EXPERIMENTAL DESIGN

How To Recognize A Split-Plot Experiment

by **Scott M. Kowalski and Kevin J. Potcner**

The application of statistically designed experiments is becoming increasingly important in organizations engaged in Six Sigma and other quality initiatives. In conducting these experiments and analyzing the resulting data, experimenters become aware of the treatment structure of the design: the number of factors to be studied and the various factor level combinations.

In 50 Words Or Less

- **Not incorporating the experimental approach into an analysis can result in incorrect conclusions.**
- **One type of statistical experimental design, known as the split-plot, is often more common in experimental situations than the completely randomized design.**
- **Several examples will help practitioners recognize the split-plot design.**

For example, most practitioners know a 2^3 full factorial design consists of three factors, each at two levels, where all eight treatment combinations are studied. However, practitioners often neglect the details of how the experimental runs are performed and thus fail to see how this component, along with the treatment structure, determines which statistical approach to use.

Most would choose to run the eight treatment combinations in a completely randomized order, known as a 2^3 full factorial completely randomized design. Unfortunately, limitations involving time, material, cost and experimental equipment can make it inefficient and, at times, impossible to run a completely randomized design. In particular, it may be difficult to change the level for one of the factors. In this case, practitioners typically fix the level of the difficult-to-change factor and run all the combinations of the other factors—the split-plot design.

Recognizing a Split-Plot Design

Split-plot experiments began in the agricultural industry. Because one factor in the experiment is usually a fertilizer or irrigation method, it can only be applied to large sections of land called whole plots. The factor associated with this is therefore called a whole plot factor.

Within the whole plot, another factor, such as seed variety, is applied to smaller sections of the land, which are obtained by splitting the larger section of the land into subplots. This factor is therefore referred to as the subplot factor.

These same experimental situations are also common in industrial settings. Split-plot designs have three main characteristics:

1. The levels of all the factors are not randomly determined and reset for each experimental run. Did you hold a factor at a particular setting and then run all the combinations of the other factors?
2. The size of the experimental unit is not the same for all experimental factors. Did you apply one factor to a larger unit or group of units involving combinations of the other factors?
3. There is a restriction on the random assignment of the treatment combinations to the experimental units. Is there something that prohibits assigning the treatments to the units completely randomly?

The following industrial examples will help you recognize when it would be best to use a split-plot experiment.

Example A

Let's say you want to examine the image quality of a printing process by varying three factors:

- A = blanket type.
- B = cylinder gap.
- C = press speed.

Figure 1 illustrates a simple image of this part of a printing press.

You plan to study two different blanket types (1 and 2), three different cylinder gaps (low, medium and high) and two press speeds (low and high), and will run all 12 treatment combinations (see Figure 2) in the experiment. A completely randomized design would require you to run the 12 treatment combinations in a random order.

To change the cylinder gap and press speed, you simply make an adjustment on a control panel while the printing press is still running. Factors such as these are called easy-to-change factors. To change the blanket type, however, you must stop the press and manually replace the blanket. A change such as this is called a hard-to-change factor.

Now imagine the first three runs in your experiment are (A = 1, B = -1 and C = -1), (A = 2, B = 1 and C = -1) and (A = 1, B = 0 and C = 1). This means you would have to install the blanket three times (1 to 2, then back to 1), and using a completely randomized design would require you to frequently stop the press, thereby extending the time required to run the experiment.

A more time efficient approach, and one that fits

FIGURE 1 The Printing Press

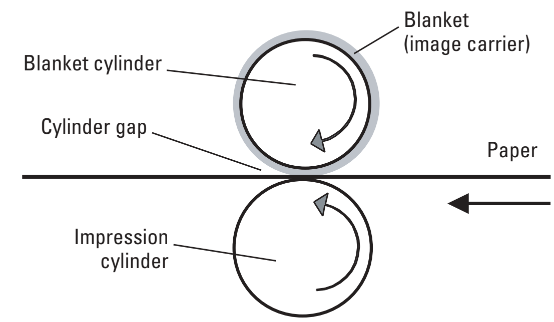
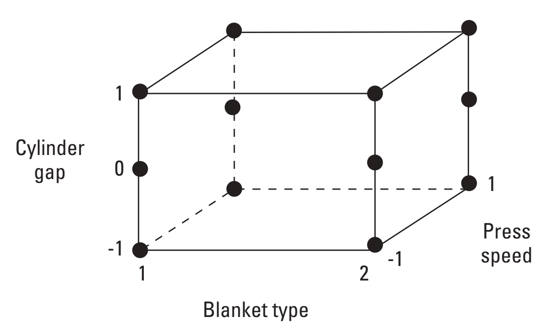


FIGURE 2 Factors That Affect Image Quality



into the split-plot framework, would be to randomly choose one of the blanket types (1 or 2), install it on the printing press, and run the six treatment combinations in cylinder gap and press speed in a random order. Then you would change the blanket type and run the six treatments in another random order, repeating the process until you reached the desired number of replicates for the blanket type factor.

