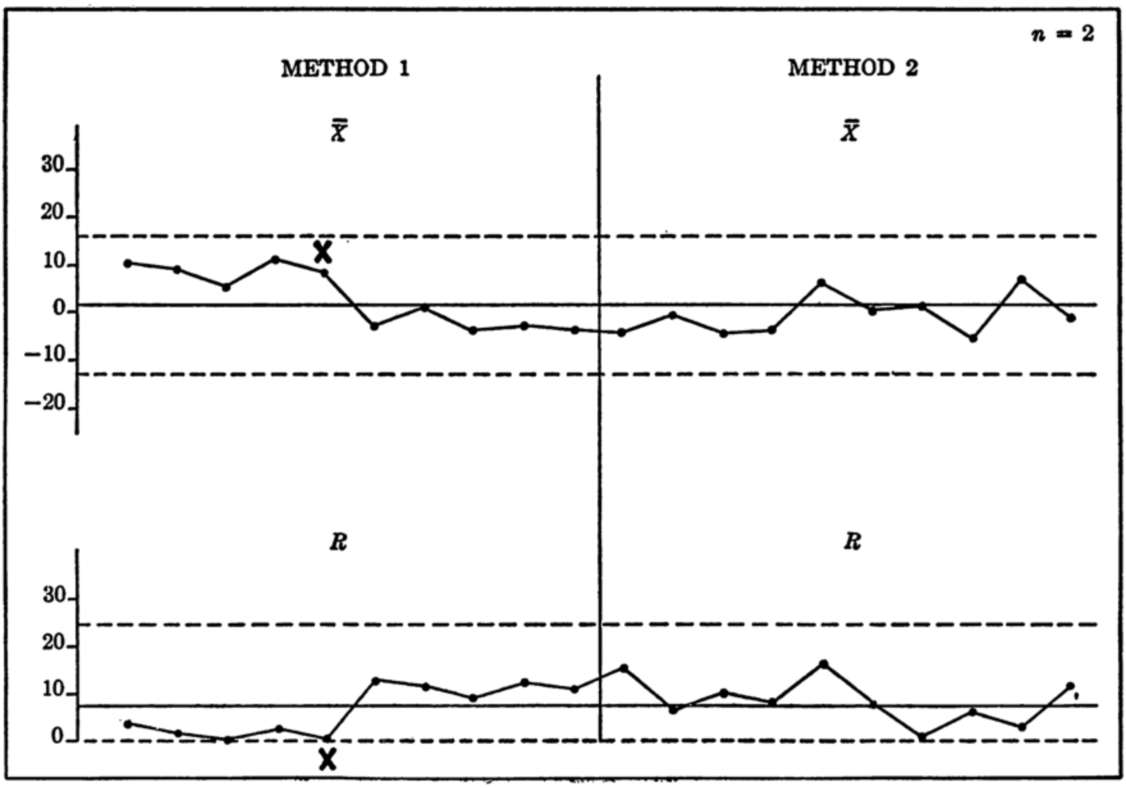

Fig. 79. Control chart plotted for Method 1 and Method 2.

data, take the measurements in the order in which they were obtained. Since the quantity of data is limited, break it up into samples of two. In calculating control limits use the ordinary control chart factors which are given on page 12.

Plot both sets of data against the same set of control limits, as shown in Figure 79. In marking x's on the R chart, use the tests for samples of two which are given on page 182.

The chart is interpreted like any other control chart.

B-2.6 Conclusions reached from studying the control chart

In Method 1, the $\bar{X}$ chart and the R chart are both out of control. We therefore arrive at the following statistical conclusions from this chart:

(1) Since Method 1 shows a definite shift in both average and spread, its data cannot be treated as having come from a normal population.

(2) Since the variability in Method 1 is out of control, its variance cannot be treated as equal to that of Method 2.

(3) The data for Method 1 did not come from a constant system of causes.

(4) Method 1 will evidently be superior to Method 2 *provided we maintain the conditions which existed during the earlier portion of the experiment.*

Note that, in this experiment, the control chart contradicts all the assumptions made in other methods.

On the practical side, it is evident that the early data in Method 1 must have come from a very desirable distribution. Its average was significantly higher than the general average of the data and its spread was extremely narrow. If the engineer is able to discover the variable which entered the process later and disturbed the results, it is clear that Method 1 will be much more uniform than Method 2 and will also give higher readings.

There is no similar indication of possible improvement in the case of Method 2.

B-2.7 Practical results from studying the control chart

The engineer investigated what had happened in Method 1. He found there was trouble with a small locating device which had jammed. This accounted for the drop in average and the immediate increase in spread. By making a modification which would virtually eliminate the possibility of jamming, the engineer was able to get consistently superior results from Method 1.

When experimental data are plotted on control charts, the results often reach farther than the experiment itself. In the above case the engineer had found a mechanical condition which was able to affect an electrical measurement. This caused him to investigate various other conditions having to do with the mechanical location and feeding of the parts. Eventually, by studying successive sets of data with control charts, he was able to get the average electrical measurement up to almost 16 and still retain the very uniform spread which had been indicated originally by his R chart.

B-2.8 Comparison between the control chart and other statistical methods

In spite of the fact that different conclusions were reached in this experiment, there is no theoretical conflict between the control chart and other sound statistical methods. If these had been reliable averages for Method 1 and Method 2 ($\bar{X}$ charts in control) and if the variability had been constant and equal (R charts in control), the control chart and the other methods would have agreed very closely.

In general, the control chart will agree with an F-test or Bartlett's test provided the patterns on the $\bar{X}$ chart and R chart would not show any unnaturalness if the individual factors were plotted separately. It will agree with a "t" test or analysis of variance if, in addition to the conditions cited above, none of the R patterns would show any evidence of unnaturalness when plotted on the same R chart. However, if a control chart pattern shows instability, freaks, stratification, etc., the control chart analysis is likely to disagree with other methods. In such cases, the information given by the control chart will cause us to modify any conclusions arrived at by other methods.

For this reason it will be desirable to use control charts in analyzing the experiments described in this Handbook.

B-3 EXPERIMENT II (Error of Measurement)

B-3.1 Background

This is a somewhat more complicated experiment than Experiment I. It involves two different factors whose effects are to be separated and studied. The background for this experiment is as follows:

> Two instruments are available for measuring a certain product. Instrument 1 is believed to be a finer piece of equipment than Instrument 2. However, the engineer and others have measured the product with both instruments and attempted to compare them. It does not appear from the initial results that either instrument is superior. We wish to test the two instruments statistically to determine which is better for measuring this product.

To solve this problem it is necessary to "design an experiment" which will separate the error of measurement from the variations in the product. This is different from Experiment I where we were concerned only with variations in the product.

B-3.2 Original data

The following example shows why a poorly designed experiment may fail to give proper conclusions. Originally, a number of units of product were measured on Instrument 1. The same units were then re-measured in the same order on Instrument 2. All measurements were made by the same experienced operator. The measurements are given in Figure 80.

The engineer took these in groups of 5, preserving the order in which the measurements had been taken. The first sample consisted of 20, 24, 19, 21 and 25. The other samples followed in order. He labeled the chart Instrument 1 and Instrument 2. See Figure 81.

On studying this chart he was disappointed to find that there was apparently no difference between the instruments. As a matter of fact, note that there is remarkable correspondence between the two halves of the chart. The pattern of Instrument 1 is repeated very closely under Instrument 2.

Interpretation of this chart

A little consideration will show that the above chart is not really comparing instruments. This chart shows certain pieces of product measured on Instrument 1 (first half of the chart) and then the same pieces of product measured on Instrument 2 (second half of the chart). The similarity between the two halves of the chart tells us that we had the same pieces of product measured in the same order in both cases; also that the effect of the measuring instrument, whatever it may have been, was not large enough to hide the variations that actually existed in the product. This is a chart on product rather than a chart on instruments. There is nothing in the "design" or planning of this experiment that makes it possible to compare instruments or to study the measurement error.

B-3.3 Design for an error of measurement study

The design of an experiment for this type of study is very simple. All that is necessary is to

Instrument 1					*Instrument 2*				
20	24	28	26	18	21	28	27	28	23
24	23	28	26	21	26	27	32	27	20
19	22	21	18	24	23	20	23	19	24
21	15	29	16	24	19	15	30	14	24
25	22	27	25	25	22	13	26	23	25
18	29	24	24	21	17	24	23	19	25
19	26	22	20	18	23	24	19	20	15
27	32	24	20	25	31	33	22	23	24
21	20	23	23	28	24	22	25	23	32
25	24	17	22	20	22	25	18	22	19

Time ↓

Fig. 80. Data for an experiment: error of measurement study.

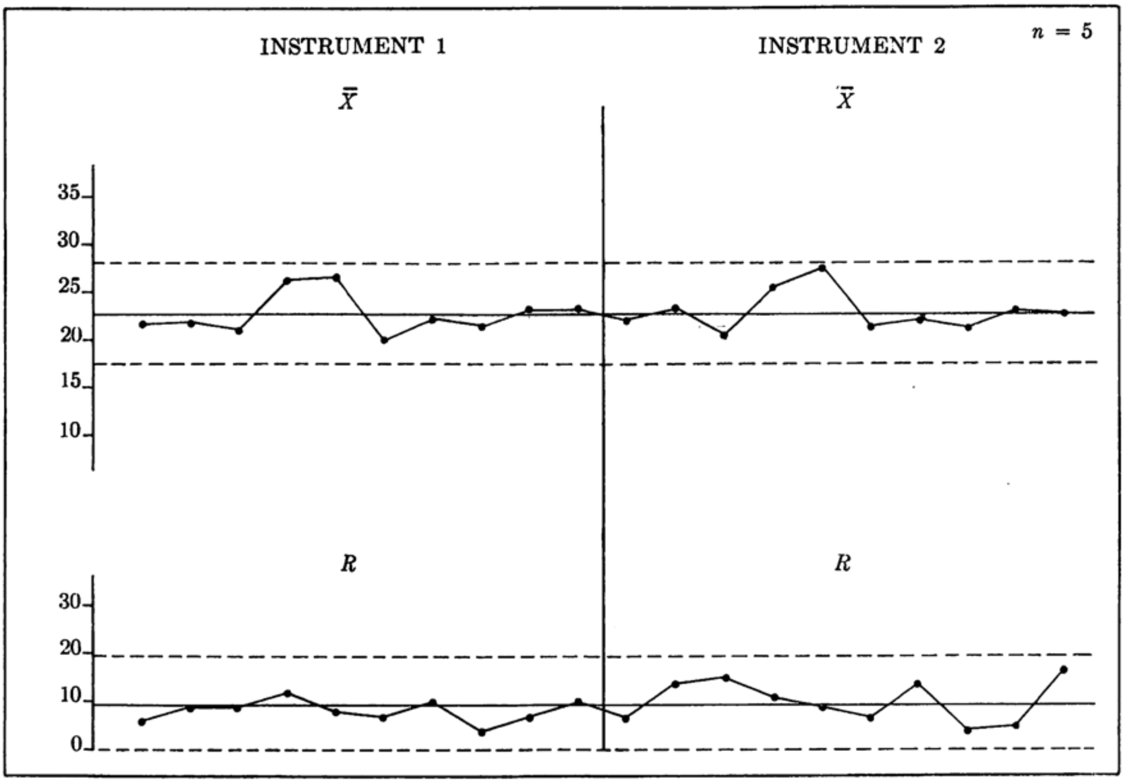

Fig. 81. Does this chart show a comparison of Instrument 1 and Instrument 2?

measure the same parts again, as follows:

> Retain the 50 measurements already recorded under Instrument 1. This time however, they should be identified as Measurement A.
>
> Then take the same parts and measure them a second time in the same order on the same instrument. Record the new set of measurements as Measurement B.
>
> This procedure is repeated for the second instrument in the same way.

The two sets of measurements must be taken under the "same essential conditions." (See pages 89–90.)

The measurements obtained for the two instruments are shown in Figure 82.

This is now a "designed experiment" because it is possible to plot the data in more than one way. If we form samples as we did in Figure 81, by grouping the data vertically, we will obtain, as before, a chart on product. However, it is now possible to form samples by taking the measurements horizontally instead of vertically, so that Measurement A and Measurement B are included in one group. We can use these horizontal groups to plot an ordinary $\bar{X}$ and R chart for samples of two. This will make it possible to compare instruments, and also to obtain some interesting information about the suitability of the instruments for measuring this type of product.

B-3.4 Chart for the error of measurement study

Figure 83 on page 86 shows the $\bar{X}$ and R values for horizontal samples of two. Plotting these values on a control chart (Instrument 1 only), we obtain Figure 84.