The way in which the factor levels for blanket type are changed in the second approach involves a different randomization scheme from that of the factor levels for the other two factors. This different randomization structure is one feature of a split-plot design and is common when some of the factors are difficult to change. If the experiment were conducted in this manner, it would be incorrect to analyze the data as if you had run the experiment as a completely randomized design.

FIGURE 3 Factors That Affect Wood's Water Resistance

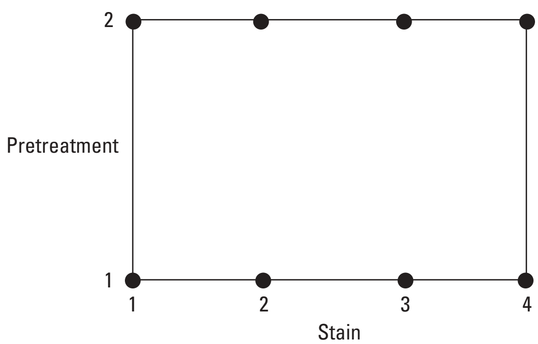
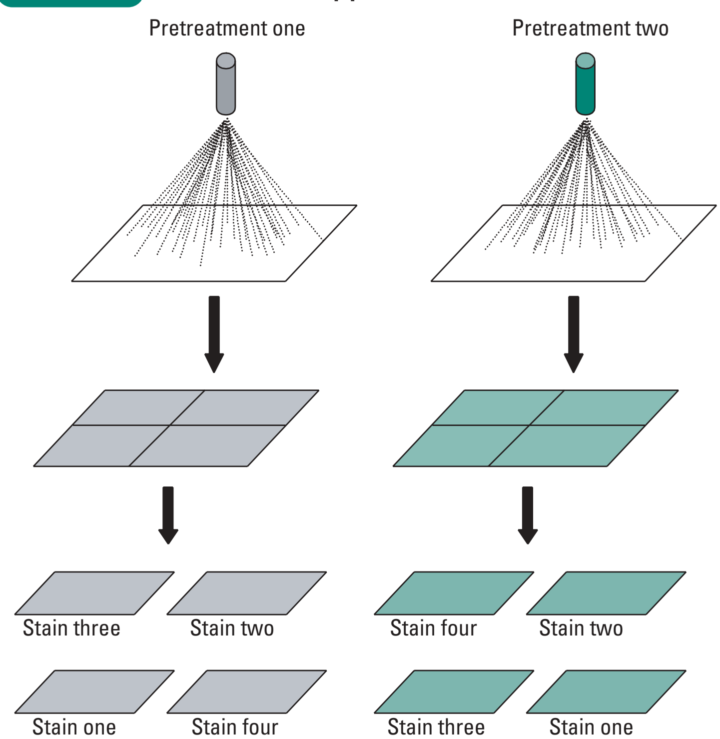


FIGURE 4 Treatment Application



Example B

Now let's look at an experiment involving the water resistance property of wood in which you select two types of wood pretreatment (1 and 2) and four types of stain (1, 2, 3 and 4) as variables (see Figure 3).

To conduct this experiment in a randomized fashion, you would need eight wood panels for each full replicate of the design. You would then randomly assign a particular pretreatment and stain combination to each wood panel.

That's when you discover how difficult it is to apply the pretreatment to a small wood panel. The easiest way to do it would be to apply each of the pretreatment types (1 and 2) to an entire board, then cut each board into four pieces and apply the four stain types to the smaller pieces (see Figure 4).

The experimental units for the two factors in this experiment are not the same. For the pretreatment factor, the experimental unit is the entire board, but for the stain factor, the experimental unit is one of the small panels cut from the large board. Varying sizes of experimental units is another feature of split-plot designs.