Note that the $\bar{X}$ chart is out of control throughout the data. This type of chart requires a special interpretation which is very different from the interpretation given to the ordinary shop chart or the ordinary process capability study.

Instrument 1				Instrument 2			
A	B	A	B	A	B	A	B
20	21	24	23	21	18	23	23
24	24	22	21	26	29	19	22
19	18	24	25	23	22	22	23
21	20	23	23	19	21	25	23
25	25	17	18	22	25	18	15
18	20	26	26	17	19	28	16
19	20	26	25	23	23	27	25
27	27	18	15	31	21	19	18
21	20	16	16	24	23	14	13
25	25	25	24	22	27	23	23
24	26	24	24	28	19	19	26
23	22	20	21	27	23	20	22
22	21	20	23	20	21	23	24
15	16	23	25	15	20	23	23
22	22	22	21	13	25	22	18
29	28	18	18	24	37	23	22
26	27	21	20	24	25	20	24
32	32	24	25	33	31	24	25
20	20	24	24	22	19	24	26
24	22	25	27	25	26	25	20
28	27	21	20	27	29	25	13
28	29	18	17	32	31	15	21
21	21	25	25	23	21	24	23
29	30	28	29	30	33	32	29
27	27	20	20	26	24	19	24

Fig. 82. Simple design for an error of measurement study: data on instruments.

Instrument 1				Instrument 2			
$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R	$\bar{X}$	R
20.5	1	23.5	1	20.5	3	23.0	0
24.0	0	21.5	1	27.5	3	20.5	3
18.5	1	24.5	1	22.5	1	22.5	1
20.5	1	23.0	0	20.0	2	24.0	2
25.0	0	17.5	1	23.5	3	16.5	3
19.0	2	26.0	0	18.0	2	22.0	12
19.5	1	25.5	1	23.0	0	26.0	2
27.0	0	16.5	3	26.0	10	18.5	1
20.5	1	16.0	0	23.5	1	13.5	1
25.0	0	24.5	1	24.5	5	23.0	0
25.0	2	24.0	0	23.5	9	22.5	7
22.5	1	20.5	1	25.0	4	21.0	2
21.5	1	21.5	3	20.5	1	23.5	1
15.5	1	24.0	2	17.5	5	23.0	0
22.0	0	21.5	1	19.0	12	20.0	4
28.5	1	18.0	0	30.5	13	22.5	1
26.5	1	20.5	1	24.5	1	22.0	4
32.0	0	24.5	1	32.0	2	24.5	1
20.0	0	24.0	0	20.5	3	25.0	2
23.0	2	26.0	2	25.5	1	22.5	5
27.5	1	20.5	1	28.0	2	19.0	12
28.5	1	17.5	1	31.5	1	18.0	6
21.0	0	25.0	0	22.0	2	23.5	1
29.5	1	28.5	1	31.5	3	30.5	3
27.0	0	20.0	0	25.0	2	21.5	5

Fig. 83. $\bar{X}$ and R calculations: error of measurement study.

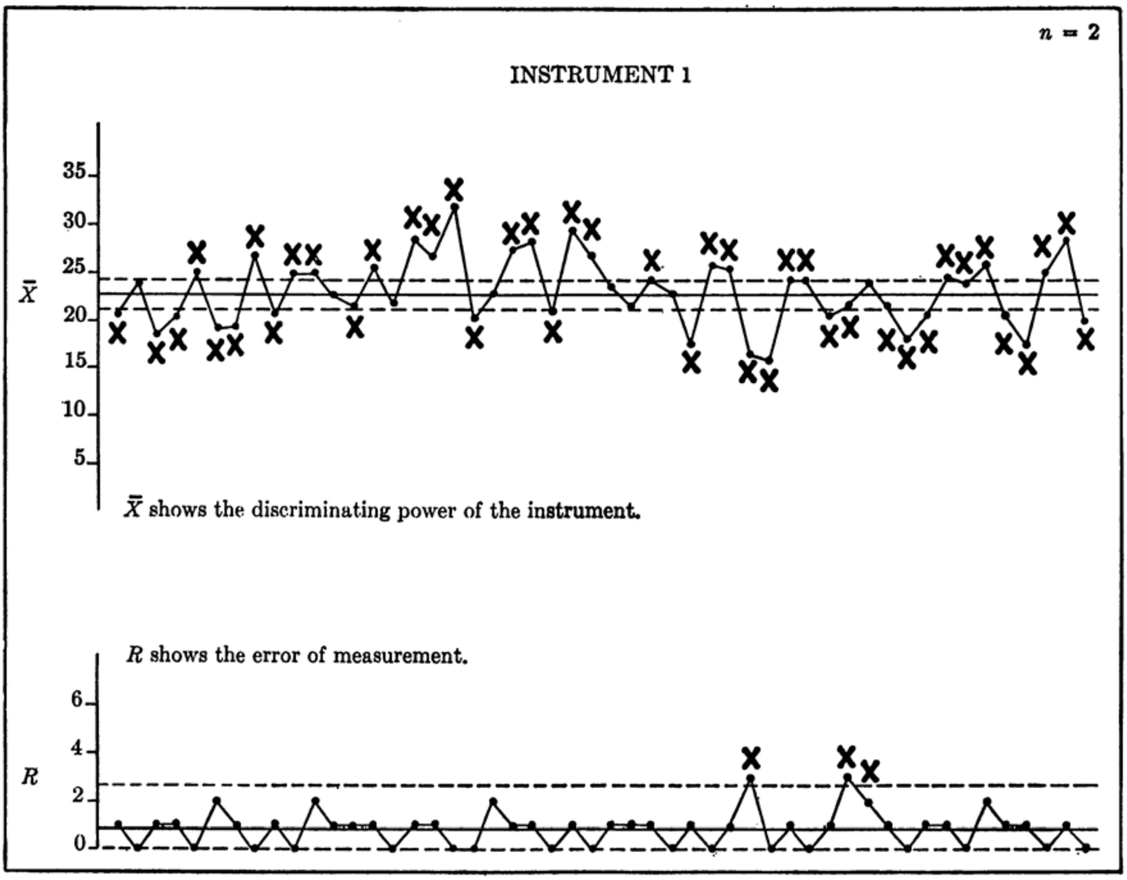

Fig. 84. Error of measurement chart for Instrument 1.

B-3.5 Interpretation of the error of measurement chart

The R chart shows, directly, the magnitude of the *error of measurement*. This is because the R values represent differences between successive measurements made by the same instrument on the same piece of product. When the R chart shows instability, as this one does in the latter portion, it means that the operator of the instrument is having difficulty in taking the measurements consistently. This in itself may be one of the factors which the engineer will wish to consider in deciding which instrument to use for a given purpose. In general, a good instrument should have a low centerline on the R chart and the indications of instability should be as few as possible.

The $\bar{X}$ chart, on the other hand, shows the *discriminating power* of the instrument. The plotted points represent different pieces of product. The control limits, being derived from the error of measurement chart, represent the inability of the instrument to tell one piece from another—that is, the area over which this instrument is not capable of discriminating. If the $\bar{X}$ points stayed in control, it would mean that the measuring instrument could see no difference between the pieces of product.

A good measuring instrument, suitable for measuring this product, should have such narrow control limits that all or most of the $\bar{X}$ points will be thrown out of control. Note that this is quite different from the ordinary control chart. *In an error of measurement study we want the $\bar{X}$ points to go out of control.*

The chart shown above should be interpreted as follows:

> Instrument 1 is capable of reproducing its results very closely and can readily distinguish between units of product. It is

suitable to use for the purpose being considered. The instrument is also capable of being used with consistent results. Its present good performance can even be improved somewhat as the operator using the instrument learns how to improve his technique and eliminates the indications of instability on the R chart.

B-3.6 Calculation of measurement error

To calculate the actual magnitude of the measurement error proceed as follows:

$$\bar{R} \text{ from Fig. 84} = \text{approx. } 0.8 \text{ unit}$$

$$\sigma' = \frac{\bar{R}}{d_2} = \frac{0.8}{1.128} = \text{approx. } 0.7 \text{ unit}$$

where d_2 is the standard control chart factor for samples of 2. See page 131.

The distribution of measurement errors is known to be approximately normal. Consequently, the spread of these errors (assuming that the R chart can be brought into control) will be approximately $\pm 3\ \sigma'$. Individual measurements can be expected to vary as much as ± 2.1 units (on repeated measurements) in extreme cases.

About two-thirds of the measurements will vary less than ± 0.7 unit ($\pm 1\ \sigma'$).

B-3.7 Comparing instruments

Taking the $\bar{X}$ and R values in the same way for Instrument 2, we obtain Figure 85. Again the R chart shows, directly, the error of measurement. We see at a glance that this is much higher than in the case of Instrument 1. Not only is it at least four times as large, but there are many more indications of instability. The operator using this instrument will have

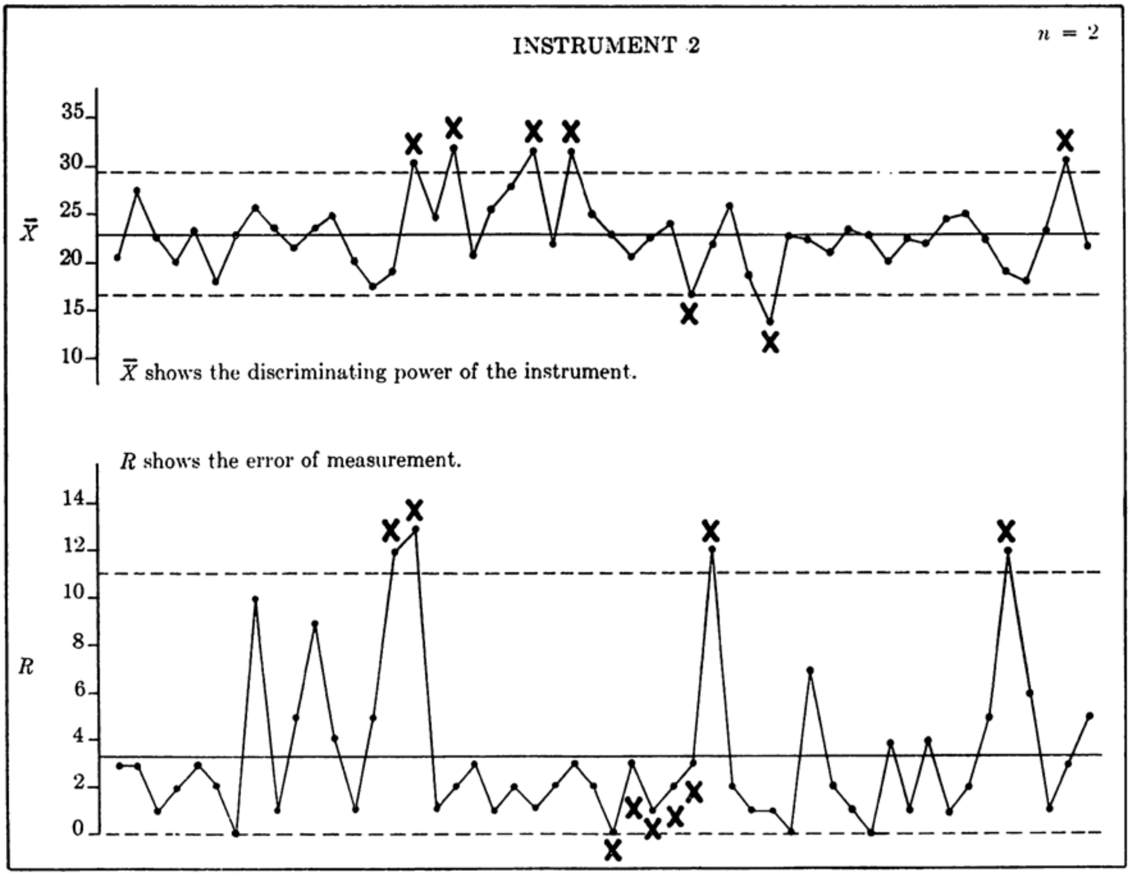

Fig. 85. Error of measurement chart for Instrument 2.

much more trouble reproducing his results (unless, of course, the operator can be trained to use the instrument much more consistently than he is doing now).

Again the $\bar{X}$ chart shows the discriminating power of the instrument. This time, however, because of the high value of $\bar{R}$, the control limits on the $\bar{X}$ chart are very wide. This means that the inaccuracies of measurement are large enough to swallow up most of the variations from unit to unit in the product. Only the very largest units or the very smallest can be reliably distinguished from the others.

It is not possible to calculate the actual magnitude of the measurement error of Instrument 2, because the R chart is out of control. A rough estimate, however, would be that σ' must be close to 3. ($\bar{R}/d_2 = 3.4/1.128 =$ approx. 3.) Since individual errors are likely to spread about $\pm 3\sigma'$, the measurements made by this instrument may vary in extreme cases as much as ± 9 units. Two-thirds or more of the measurements can be expected to vary up to ± 3 units.

B-3.8 Comparing the error of measurement with the variability of the product

Consider the case of Instrument 1, where the error of measurement was essentially in control. From the first half of the chart on product, which was shown on page 85, we observe that $\bar{R}$ for the product as measured was 8.2 units. The total variability, therefore (including both product variability and measurement variability), can be expressed in terms of σ' as follows:

$$\sigma'_{\text{total}} = \frac{\bar{R}}{d_2} = \frac{8.2}{2.326} = \text{approx. } 3.5 \text{ units}$$

Note that, since the chart in question was based on samples of 5, we must use the d_2 factor for samples of 5 rather than for samples of 2.

We know that the measurement variability of Instrument 1 (in terms of σ') is about 0.7 unit. See paragraph B-3.6. By the law of the addition of standard deviations or variances, as given on page 123, we have:

$$(\text{Total Variability})^2 = (\text{Product Variability})^2 + (\text{Measurement Variability})^2$$

This can be solved to find the product variability, x.

$$\begin{aligned} (3.5)^2 &= (x)^2 + (.7)^2 \\ 12.2 &= (x)^2 + .5 \\ (x)^2 &= 11.7 \\ x &= 3.4 \text{ units} \end{aligned}$$

Therefore, the standard deviation of the product is approximately 3.4 where the standard deviation of Instrument 1 is approximately 0.7. The standard deviation of the instrument is about $^1/_5$ as great as the standard deviation of the product.

The measurement error can also be expressed as "percentage of the total variance" as follows:

$$\frac{(\text{Measurement variability})^2}{(\text{Total variability})^2} = \frac{(.7)^2}{(3.5)^2} =$$

0.040, or 4% of the total variance.

This is sometimes spoken of as finding the "components of variance."

In all cases involving error of measurement, the observed distribution of product is the statistical sum of the real distribution of product, whatever that may be, and the distribution of measurement error. If the measurements are precise, the distribution of measurement errors will be narrow. If the measurements are accurate, the center of the distribution of measurement errors will be zero.

The above discussion covers the effect of measurement precision on the observed *variability*. It does not cover the effect of measurement accuracy on the observed distribution *center*. Measurement accuracy is discussed on pages 90–91.

B-3.9 Meaning of "measurement error." Positional variability, drift, etc. vs. the error of actual measurement

Before leaving Experiment II the engineer should note that the "measurement error" referred to above is really a combination of the instrument error itself and the error of the operator using the instrument. It would be possible to carry the experiment one step further by designing it in such a way as to separate the operator error from the instrument error. The experiment would then become a "three-factor" experiment—product, instrument and operator.

In the same way, there may be more than one source of variability in the product itself. For example, if the parts are tapered, we will obtain a different measurement from place to place on the same part, and this will be in addition to the normal variability from one part to another. If the parts are out of round, there will be a similar difference from place to place, depending on the position or point where the measurement is taken. Variability from position to position on the same piece is called "positional variability." We can eliminate it from the experiment by taking the repeat measurements in exactly the original place.

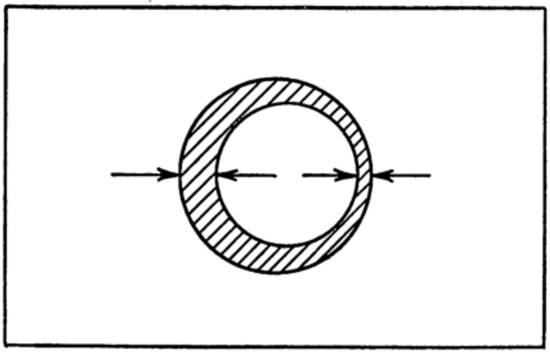

Fig. 86. Positional variability.

If we wish to study the amount of taper, out-of-round etc., as well as the error of measurement, we should take repeat measurements at *both* ends or *both* sides, and compare the two positions just as we would compare two methods or machines.

If the product is one which can drift or change from one measurement to the next (as in certain electrical properties) or if the mere act of measurement may change it (for example, by causing distortion), it will not be possible to obtain repeat measurements under the "same essential conditions." In that case it is necessary to consider the drift or distortion as if it were an additional variable—that is, we set up a three, four or five factor experiment, as explained on pages 91–101, instead of treating it as a simple Error of Measurement study.

Positional variability, or variability due to drift, distortion etc., may increase either the apparent variability of the product or the apparent error of measurement.

B-3.10 Amount of data required for the experiment

In the preceding experiment the engineer measured 50 parts. It is a general rule that the more measurements are available the more information we will be able to get from the experiment. On the other hand, the major conclusions from the experiment are likely to show up with only a moderate amount of data.

It is also a rule that the amount of data required for the experiment is governed by the amount needed to obtain the necessary precision in the estimate of variability.

If a good estimate of variability is already available before the experiment is started (for example, from process capability studies), or if there is reason to believe that the variability is not different in different parts of the experiment, then only a small amount of data will be needed. Where the variability is not known in advance, and especially where there is a possibility that there may be more variability in certain portions of the experiment, considerably more data will be needed.

B-3.11 Measurement accuracy

Experiment II was concerned with precision of measurement rather than with measurement accuracy. Precision refers to the reproducibility of the measurements: that is, the ability of the measurer or the measuring instrument to repeat or duplicate readings. Accuracy on the other hand refers to the absolute correctness of the measurements as compared with some known standard. The accuracy of measurement can be checked in either of two ways.

(1) Obtain a standard whose true value is known or has been fixed by authority: for example, a standard Jo-block, a line of standard length or width etched in glass, an oscillator of known frequency, etc. Make a series of measurements on this standard using the measurement technique which is to be studied. Plot an $\bar{X}$ and R chart. As in Experiment II the R chart will show the error or precision of measurement. However, since all measurements have been obtained on the same standard, the $\bar{X}$ chart should stay in control. The accuracy of measurement is checked by comparing the centerline on the

$\bar{X}$ chart with the known true value of the standard.

(2) If no standard of the above type is available it is necessary to have a method of measurement which is itself considered to be a standard. For example, a specific test set, master gage or micrometer may constitute the standard method of measurement. Check one or more pieces of product repeatedly using the standard method and plot an $\bar{X}$ and R chart. Check the same unit or units of product repeatedly using the method which is to be studied, and plot a similar $\bar{X}$ and R chart. The accuracy of the measurements is determined by comparing the centerlines on the two $\bar{X}$ charts.

The $\bar{X}$ chart for any one piece of product should stay in control.

In either of the above two methods, if lack of control is indicated on the R chart or $\bar{X}$ chart, check for assignable causes and eliminate them before attempting to determine the measurement accuracy. In particular, observe the basic rule which applies to all $\bar{X}$ and R charts: Do not attempt to draw conclusions from an $\bar{X}$ chart when the R chart is out of control.

B-3.12 Other statistical methods

Error of measurement data can be analyzed by other statistical methods, including the "sum of squares" method of analysis of variance. The control chart, however, has the usual advantages of

(a) Simplicity.
(b) The R chart.
(c) A plotted pattern.

B-4 EXPERIMENT III (Four Factor Experiment)

B-4.1 Background

This is a more complicated experiment than either Experiment I or Experiment II. It involves four different factors whose effects are to be separated and studied. The background for this experiment is as follows:

A certain shop is manufacturing a product which has a relatively low yield. There already are charts in the shop, and a number of process capability studies have been made. These studies indicate that one reason for low yields is the fact that a certain parameter tends to run at too low an average. If a suitable way to raise the average could be found, this would increase the yield.

Raising the average, however, is not a simple problem. Many things have been tried by the shop and engineers with no consistently good effects. There are many conflicting opinions on what should be done to produce the desired results. Among these conflicting ideas are the following:

The design engineers reported good results some time ago by using a special cleaning procedure. The Western Electric engineers have not been able to get similarly good results. On the other hand, since the trouble seems to be worse recently, they feel that it may be related to the thickness of plating on certain parts. (The plating was made thinner a short time ago in connection with certain other design changes.)

It is also possible that the trouble may be due to Western's method of activation, since this is different from the method used during laboratory development. The shop has added further confusion to the picture by claiming that it is possible to improve the yields by lengthening the drying period at a certain stage in the process. The engineers see no reason to believe that drying should be a factor, and at least one engineering experiment has indicated that longer drying may make matters worse.

In short, the results so far have been inconclusive and in some cases contradictory. A number of factors seem to be involved here, which may or may not be important individually and which may or may not be interrelated with each other.

This situation calls for a special kind of Designed Experiment, which will make it possible to study several variables simultaneously.

B-4.2 Old style experiment (without statistical design)

Originally, the engineers on this job made no attempt to design the experiment statistically. They tried to find answers to these problems by running "special lots" as follows:

(1) They held all conditions constant except the method of cleaning. They processed a certain number of units with special cleaning and a similar number with ordinary cleaning. They then measured the parameter in question and compared the results.

They found that results were slightly in favor of special cleaning.

(2) Having finished the first experiment, they ran a second "special lot." They held all conditions constant except the method of activation. They processed a certain number of units with the laboratory method and a similar number with the Western Electric method. They then measured the parameter in question and compared the findings.

They found that the Western Electric method gave slightly better results.

(3) In the same way, they held all conditions constant except the plating thickness and ran an experiment to determine the effect of this factor.

They found that thicker plating appeared to give better results.

(4) Finally, they held all conditions constant except the drying period and found that there seemed to be very little difference between the short and long drying.

(5) On the basis of the above results, they decided to use in the process:

(a) Special cleaning.
(b) The Western Electric method of activation.
(c) Thick plating.
(d) Short drying.

To arrive at these conclusions, the engineers had to process a large number of pieces of product in each of four separate experiments. This involved the wasting of many experimental pieces as well as the expenditure of large amounts of effort and time. Furthermore, when they introduced the above combination of variables into the process, they did not obtain the anticipated good results. The process continued to run with a low yield and it became obvious that the experiment had not produced the correct answers.

The engineers then designed an experiment statistically with the results shown below.

B-4.3 Designed experiment

By proper planning or design of this experiment it was possible to study all four of the above variables in a single experiment involving only 16 pieces of product. First the variables were labeled A1, A2, B1, B2 and so forth as shown in Figure 87. This is a "balanced block design" in the form of a 4 x 4 square. The design is arranged in such a way that half of the squares are reserved for condition A1 and the other half for A2 (in this example, the left and right halves of the design respectively). At the same time, using some other method of division, half of the squares are reserved for B1 and half for B2 (in this example, the first and

MEASUREMENTS ON A CERTAIN PARAMETER OF A CERTAIN PRODUCT

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	3	4	0	6
	D2	0	5	−1	4
C2	*D1*	3	−4	4	−1
	D2	5	0	3	2

A1—Original plating
A2—Thin plating
B1—Laboratory activation
B2—Western Electric activation
C1—Regular cleaning
C2—Special cleaning
D1—Short drying period
D2—Long drying period

Since we wish to raise the value of the measured parameter, we will consider effects good if they tend to result in higher values.

Fig. 87. Simple example of a Designed Experiment.

third columns for B1, the second and fourth for B2). By dividing the space horizontally instead of vertically, it is possible to reserve half of the squares for C1 and C2 respectively, and similarly for D1 and D2. This results in a completely balanced statistical design with eight squares reserved for each of the eight conditions to be studied. This completely balanced design is called a "factorial design."

With this design each variable can be studied separately if we wish, just as if the other variables were not present. At the same time, it is possible to study the variables in combination. For example, the special cleaning may have a certain effect when the plating is thin, and it may have a totally different effect when combined with thicker plating. Effects like this, which are due to a combination of variables, are called "interactions."

In the above design, it is possible to study the following variables and combinations of variables:

A alone
B alone
C alone
D alone

A combined with B
A combined with C
A combined with D
B combined with C
B combined with D
C combined with D

A combined with various conditions of both B and C
A combined with various conditions of both B and D
A combined with various conditions of both C and D
B combined with various conditions of both C and D, etc.

There will also be certain effects not attributable to any combination of variables. These additional (unidentified) effects are known as "residual."