Example C

Let's say you want to examine the effect the four following

FIGURE 5 Factors That Affect the Strength of Plastic

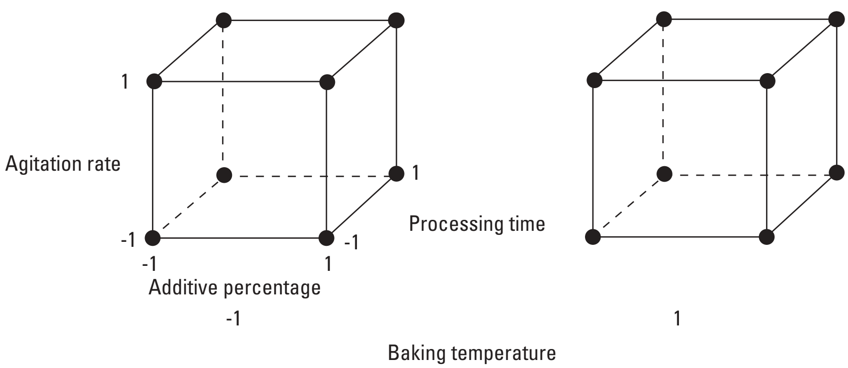
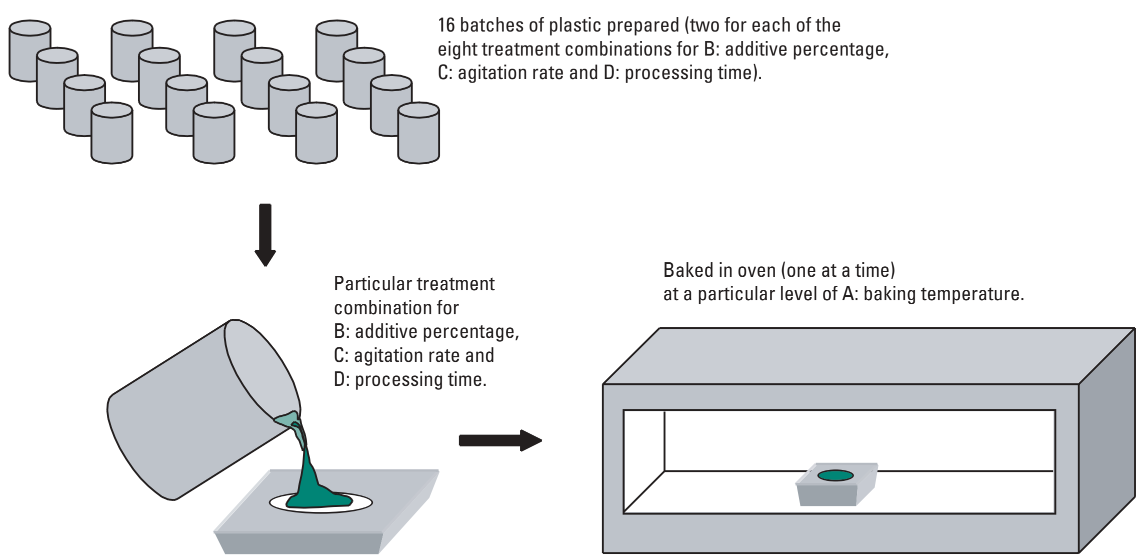


FIGURE 6 Individual Baking Process



various factors have on the strength of plastic:

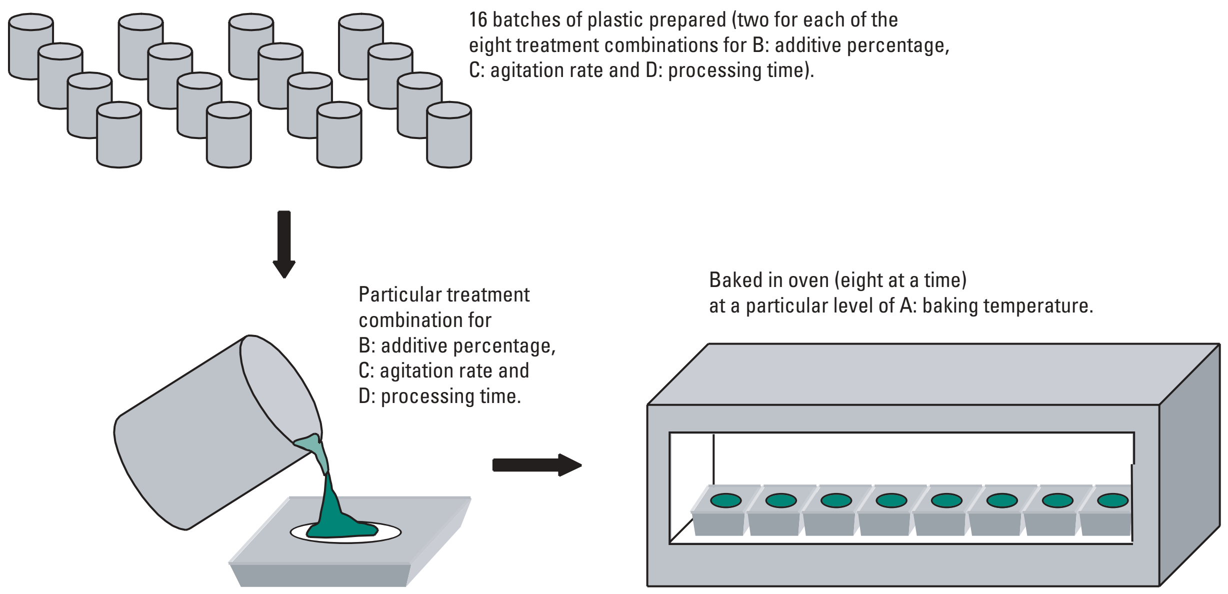
- A = baking temperature.
- B = additive percentage.
- C = agitation rate.
- D = processing time.