In running the experiment, the engineers processed one unit under the combination of conditions A1B1C1D1—that is, using the original plating thickness, the laboratory method of activation, the regular cleaning method and a short period of drying. They measured the parameter in question on this unit and recorded the result in the first box or square. In the same way they processed another unit under A1B2C1D1—that is, with original plating, the Western method of activation, regular cleaning and a short period of drying—and recorded this in the second box. All the other boxes in the experiment were filled in similarly, using a random order for filling the boxes as explained on pages 114–115.

The above example shows the minimum amount of data which can be used to study four variables at two levels each. If possible, more than one unit should be processed for each box.

B-4.4 Method of analysis

A designed experiment of this type is analyzed by the method known as "Analysis of Variance." This method involves studying all the variability (or "variance") found in the data and partitioning it off into a number of separate parts in such a way that it is possible to distinguish the variability associated with each variable or combination of variables in the experiment.

There are two methods of performing the "Analysis of Variance."

(1) *Sum of squares method.* This involves rather complicated calculations, together with the application of various statistical tables. There are many articles on this in the statistical literature.

(2) *Control chart method.* This accomplishes the same results as the sum of squares method, but does it by means of addition and subtraction plus the plotting of one or more control charts.

The following is a comparison between these two methods.

SUM OF SQUARES METHOD

The method of calculating the Sums of Squares is shown in Figure 88.

The following notes relate to the numbered columns.

(1) In this column list (a) the individuals, (b) the 4 main effects, (c) the 6 first order interactions, (d) the 4 second order interactions and (e) the residual. Each of these is a possible source of variation.

Data for the Experiment Shown in Figure 87

n (the number of observations in the experiment) $= 16$
T (the grand total of all observations) $= 33$
C (the correction factor) $= \dfrac{T^2}{n} = \dfrac{(33)^2}{16} = 68.0625$

(1)		(2)	(3)	(4)								(5)	(6)
Source of Variation		No. of Ind. (i)	No. of Totals	Numbers to Be Squared								Grand Total	Sum of Squares
Individuals		1	16	3	4	0	6	0	5	−1	4		
				3	−4	4	−1	5	0	3	2	33	183.0000
Main Effects	A	8	2	16	17							33	545.0000
	B	8	2	17	16							33	545.0000
	C	8	2	21	12							33	585.0000
	D	8	2	15	18							33	549.0000
First Order Interactions	AB	4	4	11	5	6	11					33	303.0000
	AC	4	4	12	4	9	8					33	305.0000
	AD	4	4	6	10	9	8					33	281.0000
	BC	4	4	2	19	15	−3					33	599.0000
	BD	4	4	10	5	7	11					33	295.0000
	CD	4	4	13	8	2	10					33	337.0000
Second Order Interactions	ABC	2	8	3	8	9	−4	−1	7	10	1	33	321.0000
	ABD	2	8	6	0	4	5	5	5	2	6	33	167.0000
	ACD	2	8	7	6	5	3	−1	3	5	5	33	179.0000
	BCD	2	8	3	10	−1	9	7	−5	8	2	33	333.0000
Residual		—	—	—	—	—	—	—	—	—	—	—	—

Fig. 88. Sum of squares method: sheet for calculations.

(2) In this column list the number of individual measurements in each level, or combination, associated with the source listed in column (1). For example, variable A has 8 measurements in each level. Each of the A and B combinations (such as A1B1) has 4 measurements. Each of the A, B, C combinations (such as A1B1C1) has 2 measurements, etc.

(3) In this column list the number of levels, or combinations, associated with each source. The product of (2) and (3) should in each case be equal to "n."

(4) In this column list the totals separately for each level or combination. In the case of individuals, list the individual observations. In the case of variable A, list the total of A1 and the total of A2. For the combination AB list separately the totals of A1B1, A1B2, A2B1 and A2B2. Similarly for all other variables.

(5) In this column write the total of all the numbers in column (4). In each case this should be equal to the grand total of all the data.

(6) In this column write the total obtained by squaring all the numbers in column (4) and then adding the squares. For example, taking the source listed as "individuals":

$$(3)^2 + (4)^2 + (0)^2 + (6)^2 + (0)^2 + (5)^2 + (-1)^2 + (4)^2 + (3)^2 + (-4)^2 + (4)^2 + (-1)^2 + (5)^2 + (0)^2 + (3)^2 + (2)^2 = 9 + 16 + 0 + 36 + 0 + 25 + 1 + 16 + 9 + 16 + 16 + 1 + 25 + 0 + 9 + 4 = 183$$

Carry out the same number of decimal places as in the correction factor at the top of the sheet.

For variable A:

$$(16)^2 + (17)^2 = 256 + 289 = 545.0000$$

Sum of Squares Method (*Continued*)

	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Source of Variation	Sum of Squares ÷ i	Col. (7) − C	Subtract Other Corrections	Corrected Sum of Squares	*df*	Mean Square	Significance
Individuals	183.0000	114.9375	—	114.9375	15	—	—
A	68.1250	.0625	—	.0625	1	.0625	None
B	68.1250	.0625	—	.0625	1	.0625	—
C	73.1250	5.0625	—	5.0625	1	5.0625	—
D	68.6250	.5625	—	.5625	1	.5625	None
AB	75.7500	7.6875	(A) .0625 (B) .0625 .1250	7.5625	1 x 1	7.5625	None
AC	76.2500	8.1875	(A) .0625 (C) 5.0625 5.1250	3.0625	1 x 1	3.0625	None
AD	70.2500	2.1875	(A) .0625 (D) .5625 .6250	1.5625	1 x 1	1.5625	None
BC	149.7500	81.6875	(B) .0625 (C) 5.0625 5.1250	76.5625	1 x 1	76.5625	0.1%
BD	73.7500	5.6875	(B) .0625 (D) .5625 .6250	5.0625	1 x 1	5.0625	None
CD	84.2500	16.1875	(C) 5.0625 (D) .5625 5.6250	10.5625	1 x 1	10.5625	None (or "not quite")
ABC	160.5000	92.4375	(A) .0625 (B) .0625 (C) 5.0625 (AC) 3.0625 (AB) 7.5625 (BC) 76.5625 92.3750	.0625	1 x 1 x 1	.0625	None
ABD	83.5000	15.4375	(A) .0625 (B) .0625 (D) .5625 (AB) 7.5625 (AD) 1.5625 (BD) 5.0625 14.8750	.5625	1 x 1 x 1	.5625	None
ACD	89.5000	21.4375	(A) .0625 (C) 5.0625 (D) .5625 (AC) 3.0625 (AD) 1.5625 (CD) 10.5625 20.8750	.5625	1 x 1 x 1	.5625	None
BCD	166.5000	98.4375	(B) .0625 (C) 5.0625 (D) .5625 (BC) 76.5625 (BD) 5.0625 (CD) 10.5625 97.8750	.5625 111.8750	1 x 1 x 1	.5625	None
Residual = 114.9375 − 111.8750 = 3.0625					15–14	3.0625	—

Fig. 89. Sum of squares method: table of components of variance.

and similarly for all the other sources of variation.

No Sum of Squares is calculated for the residual.

The Sums of Squares must now be tested for significance as shown in Figure 89.

The following notes relate to the numbered columns.

(7) To obtain the values in this column divide column (6) by column (2).

(8) For this column subtract the correction factor C from column (7).

(9) In this column list any of the values from column (10) which are associated with the source of variation in question. Leave column (9) blank in the case of individuals, main effects, and residual. In the case of the interaction AB, fill in the column (10) values for A and B. In the case of the interaction ABC, fill in the column (10) values for A, B, C, AC, AB and BC. Add these values to obtain the total as indicated.

(10) To obtain the values in this column, subtract the total in column (9) from column (8).

(11) In this column list the number of degrees of freedom. In the case of individuals, the degrees of freedom are $n - 1$. For a main effect, the degrees of freedom are the number of levels minus 1. For a first order interaction, the degrees of freedom will be the product of the degrees of freedom associated with the two variables involved. In the case of a second order interaction, the degrees of freedom will be the product of the degrees of freedom for the three variables involved. The number of degrees of freedom in the residual will be the degrees of freedom for individuals minus the total of all degrees of freedom for the other sources of variation.

(12) To obtain the values in this column divide column (10) by column (11). These are "components of variance" for each of the possible sources of variation.

(13) To find the significance of the different sources of variation, use the "*F*-test" as follows:

F-test:

Form various ratios consisting of

$$\frac{\text{Mean Square to Be Tested}}{\text{Residual}}$$

Start by testing the second order interactions. If these are found to be nonsignificant, pool them with the residual, obtain a new residual and use this to test the first order interactions.

In testing the first order interactions, start with the smallest. Pool those which are nonsignificant to obtain various new residuals, and proceed in this manner until all the mean squares have been tested or until one is found to be significant. When a first order interaction is significant, do not test the main effects associated with that interaction.

The method of making an *F*-test is explained in Reference No. 5. See also the example in this book on page 79.

While the rules for pooling the residual will not be given here in detail, the following calculations will show how the pooling is done.

Pooled Residual =	3.0625	Res.
	.5625	BCD
	.5625	ACD
	.5625	ABD
	.0625	ABC
	3.0625	AC
	1.5625	AD
7	9.4375	
	1.3482	

To test for BD:

$$\frac{5.0625}{1.3482} = 3.76 \text{ at } df\ 1, 7$$

Not significant.

Second Pooled Residual =	9.4375	Former Residual
	5.0625	BD
8	14.5000	
	1.8125	

To test for AB:

$$\frac{7.5625}{1.8125} = 4.17 \text{ at } df\ 1,8$$

Not significant.

Third Pooled
Residual $= 14.5000$ Former Residual

$$\begin{array}{rl} & 14.5000 \quad \text{Former Residual} \\ & \underline{7.5625} \quad \text{AB} \\ 9 \,| & \underline{22.0625} \\ & 2.4514 \end{array}$$

To test for CD:

$$\frac{10.5625}{2.4514} = 4.31 \text{ at } df \text{ 1,9}$$

Not significant. (Or "not quite significant.")

Fourth Pooled
Residual $= 22.0625$ Former Residual

$$\begin{array}{rl} & 22.0625 \quad \text{Former Residual} \\ & \underline{10.5625} \quad \text{CD} \\ 10 \,| & \underline{32.6250} \\ & 3.2625 \end{array}$$

To test for BC:

$$\frac{76.5625}{3.2625} = 23.47 \text{ at } df \text{ 1,10}$$

Significant at 0.1% level.

For tables of the Variance Ratio and for a detailed discussion of pooling, significance testing and other calculations involved in the Sum of Squares method, see Reference No. 5, pages 77–97 and 146–149.

Effects which are found to be "significant at the 5% level" are usually called merely "significant"; those at the 1% level "very significant"; and those at the 0.1% level "extremely significant." Some workers identify these with one, two and three asterisks respectively.

Conclusions from this analysis

The result of this analysis is that there is something "extremely significant" about the BC interaction. The Sum of Squares method does not tell us just what combination is significant, or how the B and C effects are related to each other or to the other variables. To obtain this information, it is necessary to go back and study the original data.

CONTROL CHART METHOD

A large amount of information is available on the use of control charts in analyzing multifactor experiments. The control charts have the general advantages of (a) simplicity and (b) the information contained in patterns. The control charts are usually much easier to understand and interpret than other forms of analysis. However, it is necessary to learn certain new techniques in connection with (a) the calculation of control limits and (b) plotting the data.

The following material covers only that part of the control chart analysis which is directly comparable to the Sum of Squares method shown above. Further information is available in the advanced Engineering Courses on this subject which are given from time to time in the various manufacturing locations.

Basis of control chart analysis

When control charts are used in multifactor experiments, it is customary to base the control limits on the "Residual" rather than on a series of sample ranges. This is the only essential difference between this and other uses of control charts. The Residual is the same value which is obtained in the Sum of Squares method, but in plotting the control chart it can be obtained very rapidly. The following instructions apply to a four factor experiment with each factor at two levels.

(1) Visualize the boxes in the experiment as shaded or unshaded according to the following diagram.

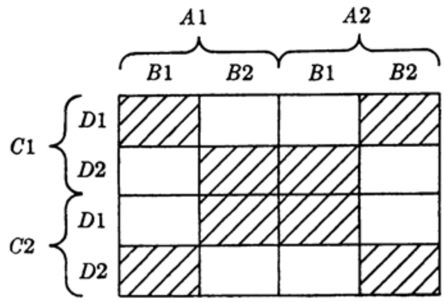

Fig. 90. Diagram for calculating residual: control chart method.

(2) Add the numbers in the shaded squares, subtract from these the numbers in the unshaded squares, and divide by 4. In the present example,

$$\frac{20 - 13}{4} = +1.75$$

The absolute value of this Residual (disregarding signs) is called σ' and is used as the basis for calculating control limits. Note that if we take the square root of 3.0625 (the mean square for the Residual which was obtained at the end of the calculations in Figure 89) we obtain this value of 1.75.

For the control chart analysis, it is not necessary to make the other calculations which were used in the Sum of Squares method. However, if it is desired to reproduce all the other numerical effects, this can be done as indicated under the heading "Optional Calculations" below.

To obtain a value of $\bar{R}$ for use on the range chart, first multiply σ' by the appropriate d_2 factor for the sample size to be used. In the present example we would probably decide to use samples of 2.

$$\bar{R} = d_2 \times \sigma' = 1.128 \times 1.75 = 1.97$$

Then use this value of $\bar{R}$ in the usual way to obtain standard control limits. Complete directions are given on pages 107–109.

Optional calculations

To reproduce the numerical effects obtained in Figure 89, proceed as follows:

(1) Identify each box in the experiment as shown in Figure 91.

(2) Calculate the numerical effects as shown in Figure 92. "Res." stands for the algebraic value of Residual already calculated (+1.75). Add or subtract as indicated, using the values in the designated boxes. For example:

$$\text{A effect} = +1.75 + 1/2(-3 + 0 - 5 + 4 - (-4) - 1 - 5 + 3) = +.25$$

Each of these values is the square root of the corresponding value in Figure 89.

These calculations show that the largest potential effect is the BC interaction. This information is useful, but not essential, in plotting the control charts.

Method of plotting

The following procedure may be used in plotting the control charts.

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	*a1*	*a2*	*a3*	*a4*
	D2	*b1*	*b2*	*b3*	*b4*
C2	*D1*	*c1*	*c2*	*c3*	*c4*
	D2	*d1*	*d2*	*d3*	*d4*

Fig. 91. Identification of boxes in a four factor experiment.

	Main Effects	
A	Res. + 1/2(−a1 + a3 − b2 + b4 − c2 + c4 − d1 + d3) =	+ .25
B	Res. + 1/2(−a1 + a2 − b3 + b4 − c3 + c4 − d1 + d2) =	− .25
C	Res. + 1/2(−a1 − a4 − b2 − b3 + c1 + c4 + d2 + d3) =	−2.25
D	Res. + 1/2(−a1 − a4 + b1 + b4 − c2 − c3 + d2 + d3) =	+ .75
	First Order Interactions	
AB	Res. + 1/2(+b1 − b2 − b3 + b4 + c1 − c2 − c3 + c4) =	+2.75
AC	Res. + 1/2(+a2 − a4 + b1 − b3 − c2 + c4 − d1 + d3) =	+1.75
AD	Res. + 1/2(+a2 − a4 − b2 + b4 + c1 − c3 − d1 + d3) =	−1.25
BC	Res. + 1/2(+a3 − a4 + b1 − b2 − c3 + c4 − d1 + d2) =	−8.75
BD	Res. + 1/2(+a3 − a4 − b3 + b4 + c1 − c2 − d1 + d2) =	+2.25
CD	Res. + 1/2(+a2 + a3 − b2 − b3 − c2 − c3 + d2 + d3) =	+3.25
	Second Order Interactions	
ABC	Res. + 1/2(−a1 + a2 + a3 − a4 + c1 − c2 − c3 + c4) =	+ .25
ABD	Res. + 1/2(−a1 + a2 + a3 − a4 + b1 − b2 − b3 + b4) =	− .75
ACD	Res. + 1/2(−a1 + a3 + b1 − b3 + c1 − c3 − d1 + d3) =	− .75
BCD	Res. + 1/2(−a1 + a2 + b1 − b2 + c1 − c2 − d1 + d2) =	+ .75

Fig. 92. "Effects" or components of variance: control chart method.

(1) Select two of the factors in the experiment which you particularly wish to study. These may be (a) factors in which you are particularly interested for engineering reasons, (b) factors which appear, in the original data, to be associated with large numerical effects, or (c) factors which show the largest main effects or interactions when calculations are made as in Figure 92. In the present example we would select B and C.

(2) Look up these factors in the Plotting Guide on page 109. The Guide will show (a) the headings to put at the top of the chart, (b) the identification to put at the bottom of the chart and (c) the order in which to plot the data. The order is shown by the series of symbols

a1b1a3b3 etc.,

each symbol referring to one of the experimental boxes in Figure 91.

B1		B2	
C1	C2	C1	C2
a1 b1 a3 b3	c1 d1 c3 d3	a2 b2 a4 b4	c2 d2 c4 d4
A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2
D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2

Fig. 93. Plotting guide for factors B and C.

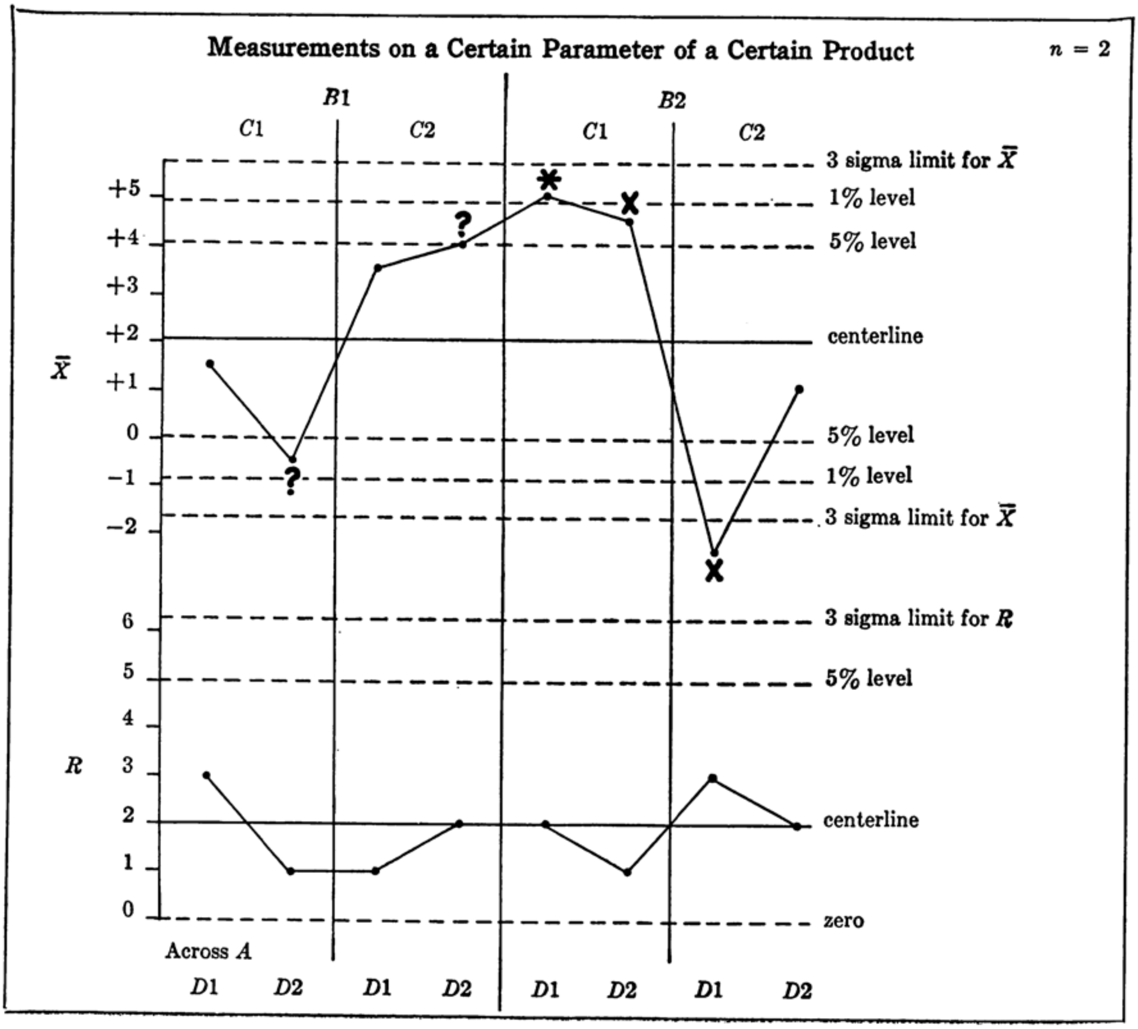

Fig. 94. Control chart analysis: $\bar{X}$ and R chart. Three sets of limits are shown (1% and 5% levels in addition to the 3 sigma limits).

Directions for making this chart are given on pages 107–109.

In plotting the control chart, it is customary to show one or two sets of "inner control limits" in addition to the standard 3σ limits which are used on other control charts. These inner control limits are explained on page 107.

To interpret the chart, first look at the standard 3σ limits and mark x's in accordance with the usual tests. The x's are interpreted as in the case of any other control chart. Then look at the inner control limits and mark any significant points with an asterisk or a question mark as explained on pages 107 and 109. The asterisks and question marks are interpreted in the same way as x's, except that we recognize that the conclusions are less certain.

Drawing conclusions

The conclusions from the experiment are usually obvious, once the chart has been plotted, since each point is completely identified in terms of the variables being studied. However, a detailed analysis of the present example is given below to serve as a guide in interpreting other experiments.

Detailed analysis of the chart in Figure 94

$\bar{X}$ chart

(1) Looking at the 3 σ control limits on this chart, we note that one point is out of control on the low side. A check of the headings at the top of the chart and the identification at the bottom shows that the low point is B2C2D1. Since we wish to avoid low readings on the parameter being studied (page 91), this means that we should avoid the combination B2C2D1.

By checking the list of variables on page 92 we find that the combination to be avoided is Western activation, special cleaning, and short drying.

(2) Still looking at the 3 σ limits, we note that the four points at the top of this chart react to a test for "4 out of 5." This means that these four points are significantly higher than the rest of the data. The desirable combinations are B1C2 (laboratory activation with special cleaning) and B2C1 (Western method with regular cleaning).

The above points, which react to the 3 σ limits, have been marked in the usual way with x's.

(3) Looking now at the inner control limits, which represent lesser degrees of certainty, we note the following:

At the 1% level (that is, with less certainty than if we were basing this on 3 σ limits), the first point under B2C1 is significant. There is a little more evidence of the desirability of B2C1 than there is of the desirability of B1C2.

At the 5% level (that is, with less certainty than if we were basing this on the 1% limits), the point marked B1C1D2 is significant. It may be that, if we are already using the combination B1C1, longer drying will make matters worse.

R Chart

The $\bar{X}$ chart was plotted in such a way as to show the effect of variables B, C and D. The *R* chart will show the effect of the remaining variable, A. The *R* chart compares this variable (plating) directly with the Residual. The Residual is represented by the centerline on the *R* chart.

Since none of the *R* points are significantly different from the Residual, this indicates that there is no significant effect due to plating. If there had been a significant effect in some portion of the *R* chart (for example, in the portion marked B1C2), this would have warned us that there was an interaction between variables A, B and C.

Summary

The conclusions from this experiment are summarized as follows:

(1) It is desirable to use the Western Electric method (B2) but only if it is to be followed by regular cleaning (C1).

(2) Very bad results may be obtained by using the Western method in combination with special cleaning (B2C2).

(3) If by any chance it should be necessary

to use the Western method with special cleaning, at least we should try to make the best of a bad situation by using longer drying (D2).

(4) It does not matter whether we use thick or thin plating.

(5) If we decided to adopt the laboratory method (B1), it would probably be necessary to use special cleaning.

(6) If we wished to use the laboratory method without going to special cleaning, it would probably not be advisable to use prolonged drying.

B-4.5 Comments on this experiment

(1) The control chart explains the conflicting ideas described on page 91. The design engineers said that special cleaning was better. This checks with one part of the chart. The Western engineers did not agree with this idea. This checks with another part of the chart. The shop said longer drying would improve the yield. This is true—provided we are unfortunate enough to be using the combination B2C2. The engineers found just the opposite effect with longer drying. This would be true if they were using B1C1.

(2) The best combination of variables in this process would be Western activation and ordinary cleaning. Note that this was not the conclusion reached in the non-statistical experimenting on page 92.

(3) As a result of this experiment, the variables chosen for the process were:

a. Western activation.
b. Ordinary cleaning.
c. Thin plating (desired for other reasons).

The drying time was not changed.
This combination improved the average considerably, and helped to bring about a significant increase in yield.

B-4.6 General comment

Before attempting to analyze an experiment by the method shown above, study the "Directions for Plotting" on pages 107–109 and the material on "Drawing Conclusions" on pages 111–112.

For further information on the Sum of Squares method, see References No. 5, 15, 16, 17 and 41. For further information on the Control Chart method, see References No. 7, 32, 33, 40 and 47.