You plan to study each factor at two levels: low = -1 and high = 1. See Figure 5 for a graphical representation of this treatment design.

To conduct this experiment as a completely randomized design, you would run all 16 treatment combinations in a random order. After obtaining the required 16 batches of plastic, two for each of the eight different combinations of factors B, C and D, you would pour the plastic into molds and bake each individually at one of two temperatures (see Figure 6).

If you conducted the experiment in this way, you would have to frequently change the baking oven's temperature, which may take some time to stabilize. A completely randomized design also implies each run of the oven is a true experimental run. That means 16 separate runs of the oven are

FIGURE 7 Group Baking Process



needed, therefore adding considerable time to the experiment.

A more efficient approach would be to bake all eight molds for one temperature setting at the same time. You would follow this by a single run of the oven at the other temperature level, repeating the process until you have run the desired number of replicates for the temperature factor (see Figure 7).

Now you no longer have a completely randomized design but a split-plot design. Why? There are three reasons:

1. For each of the three factors—additive percentage, agitation rate and processing time—one experimental unit equals one batch of plastic, while for the temperature factor, one experimental unit equals all eight batches.
2. Temperature can be thought of as a hard-to-change factor, and the three easy-to-change factors are varied within a level of the hard-to-change factor.
3. The temperature factor uses a different randomization scheme from the other factors. The molds are assigned to the temperature factor in groups of eight as opposed to individually.

Split-Plot Design Affects Analysis

Many practitioners fail to see there is more to knowing the correct analysis than just being able to identify the treatment structure. The analysis of designed experiments directly follows from the way the runs were carried out.

For example, when a designed experiment uses blocks such as days or batches, the analysis of the experiment includes a term for these blocks. When a designed experiment is performed by fixing a factor and then running the combinations of the other factors, using different sized experimental units or using a different randomization for the factors (a split-plot design), the analysis should incorporate these features.

In example C, the complete 2^3 factorial treatment design was replicated twice using the split-plot approach. This resulted in the 32 response values shown in Table 1. The responses were first analyzed incorrectly as if they came from a completely randomized design. The responses were then correctly analyzed as a split-plot experiment. (Our intention is not to teach the analysis, but interested readers can look at Table 2 (p. 66)

for a summary of the two different analyses.)

The results of the incorrect analysis, a completely randomized design, indicate that, at the 0.05 significance level, the main effects for A (baking temperature) and D (processing time) are significant, as are the AC (baking temperature/agitation rate) and AD (baking temperature/processing time) interactions.

The results of the correct, split-plot analysis indicate the main effects for B (additive percentage) and D (processing time) are significant at the 0.05 level and A (baking temperature) is not. In addition to the AC (baking temperature/agitation rate) and AD (baking temperature/processing time) interactions, the CD (agitation rate/processing time) interaction is also significant.

Experimental Error

Two interesting results appear when the two analysis approaches are compared:

1. The effect of the baking temperature was thought to be significant when analyzed as a completely randomized design but was actually insignificant when analyzed correctly. The whole plot error of 56.29 is much larger than the error of 14.21 from the completely randomized design analysis. This would cause you to incorrectly assume baking temperature is an important effect.
2. Effects at the subplot level that were not significant when analyzed as a completely randomized design are seen as significant when analyzed correctly. The subplot error of 9.78 is smaller than the one that arises when a completely randomized design is incorrectly assumed. As a result, important effects at the subplot level that were missed in the incorrect analysis are now seen.

Why did this happen? In the completely randomized design, all factor effects use the mean square error as the estimate of experimental error. In a split-plot experiment, however, there are two different experimental error structures: one for the whole plot factor and one for the subplot factors. This is a result of the two separate randomizations that occur when the experiment is run.