B-5 EXPLANATION OF THE FOUR FACTOR ANALYSIS

(With Special Reference to the Control Chart Method)

The following explanation is not essential for the analysis, but will aid in understanding the theory of the factorial design. The engineer is asked to imagine a process which contains initially no variation whatsoever. As the example proceeds, variables are deliberately introduced into this process in such a way that their separate and combined effects can be studied. This will help to show:

(a) The meaning of "balance" in a factorial design.
(b) The fact that it is possible to study the factors separately, even though many factors have been combined in the same experiment.
(c) The meaning of "residual."
(d) The meaning of "interactions."

Because the example starts with zero variation, it is easy to check the effects and also to cross-check the calculations.

B-5.1 Explanation of "factorial design"

Imagine a process which contains no variation whatsoever. A series of measurements from this hypothetical process could be represented as follows:

0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

Fig. 95. Original measurements from a hypothetical process containing zero variation.

When variables are introduced into an experiment on this process, they will tend to produce changes in the original numbers. For example, suppose we introduce a change in a certain alloying temperature, the effect of the change being equivalent to adding 2 to each of the basic measurements. To represent this, divide the experiment vertically into two halves, label the halves A1 and A2, and add 2 to each of the A2 measurements as shown below.

A1		*A2*	
0	0	2	2
0	0	2	2
0	0	2	2
0	0	2	2

Fig. 96. Introduction of Variable A: +2 added to all A2 measurements.

A second variable could be introduced into the same experiment as follows:

> Suppose, in addition to alloying temperature, we are interested in the effect of curing in a bake-out oven overnight. Let the effect of baking be equivalent to subtracting 6 from each of the original measurements. Introduce this second variable into the experiment in the following way:
>
> Let B1 represent the units processed without baking and B2 the units which have been baked. Divide the experiment vertically into quarters. Label the first and third quarters B1 and the second and fourth quarters B2. Now subtract 6 from each of the B2 measurements as shown below. These measurements will contain the combined effect of variables A and B.

A1		*A2*	
B1	*B2*	*B1*	*B2*
0	−6	2	−4
0	−6	2	−4
0	−6	2	−4
0	−6	2	−4

Fig. 97. Introduction of Variable B: −6 added to all B2 measurements.

In the same way, let C1 represent a certain capacitance level and C2 another. The data below would be obtained by adding 4 to each of the C2 measurements:

	A1		*A2*	
	B1	*B2*	*B1*	*B2*
C1	0	−6	2	−4
	0	−6	2	−4
C2	4	−2	6	0
	4	−2	6	0

Fig. 98. Introduction of Variable C: +4 added to all C2 measurements.

Finally, let D1 represent a "bright dip" finish and D2 the regular finish. The data below would be obtained by subtracting 2 from each of the D2 measurements:

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	0	−6	2	−4
	D2	−2	−8	0	−6
C2	*D1*	4	−2	6	0
	D2	2	−4	4	−2

Fig. 99. Complicated data containing four different variables.

We now have a number of variables introduced into the same data. It would be difficult to tell, by looking at the numbers, just what effect was contributed by each variable.

B-5.2 Method of separating the effects of the variables

To separate the effects of the different variables, proceed as follows:

Starting with the complicated data in Figure 99, calculate the average for each level of each variable and record these averages as shown in Figure 100. The "Effect" of the variable is the amount obtained by subtracting Level 1 from Level 2.

Note that these are the numbers which we added to the data originally.

It would now be possible to "remove" the

Variables		*Averages*		*Main Effects*
A2 − A1	=	0 − (−2)	=	+2
B2 − B1	=	−4 − (+2)	=	−6
C2 − C1	=	+1 − (−3)	=	+4
D2 − D1	=	−2 − 0	=	−2

Fig. 100. Main effects.

effect of any variable from the data by subtracting the calculated difference in average from each of its Level 2 measurements. For example, to remove the effect of Variable B, start with the data in Figure 99 and subtract −6 from (or add +6 to) each of the B2 measurements.

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	0	0	2	2
	D2	−2	−2	0	0
C2	*D1*	4	4	6	6
	D2	2	2	4	4

Fig. 101. Effect of removing Variable B.

To remove the effect of Variable C (in addition to B) start with the data in Figure 101 and subtract 4 from each of the C2 measurements.

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	0	0	2	2
	D2	−2	−2	0	0
C2	*D1*	0	0	2	2
	D2	−2	−2	0	0

Fig. 102. Effect of removing Variable C.

To remove the effect of Variable D (in addition to B and C) start with the data in Figure 102 and subtract −2 from each of the D2 measurements. See Figure 103.

This uncovers the original simple effect of Variable A, uncomplicated by any other variable. If we now removed Variable A also, in the same manner, all the numbers would be reduced to the original zeros.

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	0	0	2	2
	D2	0	0	2	2
C2	*D1*	0	0	2	2
	D2	0	0	2	2

Fig. 103. Effect of removing Variable D: all variables are now removed except A.

B-5.3 Meaning of "Residual"

Suppose we had started, not with zeros, but with other numbers such as 1, 2 and 3. Assume that these numbers are scattered more or less uniformly through the experiment, in such a way that the numbers in one portion of the experiment are about the same as the numbers in other portions of the experiment. It would be possible to introduce variables into this experiment just as we did before, then calculate the effect of each variable and remove these effects from the data. This time the numbers would be reduced, not to zero, but to something closely approximating the original 1's, 2's and 3's.

In the same way, we could start with any set of numbers (provided only they are uniformly scattered) and reproduce these numbers more or less closely by calculating and removing effects.

In a real experiment we assume that we start with random variation (uniformly scattered). We add several variables to this in conducting the experiment. We calculate the effects of these variables in the manner shown above. *When the effects are removed, we obtain an estimate of the original random variation.*

The variability left in the numbers after the known effects are removed is called the "Residual." It is a measure of random variation. Some writers refer to the Residual as the "experimental error."

If any of the identified variables in the experiment are significantly larger than the Residual, this is statistical evidence that these variables must have a real effect. If any of the variables are not significantly larger than the Residual, we conclude that their apparent effects may be due to chance, or to ordinary random fluctuation.

B-5.4 Interactions

When variables have a simple direct effect like those introduced in paragraph B-5.1 (that is, when we can represent the effect of A2 by adding or subtracting a constant amount from all the A2 measurements) such variables are said to have a "main effect." When the effect of a variable is more complicated, it is said to exhibit "interaction." A simple example of interaction would be the following:

> In paragraph B-5.1 we introduced a certain variable, B, representing the effect of curing in a bake-out oven overnight. This variable had a "main effect" indicated by subtracting 6 from all B2 measurements. Imagine now that the effect of baking is not always the same. Suppose it normally tends to *increase* the measurements by 6, but if the material has already been processed at alloying temperature A2, then baking will *reduce* the measurements by an amount equal to 12. We then say that there is an "interaction" between Variable B and Variable A.
>
> Interaction is defined as the effect of one variable acting on another. It can also be defined as the effect produced by two variables acting in combination where the effect would not be produced by either variable acting alone. These are two ways of saying the same thing.

To represent the above interaction between B and A, start with the hypothetical data in Figure 95 and add 6 to all of the B2 measurements. Then subtract 12 from the A2B2 measurements only (those in the last column). Then introduce the main effects of variables A, C and D in the same manner as in paragraph B-5.1.

		A1		*A2*	
		B1	*B2*	*B1*	*B2*
C1	*D1*	0	6	2	−4
	D2	−2	4	0	−6
C2	*D1*	4	10	6	0
	D2	2	8	4	−2

Fig. 104. Effect of introducing 3 main effects and one interaction.

The results of this further complication of the data are shown in Figure 104.

By working backward from Figure 104, the numerical effect of the interaction can be discovered as follows. First find the averages of all possible combinations of B and A. There are four of these:

$$\begin{aligned} A1B1 &= +1 \\ A1B2 &= +7 \\ A2B1 &= +3 \\ A2B2 &= -3 \end{aligned}$$

Then take the difference between the A1B2 average and the A1B1 average:

$$A1B2 - A1B1 = 7 - 1 = 6$$

This is the effect of B, considering the A1 data only.

Now take the difference between the A2B2 average and the A2B1 average:

$$A2B2 - A2B1 = (-3) - 3 = -6$$

This is the effect of B, considering the A2 data only.

Finally take the difference between these two differences. This is the interaction.

$$(A2B2 - A2B1) - (A1B2 - A1B1) = (-6) - 6 = -12$$

This shows that the A2B2 measurements must have been reduced by 12.

The above expression is usually written as follows:

Interaction between A and B:
$A2B2 - A2B1 - A1B2 + A1B1$
$(-3) - (+3) - (+7) + (+1) = -12$

In the same way, it would be possible to calculate the interactions between all other combinations of variables. In the present example we did not actually introduce any other interactions. Consequently these interactions should all turn out to be zero.

Between A and C:
$A2C2 - A2C1 - A1C2 + A1C1$
$(+2) - (-2) - (+6) + (+2) = 0$

Between A and D:
$A2D2 - A2D1 - A1D2 + A1D1$
$(-1) - (+1) - (+3) + (+5) = 0$

Between C and B:
C2B2 − C2B1 − C1B2 + C1B1
(+4) − (+4) − (0) + (0) = 0

Between C and D:
C2D2 − C2D1 − C1D2 + C1D1
(+3) − (+5) − (−1) + (+1) = 0

Between D and B:
D2B2 − D2B1 − D1B2 + D1B1
(+1) − (+1) − (+3) + (+3) = 0

B-5.5 How to remove the effect of interactions

To remove from the data the effect of an interaction, subtract the indicated amount from the measurements which represent Level 2 for both variables. For example, to remove the AB interaction found above, subtract −12 from (or add +12 to) the A2B2 measurements.

B-5.6 Correcting the "main effects" by taking account of interactions

When main effects were calculated in paragraph B-5.2, these were not complicated by the presence of any interactions. If interactions are present, however, they will tend to "throw off" or distort the corresponding main effects. An example of this is the following: Suppose we try to determine the main effect of variable A in Figure 104. We find

$$A2 - A1 = 0 - 4 = -4$$

This is not the effect which we really introduced into variable A. *The A2 average has been changed by the presence of the interaction.*

A simple additional calculation, however, can easily remove this difficulty. Merely subtract from the A2 average one-half of any interactions involving A.

Corrected "A" main effect =

$$A2 - A1 - \frac{\text{(the sum of any interactions involving A)}}{2}$$

$$= 0 - 4 - \left(\frac{-12}{2}\right)$$

$$= -4 + 6$$

$$= +2$$

This is the true main effect which we introduced into Variable A.

In the same way, corrected effects can be calculated for all the other variables. This has been done in Figure 105. Note that in all cases we obtain the values which were originally introduced.

Corrected Main Effects

Variable	Level 2 − Level 1 − $^1/_2$ of any interactions which involve this variable	
A	$0 - (+4) - \left(\frac{-12}{2}\right)$	= +2
B	$+2 - (+2) - \left(\frac{-12}{2}\right)$	= +6
C	$+4 - 0 - 0$	= +4
D	$+1 - (+3) - 0$	= −2

Fig. 105. Corrected main effects for all four variables.

Since all of these effects are the true values (those actually introduced), it would be possible to subtract them all from the data and thus get back to the original zeros. To do this, start with Figure 104, subtract 2 from the A2 measurements, 6 from the B2, 4 from the C2 and (−2) from the D2. Finally, subtract (−12) from the A2B2 column. All the numbers will now be reduced to zero.

In a real experiment, they would be reduced to the Residual.

B-5.7 Higher order interactions

It is possible for the data to be complicated by other interactions than the simple one shown above. For example, it may take a combination of three or more variables to produce a certain effect on the process. Interactions involving more than two variables are called "higher order interactions." First order interactions involve two variables, second order interactions involve three variables, third order interactions involve four variables, etc. In a four factor experiment it is possible to calculate (a) Main Effects, (b) First Order interactions, (c) Second Order interactions and (d) the Third Order interaction involving all four factors. While these calculations will not be given in detail, the method is a simple extension of the method already explained.

Second Order interaction between A, B and C = A2B2C2 − A2B1C2 + A2B1C1 − A2B2C1 + A1B1C2 − A1B2C2 − A1B1C1 + A1B2C1.

Using the data in Figure 104, (−4) − (+6) + (0) − (+2) + (+8) − (−2) − (+2) + (+4) = 0.

This shows that we did not introduce any Second Order interaction ABC.

Third Order interaction between A, B, C and D = A1B1C1D1 − A1B2C1D1 − A2B1C1D1 + A2B2C1D1 − A1B1C1D2 + A1B2C1D2 + A2B1C1D2 − A2B2C1D2 − A1B1C2D1 + A1B2C2D1 + A2B1C2D1 − A2B2C2D1 + A1B1C2D2 − A1B2C2D2 − A2B1C2D2 + A2B2C2D2.

Using the data in Figure 104, (0) − (+6) − (+2) + (−4) − (−2) + (+4) + (0) − (−6) − (+4) + (+10) + (+6) − (0) + (+2) − (+8) − (+4) + (−2) = 0.

This shows that we did not introduce any Third Order interaction ABCD.

As before, interactions would tend to complicate the other effects of the variables, but this could be corrected by a simple calculation similar to that in paragraph B-5.6. In each case we would be able to find the real (corrected) effect of each variable; and if all these effects were removed, the data would be reduced to the original zeros.

B-5.8 How this can be used in analysis

If we actually went through all this, calculating the effects and obtaining their true values, we would be able to remove any effects from the data at will and study the remainder. In a four factor factorial experiment containing only one measurement per box, we might choose to remove

All 4 Main Effects
All 6 First Order interactions
All 4 Second Order interactions

and leave only the Third Order interaction. (We would not attempt to remove the Third Order interaction because it would not be possible to separate this from the Residual.)

The data from which all possible effects have been removed are called the "fully reduced data".

The reader would be able to verify, by calculating the effects, correcting them and actually removing them from the data, that

$\bar{R}$ for samples of 2 from the fully reduced data would be ABCD/8*, and

σ for the same data would be $\bar{R}/2$.

The calculated σ should be multiplied by $\sqrt{n}$ (where n is the number of boxes in the experiment) to take care of the restrictions imposed on the data in removing so many effects.

$$\text{Residual} = \sigma \times \sqrt{n} = \bar{R}/2 \times \sqrt{16} = \text{ABCD}/16 \times \sqrt{16}$$
$$= \text{ABCD}/4$$

This gives the same value of Residual that is obtained in the Sum of Squares method. The shaded boxes shown in Figure 90 are merely a convenient way of calculating ABCD.

B-5.9 Other possibilities

In the control chart method it is not necessary to remove all the main effects and interactions in the manner shown above. It is possible to remove any desired variables or combinations of variables and study the effect of this on the remainder. This makes the control chart a very flexible method.

Estimates of Residual will involve larger or smaller numbers of degrees of freedom depending on the number of effects that are removed. The method shown above is the one which corresponds directly to the Sum of Squares method prior to the pooling of any effects with the Residual.

B-5.10 Summary of technical terms

(1) The *"characteristic to be plotted"* is the characteristic which is actually measured and whose measurements are recorded in the boxes provided for data.

(2) A *"factor"* is a variable which may or may not have an effect on the characteristic to be plotted, but which has been selected as an object of study in the experiment. We run the experiment to discover the possible effect of one or more factors. An experiment containing four variables is called a four-factor experiment.

---

* This will be true regardless of which variable is "summed across". See page 107.

(3) A *"level"* is a particular condition or state of one of the factors being studied. The different conditions or states of the same variable are called its different "levels." For example, the variable C may appear in the experiment at two levels, C1 and C2. The *presence* of a certain condition may be considered one level, and its *absence* may be considered another level.

(4) The *"Residual"* in the experiment refers to the basic data which existed (or might have existed) in the process prior to the introduction of any of the factors in the experiment. The Residual is the estimated variability in the original basic data. See paragraphs B-5.3 and B-5.9.

(5) A *"main effect"* is a simple, direct, consistent effect on the characteristic being plotted. For example, if changing from D1 to D2 has a definite tendency to make the measurements higher, regardless of the presence or absence of other variables, there is a D main effect.

(6) Sometimes variables do not have a particular effect when acting alone, but produce that effect only when acting in combination with other variables. Such variables are said to exhibit *"interactions"* rather than main effects. Interactions were discussed in paragraphs B-5.4 to B-5.7.

(7) The different methods of forming samples in an experiment are spoken of as *"summing across"* variables. We say we are "summing across" a variable when we include, in the same sample, measurements representing different conditions of the same variable. For example, if the experiment covers conditions C1 and C2, we are summing across C when we include a C1 measurement and a C2 measurement in the same sample. Each different method of forming samples from the boxes (horizontally, vertically, skipping one box, etc.) may result in summing across a different variable.

B-6 DIRECTIONS FOR PLOTTING

The following are directions for plotting control charts in simple factorial experiments. The instructions include not only the calculation of standard 3 σ limits, but also the calculation of "inner control limits" at various "significance levels." The significance of the inner control limits is as follows. (All percentages are based on a normal distribution.)

On the $\bar{X}$ chart:

- 3 sigma limits correspond to 0.1% level (approx.)
- 2.33 sigma limits correspond to 1% level (approx.)
- 1.65 sigma limits correspond to 5% level (approx.)

On an R chart for samples of 2:

- 3 sigma limits correspond to 1% level (approx.)
- 2 sigma limits correspond to 5% level (approx.)

The inner control limits are used as follows:

(1) Points which react at the 5% level are less certain than those which react at the 1% level (1 in 20 chances of being wrong as compared with 1 in 100). In the same way, points at the 1% level are less certain than points at the 3 σ limits.

(2) Points which react to the 3 σ limits are marked with an "x" in the usual way. Points which react at the 1% level are marked with an asterisk (*) and points at the 5% level with a question mark (?) to distinguish them from points which react to the standard 3 σ limits.

B-6.1 Experiment with four factors, two levels, one measurement per box

Preliminary calculations

(1) Calculate the Grand Average of all the data ($\bar{\bar{X}}$).

(2) Calculate the Residual. This can be done by using the shaded boxes in Figure 90.

(3) Take σ' as the absolute value of the Residual, disregarding signs.

$\bar{X}$ and R chart for n = 2

(1) Calculate the centerline for the R chart

as follows:

$$\bar{R} = d_2 \times \sigma' = 1.128\ \sigma'$$

(2) Calculate control limits for the R chart as follows:

Upper 3 sigma limit $= 3.267\ \bar{R}$ (or $3.68\ \sigma'$).
Upper 1% limit: The 1% limit in this case is so close to 3 σ that no separate 1% limit is calculated.
Upper 5% limit $= 2.51\ \bar{R}$ (or $2.83\ \sigma'$).
Lower control limit $= 0$.

(3) Calculate control limits for the $\bar{X}$ chart as follows:

Upper 3 sigma limit $= \bar{\bar{X}} + 1.88\ \bar{R}$ (or $+\ 2.12\ \sigma'$).
Upper 1% limit $= \bar{\bar{X}} + 1.46\ \bar{R}$ (or $+\ 1.65\ \sigma'$).
Upper 5% limit $= \bar{\bar{X}} + 1.03\ \bar{R}$ (or $+\ 1.16\ \sigma'$).

The lower control limits are obtained similarly, except that the quantities are subtracted from $\bar{\bar{X}}$ instead of being added to it.

(4) Plot the control chart as follows:

(a) Decide on two factors which you particularly wish to study.
(b) Select the appropriate diagram from page 109.
(c) Set up a control chart which shows, at the top, the combination of variables given at the top of the diagram.
(d) Before plotting the chart, consider the two variables shown at the bottom of the diagram. Decide which of these variables is probably less important. Strike out the line of identification corresponding to the less important variable.
Show at the bottom of the control chart the identification which remains.
(e) Finally, consider the row of symbols at the center of the diagram. These symbols refer to the boxes in the experiment, as shown in Figure 91. Each symbol can be translated into one of the numbers in the original data.
If, in step (d), you struck out the *last* line of identification, form samples by taking these symbols (or numbers) in successive pairs. That is, use the first two numbers for the first sample, the next two numbers for the second sample, etc.
But if, in step (d), you struck out the *next to the last* line of identification, form samples by taking the symbols alternately. That is, use the first and third numbers for the first sample; the second and fourth numbers for the second sample; the fifth and seventh numbers for the third sample, etc.
(f) Calculate $\bar{X}$ and R for each sample in the usual way, and plot these values on the control chart.

Example

Consider the experiment shown in Figure 87 on page 92.

$$\bar{\bar{X}} = 2.06$$
$$\sigma' = 1.75$$
$$\bar{R} = 1.128 \times 1.75 = 1.97$$

Control limits for R chart:

$3.68 \times 1.75 = 6.44$ (3 sigma level)
$2.83 \times 1.75 = 4.95$ (5% level)

Control limits for $\bar{X}$ chart:

$2.06 \pm (2.12 \times 1.75) = 2.06 \pm 3.71$ (3 sigma level)
$2.06 \pm (1.65 \times 1.75) = 2.06 \pm 2.89$ (1% level)
$2.06 \pm (1.16 \times 1.75) = 2.06 \pm 2.03$ (5% level)

Points are plotted as follows. The steps are lettered to correspond to the instructions given above.

(a) We have previously decided that we would like to plot B and C.
(b) We therefore select Diagram No. 3.
(c) We set up a control chart in accordance with this diagram as shown on page 99.
(d) The two variables at the bottom of the diagram are A and D. Suppose we decide we are least interested in A. Strike out the upper line of identification, consisting of A's, and show only the D's on the control chart. See page 99.
(e) Since the next to the last line of identification was eliminated, we must form

samples by taking the data alternately. The symbols in Diagram No. 3 are:

a1b1a3b3   c1d1c3d3   a2b2a4b4   c2d2c4d4

Our samples will therefore be:

a1	b1	c1	d1	a2	b2	c2	d2
a3	b3	c3	d3	a4	b4	c4	d4

Note that this results in "summing across A"—that is, including in one sample an A1 measurement and an A2 measurement. This will cause variable A to appear in the R chart but not in the $\bar{X}$ chart. Label the R chart "Across A."

When we translate these symbols into the data in Figure 87 we get the following:

3	0	3	5	4	5	−4	0
0	−1	4	3	6	4	−1	2

These are the samples which will be plotted on the control chart.

(f) The $\bar{X}$ and R points for the above samples are:

$\bar{X}$	1.5	−0.5	3.5	4.0	5.0	4.5	−2.5	1.0
R	3	1	1	2	2	1	3	2

These are the points which are plotted on page 99.

Marking x's and other significant points

In looking at these charts, consider first the 3 σ limits. Apply the standard tests for unnatural patterns and mark any significant point with an "x". See pages 182–183.

Next, consider the 1% and 5% limits. Do not apply the tests for "2 out of 3" or "4 out of 5", but consider it significant if

(a) a single point exceeds the limit in question, or

(b) the average of two related points would be more than $^7/_{10}$ of the distance from the centerline to the limit in question.

Rule (b) is derived from the fact that, if we plotted the averages of samples having twice as much data, the control limits applying to these averages would have their width divided by $\sqrt{2}$.

$$\frac{1}{\sqrt{2}} = \text{approximately } .7$$

Use asterisks or question marks for these special significant points as indicated on page 107.

Guide for plotting

The diagrams referred to in the preceding instructions are shown at the bottom of this page. The symbols in the center refer to boxes, as in Figure 91 on page 98.

Diagram 1. A and C.

A1		A2	
C1	C2	C1	C2
a1 b1 a2 b2	c1 d1 c2 d2	a3 b3 a4 b4	c3 d3 c4 d4
B1 B1 B2 B2	B1 B1 B2 B2	B1 B1 B2 B2	B1 B1 B2 B2
D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2

Diagram 2. B and D.

B1		B2	
D1	D2	D1	D2
a1 c1 a3 c3	b1 d1 b3 d3	a2 c2 a4 c4	b2 d2 b4 d4
A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2
C1 C2 C1 C2	C1 C2 C1 C2	C1 C2 C1 C2	C1 C2 C1 C2

Diagram 3. B and C.

B1		B2	
C1	C2	C1	C2
a1 b1 a3 b3	c1 d1 c3 d3	a2 b2 a4 b4	c2 d2 c4 d4
A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2
D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2

Diagram 4. A and D.

A1		A2	
D1	D2	D1	D2
a1 c1 a2 c2	b1 d1 b2 d2	a3 c3 a4 c4	b3 d3 b4 d4
B1 B1 B2 B2	B1 B1 B2 B2	B1 B1 B2 B2	B1 B1 B2 B2
C1 C2 C1 C2	C1 C2 C1 C2	C1 C2 C1 C2	C1 C2 C1 C2

Diagram 5. A and B.

A1		A2	
B1	B2	B1	B2
a1 b1 c1 d1	a2 b2 c2 d2	a3 b3 c3 d3	a4 b4 c4 d4
C1 C1 C2 C2	C1 C1 C2 C2	C1 C1 C2 C2	C1 C1 C2 C2
D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2	D1 D2 D1 D2

Diagram 6. C and D.