Experimental error is caused when the actual experimental conditions are replicated. This could include the preparation and mixing of the plastic batches or the setup and temperature stabilization of the oven. For the baking temperature factor, there are only four experimental units—each set of eight molds placed together in the oven.

Even though each of these eight molds comes from a different treatment combination of the other three factors, they were all processed in a single run of the oven. They do not provide an estimate of experimental error for the whole plot factor. The experimental error for the whole plot factor comes

TABLE 1 The 32 Response Values

	Temperature	Additive	Rate	Time	Strength
	A	B	C	D	Y
1	1	1	-1	-1	51.9
2	1	1	-1	1	66.8
3	1	1	1	-1	66.2
4	1	1	1	1	70.8
5	1	-1	1	-1	61.3
6	1	-1	1	1	68.5
7	1	-1	-1	1	59.5
8	1	-1	-1	-1	58.5
9	-1	1	-1	-1	57.4
10	-1	1	-1	1	57.5
11	-1	-1	1	-1	56.5
12	-1	1	1	1	63.9
13	-1	-1	1	1	56.4
14	-1	1	1	-1	58.1
15	-1	-1	-1	1	53.2
16	-1	-1	-1	-1	59.5
	Temperature	Additive	Rate	Time	Strength
	A	B	C	D	Y
17	-1	-1	-1	-1	66.6
18	-1	-1	-1	1	63.9
19	-1	1	1	-1	62.6
20	-1	1	1	1	63.2
21	-1	-1	1	-1	56.1
22	-1	1	-1	1	63.3
23	-1	-1	1	1	62.7
24	-1	1	-1	-1	65.0
25	1	-1	-1	-1	59.5
26	1	-1	-1	1	64.2
27	1	-1	1	1	68.0
28	1	-1	1	-1	58.6
29	1	1	-1	-1	65.6
30	1	1	1	1	73.3
31	1	1	-1	1	61.5
32	1	1	1	-1	64.0

TABLE 2 Summary of Incorrect and Correct Analyses for Example

	Incorrect completely random		Correct split-plot	
Term	Significance	Variability	Significance	Variability
Temperature ▲	Significant	14.21	Not significant	56.29
Additive ▲	Not significant	14.21	Significant	9.78
Rate	Not significant	14.21	Not significant	9.78
Time	Significant	14.21	Significant	9.78
Temperature/additive	Not significant	14.21	Not significant	9.78
Temperature/rate	Significant	14.21	Significant	9.78
Temperature/time	Significant	14.21	Significant	9.78
Additive/rate	Not significant	14.21	Not significant	9.78
Additive/time	Not significant	14.21	Not significant	9.78
Rate/time ▲	Not significant	14.21	Significant	9.78
		One error term for all		Two error terms

▲ Shows terms that have a different interpretation between the two analyses.

from the variation experienced when the temperature is changed. This whole plot error is typically larger than the error from a completely randomized design.

In conducting a split-plot experiment, you need to be sure there is true replication in the whole plot factor. If each level of baking temperature was run only once and not replicated as it was here, there would be no estimate of whole plot experimental error and, therefore, no statistical test for this factor.

The challenges faced by practitioners result in completely randomized experiments being the exception, not the norm. Unfortunately, split-plot and other noncompletely randomized experimental designs have not received proper attention in most Black Belt statistical training courses because the mathematical concepts are usually more complicated or more general than those in the completely randomized design.

Fortunately, the availability of statistical software has slowly started to ease the analysis and interpretation of more complicated experimental structures, such as split-plot experiments.

Knowledge of the split-plot design gives practitioners another option with which to conduct experiments more efficiently.

BIBLIOGRAPHY

Box, G.E.P., "Split-Plot Experiments," *Quality Engineering*, Vol. 8, No. 3, pp. 515-520.

Box, G.E.P., and Stephen Jones, "Split-Plots for Robust Product and Process Experimentation," *Quality Engineering*, Vol. 13, No. 1, pp. 127-134.

Montgomery, D.C., *Design and Analysis of Experiments*, fifth edition, John Wiley & Sons, 2001.

SCOTT M. KOWALSKI *is a technical trainer at Minitab Inc. in State College, PA. He earned a doctorate in statistics from the University of Florida, Gainesville, and is a member of ASQ.*

KEVIN J. POTCNER *is a technical trainer at Minitab Inc. in State College, PA. He earned a master's degree in statistics from the Rochester Institute of Technology in New York.*

Please comment

If you would like to comment on this article, please post your remarks on the *Quality Progress* Discussion Board at www.asq.org, or e-mail them to editor@asq.org.