C1		C2	
D1	D2	D1	D2
a1 a2 a3 a4	b1 b2 b3 b4	c1 c2 c3 c4	d1 d2 d3 d4
A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2	A1 A1 A2 A2
B1 B2 B1 B2	B1 B2 B1 B2	B1 B2 B1 B2	B1 B2 B1 B2

B-6.2 Experiment with three factors, each factor at two levels, one measurement per box

Obtain the Residual as follows. Visualize the boxes in the experiment as shaded or unshaded according to the diagram shown below.

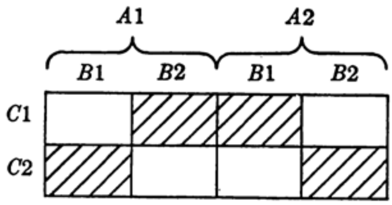

Fig. 106. Calculating residual for a three factor experiment.

Add the numbers in the shaded squares, subtract from these the numbers in the unshaded squares, and divide the result by $\sqrt{8}$.

Use this in the standard manner to obtain control limits. Follow the general method used for the Four Factor experiment as given on pages 107–109.

B-6.3 Experiment with five factors, each factor at two levels, one measurement per box

Obtain the Residual as follows. Visualize the boxes in the experiment as shaded or unshaded according to the diagram shown in Figure 107.

Add the numbers in the shaded squares, subtract from these the numbers in the unshaded squares and divide the result by $\sqrt{32}$.

Use this in the standard manner to obtain control limits. Follow the general method used for the Four Factor experiment as given on pages 107–109.

B-6.4 Experiments containing factors at more than two levels

When one of the factors in an experiment occurs at more than two levels, the experiment becomes, in effect, a combination of simpler experiments. For example, we might have an experiment containing three factors (A, B and C) with factor A occurring at three levels instead of two. This experiment can be regarded as a combination of three experiments —one comparing A1 with A2, another comparing A2 with A3, and a third comparing A1 with A3. The separate (simple) experiments are shown in Figure 108.

While 24 boxes are shown in Figure 108, it would only be necessary to obtain a total of 12 measurements. The four A1 measurements are entered first in section 1 and then copied for use in section 2. The same is done for A2 and A3.

This type of experiment may be analyzed in separate sections, if desired, following the method in paragraph B-6.2. It is also possible to calculate a "combination" Residual which can then be used for all three sections. The method is as follows:

(1) For each section, add the numbers in the shaded squares and subtract from these the numbers in the unshaded squares. Square the result and divide by 8.

(2) Average the values obtained in (1) and take the square root of the average.

This gives the same value of Residual which would have been obtained in the Sum of Squares method. There are other shortcuts for multi-level experiments, but these are beyond the scope of the present Handbook.

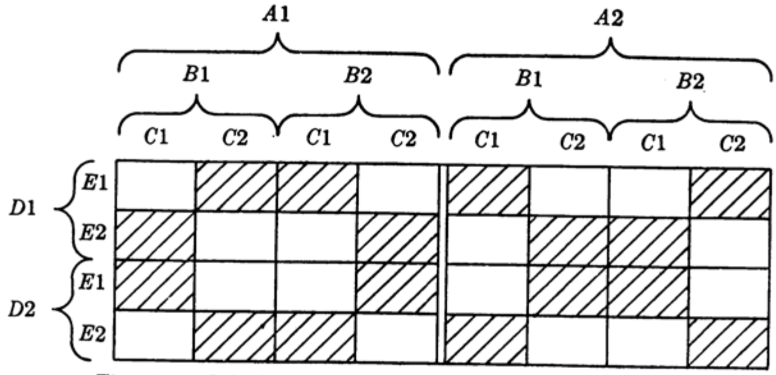

Fig. 107. Calculating residual for a five factor experiment.

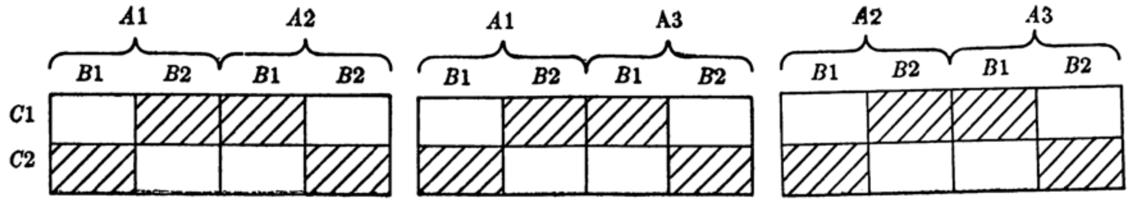

Fig. 108. Arrangement of data for a three factor experiment with Factor A occuring at three levels.

B-6.5 Experiments containing more than one measurement per box

The foregoing instructions were for experiments which have only one measurement per box. If more than one measurement is available for each box, the instructions are similar, except that it may be possible to use some other method of finding the Residual. In general, the data may be treated in three different ways, as follows. Before deciding on the appropriate treatment, use a control chart to test for the presence of assignable causes within boxes. If there are no assignable causes within boxes, use methods (1), (2) or (3). If assignable causes are found within boxes, use method (2) or (3) but not method (1).

(1) *Using the average of the measurements in each box.*

In this treatment, consider the average of each box as if it were a single measurement, and analyze the experiment in the same way as when there is one measurement per box. It is desirable to have the same number of measurements in each box. In drawing conclusions, remember that these are averages and that the conclusions will apply to averages also.

(2) *Forming samples within each box.*

In this treatment, use the measurements within a single box to form one or more "samples." Calculate $\bar{X}$ and R for each sample in the usual way. Add the values of R for all samples in the experiment and calculate $\bar{R}$. Use this value of $\bar{R}$ as if it were the $\bar{R}$ obtained from the Residual.

(3) *Treating the measurements as a separate factor in the experiment.*

In this treatment, identify the sets of measurements as M1, M2 etc., and treat "M" as an additional factor in the experiment. For example, if this is a four factor experiment containing variables A, B, C and D, treat it as a five factor experiment containing variables A, B, C, D and M. In some cases there may be more variability from measurement to measurement than there is between boxes.

B-7 DRAWING CONCLUSIONS FROM EXPERIMENTAL CONTROL CHARTS

B-7.1 Preliminary analysis

Contrary to the expectation of many engineers, experimental charts are easy to interpret once they have been plotted. The patterns are marked with x's as in the case of any control chart. They are classified as stratification, mixture, freaks, sudden shift etc. as in any process capability study. The tracing of causes is generally simpler on the experimental charts, since changes in the pattern can be associated immediately with the particular variables included in the experiment. The interpretation is simplest when the analysis includes only one chart.

If a number of charts have been plotted, or a number of experiments have been run, it is sometimes helpful to combine the conclusions and reduce them to their simplest form. One method of doing this is the following.

(1) Considering each chart as a whole, read the patterns as in any process capability study.

(2) As each conclusion is reached, record it on a suitable form as shown in Figure 109. This form may contain the conclusions from an entire group of control charts.

(3) When duplicate conclusions are obtained, as shown by duplicate entries in the same column, strike out one of the duplicates so

as to keep the outstanding conclusions in as simple a form as possible. See Figure 109.

HIGH	LOW	MORE UNIFORM	LESS UNIFORM
A1D2 A1B1C1 A1B1C2 ~~A1D2~~	A2D2	C1	C2

Fig. 109. Form for recording significant effects found on control charts.

(4) If, in a single column, entries occur which are identical, except that they contain all possible levels of one of the variables, strike out the variable which occurs at all levels. Afterward, since this will leave duplicates in the same column, strike out all but one of the duplicates so as to leave the simplest possible conclusions.

This can be illustrated by an example. Suppose we have the following entries, all in the same column:

A1D2
A1B1C1
A1B1C2
A1D2

First strike out one of the "A1D2" entries, since this is a duplicate. This leaves:

A1D2
A1B1C1
A1B1C2

In two of the entries which are otherwise identical (A1B1C1 and A1B1C2), the variable C occurs at all possible levels. Since we find the same effect at all levels of C, it is obvious that this effect must be due to other variables and not C. Strike out C1 and C2. This leaves:

A1D2
A1B1
A1B1

Since we now have duplicates in the same column (two entries of A1B1), one of the duplicates should be eliminated. This leaves:

A1D2
A1B1

These are the conclusions reduced to their simplest possible form.

(5) After the analysis is completed and all conclusions have been noted and reduced to their simplest form, express the conclusions verbally in terms of the real variables. Check the conclusions by referring back to the plotted charts. Be sure you are fully aware of the evidence on which these conclusions are based.

B-7.2 Final conclusions

All experiments of the type discussed here involve minimum amounts of data. They should be used as a means of obtaining quick indications of the best avenues to explore further. Variables that are found to be significant in the experiment are probably important variables. They can be used to improve the process, reduce costs or explain effects that were previously baffling and unexplainable.

On the other hand, variables which do not show up as significant in the experiment are not necessarily unimportant. The amounts of data used in designed experiments may be sufficient to establish certain variables as significant, but it requires much larger amounts of data to establish the absence of significance. This is true whether the experiment is analyzed by control charts or by any other method.

In any case the engineer should design his experiments on the basis of previous process capability studies, and should check all conclusions from his experiments by making other process capability studies. He should guard against the temptation to substitute conclusions from this quick type of experiment for the broader and more reliable analysis of the process which is included in a process capability study.

B-8 SOME SUGGESTIONS ON PLANNING THE EXPERIMENT

B-8.1 The problem

Define the problem as specifically as possible before starting the experiment. Consult others who have expert knowledge of this job or of the planning of designed experiments. During the planning, decide whether to measure one or several characteristics on the experimental units. By measuring the units for a number of characteristics, it is often possible to select an optimum set of variables considering their effect on all characteristics simultaneously. If the problem is complicated, attempt to subdivide it and handle the subdivisions individually.

B-8.2 Type and quantity of data

The characteristics studied may be "variable" in nature (that is, capable of measurement along a continuous scale) or they may be discrete or non-variable in nature (that is, capable of being classified into limited categories only). Characteristics of the latter type include those which are measured with attributes gages on a go, no-go basis and those which are indicated only by the presence or absence of a given condition (for example, the units did or did not crack, did or did not fail on a life test, etc).

There is also a third type of characteristic which is intermediate between the continuously variable and the entirely discrete. These are the characteristics which are ordinarily considered discrete in nature but which can nevertheless be measured on some sort of crude "semi-variables" scale. Such characteristics may include cracks, burrs, extent of warpage or damage, depth of nicks or scratches and many other characteristics for which it is possible to distinguish degrees if not actual measurements.

If the characteristic is measurable on a continuous scale, it may be necessary to process only one unit for each "box" in the table of experimental results. If the characteristic is non-measurable, it will be necessary to produce a number of units under each designated set of conditions and record in the box the percentage or count which did or did not contain the characteristic in question. The number of units to be processed for each box depends on how many are needed to obtain usable counts or percentages. The number in each group should be large enough so that most of the boxes in the table will contain a number other than zero. It is desirable to process the same number of units for each box.

Attributes data may be analyzed with p-charts or c-charts, as explained in Reference No. 32; or it is possible to use an $\bar{X}$ and R chart as explained on page 198. Most engineers find it convenient to use the $\bar{X}$ and R chart, following the same procedures for both variables and attributes data.

Unless there is good reason for doing otherwise, the experiment should not be reduced to fewer than a total of sixteen observations. This is necessary to permit reasonable estimates of the residual or experimental error. Better estimates of residual may be obtained if the experiment provides more than one observation per box.

Occasionally it is convenient to process more units than will be needed for the experiment and select the experimental units at random from the larger group. In this event, use dice, shuffled cards or a table of random numbers as the basis for selecting the units, in order to avoid unconscious bias in making the selection.

B-8.3 Reliability of measurements

The person conducting the experiment should take special precautions to make sure that the measurements are reliable. In many process capability studies the initial measurements show erratic and unexplainable patterns, particularly on the R chart. These are likely to reflect problems in measurement at least as often as they reflect difficulties in the product.

In a designed experiment, measurement peculiarities are even more important because the amount of data is very small and all conclusions are based on a few measurements obtained in a brief period of time. If practice is necessary in using the measuring instrument (or using the standard adopted as a basis for classification), this should be carried out prior to the time when the experimental data are collected. In particular, make certain that the standards of measurement are not allowed to change during the experiment. If gages and test sets must be replaced, re-calibrated or overhauled, start again to collect the experimental data.

B-8.4 Selection of variables

In designing the experiment, make a list of all variables which are suspected of being able