Paper 234-2009

Designing Experiments for the Process

George A. Milliken, Kansas State University and Milliken Associates, Inc
Manhattan, Kansas USA

ABSTRACT

Most experiments are a part of a process, not an entity unto themselves, and designs that do not account for the restrictions of the process end up with inferior analyses. In this presentation the design of experiments are described as a process instead of a single entity. A design must first be approached by stepping back to evaluate how the design of the experiment fits into the whole process, identifying the process restrictions, and, finally, using the restrictions to develop an appropriate designed experiment and then the corresponding analysis. Only when the complete process involved with the design of the experiment is known can an appropriate model be constructed. Simple to complex designs have basic characteristics in common called design structures. Basic design structures are the building blocks of complex designs. The identification of the four basic structures and their restrictions is the basis of successfully identifying or classifying an experiment and of constructing the resulting analysis. This approach is a change to the paradigm for designing experiments, and the methodology is applicable to all areas of research from agriculture to manufacturing to social sciences. The following sections present a discussion of the design of experiment tools and will present several examples from agriculture, semi-conductor manufacturing, and the social sciences.

INTRODUCTION

The traditional methods design of experiments are taught and/or discussed in text books are not the ways design of experiments are or should be used for real world applications. Design of experiments are taught as single entities such as a completely randomized design, randomized complete block, nested design, Latin square, etc.. In reality a study will generally consist of a series of individual steps that might require the use of a series of designs. A typical experiment consists of assigning animals to the levels of a drug and then measuring one or more responses on each of the animals. The analysis of the data obtained at the end of the process is generally based on the method of assigning the animals to the levels of the drug. But, there are possibly many steps between the assignments of animals to drugs and the data one extracts at the end of the study. Ignoring what possibly happens during these steps can result in inferior quality data analysis(data with a lot of unexplained variation). One or more tools of designed experiments might be used at each of the steps of the study. For example, if blood samples are obtained from each animal, what order does one obtain the samples from the animals and then what order does one have the samples analyzed by the laboratory? Will more than one technician be involved with obtaining the blood samples? Will more than one technician be involved with the analyses of the samples in the laboratory? How long will it take the samples to be analyzed in the laboratory? Will laboratory results vary from setup to setup or day to day or session to session? When discussing design of experiments, questions like these are generally not discussed. And most surely, the consequences of the answers to those (or similar) questions are ignored. The effect of ignoring the answers to the questions is to add variability to the data set. An increase in the variability in the data set dilutes the evaluation of the means of the levels of the drug through an increase in the estimated standard error of a difference and results in increases of type I error rates.

Each experiment is composed of a set of steps and the activities carried out at each step can have a big influence on the variability of the data set and unfortunately these causes often go unrecognized. It is important for the statistician, biostatistician or data analyst to understand exactly what is happening (or has happened) at each step of the process so that a decision can be made as to if one or more tools should be used to understand sources of variability in the data set and then identify and account for them by an appropriate analysis. Variability that can be accounted for by the analysis is variability that is removed from the error(s) of the model.

This presentation starts out with a detailed example of a simple experiment that involves several steps with discussion as to when it may be important to incorporate a tool such as randomization, or identify which samples were processed by which technician, etc. The second part of the discussion provides a description of a set of tools available for use at any step of the experiment. The final sections provide discussions of case studies showing the effects of ignoring the intermediate steps of the process.

It is important when one teaches design of experiments to inform the students that most designs consists of a set of

steps and that tools of design of experiments should be used at most of those steps to improve the quality of the data and resulting analyses. Examples like those discussed below should be used in the class room so the students can be aware of all of the opportunities that can occur between the assigning of treatments to experimental units and the data listed on the data sheet. Instructors need to get away from describing a design, its analysis, and the computations using software. There is much more to be learned when the experiment is considered as a process with a set of steps.

MOTOVATING EXAMPLE

An example of a simple experiment is a study of swine influenza virus. The experiment consists of evaluating two treatments, the control or placebo and the treatment as to their effect after the vaccinated pigs are challenged by a virus. There are 20 pigs of age 3-4 weeks old available for the study so the process starts by assigning unique numbers to each of the pigs, as shown in the first section of Figure 1. Next randomly assign 10 pigs to the two treatments as is demonstrated by the second section of Figure 1. For simplicity, randomly order the pig numbers and assign the first 10 pigs in the random list to the control and the second set of 10 to the treatment (or vis versa). At this point in the study, the design involves two treatments and a completely randomized design. The next step in this study is to apply the treatments to the pigs. This step must be blinded so the technician will not know which of the pigs receive the control and which receive the treatment. Each pig will be vaccinated where the control pigs will be vaccinated with saline solution and the treatment pigs will be vaccinated with the active treatment. The question is, "What order should be used to carry out the vaccination step of the study?" One could vaccinate all of the pigs assigned to the control and then vaccinate all of the pigs assigned the treatment. Could this possibly induce some bias into the experiment? If there is just a little possibility that part of the process could induce a little bias into the experiment, then use the randomization tool from the tool box. The process is to take 10 pre-drawn syringes of the saline solution and write on each syringe a number corresponding to a pig assigned to the control group and repeat the process using 10 syringes of active treatment for the treatment group. Complete this step by creating a random list of the numbers of the 20 pigs and vaccinate the pigs in that order, as shown in the third step of Figure 1.

The pigs will be challenged with a particular viral strain 14 days after vaccination where the challenge material is included in the feed as in the fourth step of Figure 1. All of the pigs are in the same pen and eat from the same feeder, so essentially all of the pigs are challenged at the same time. Nasal swabs and blood samples are to be obtained from each pig 21 days after the pigs have been challenged. Again, what order should one use to carry out the sample collection part of the study? A random order of the pigs should be used to obtain the samples as shown in the fifth box of Figure 1 where a set of swabs and tubes have the pig numbers attached so the researcher can keep track of which swab and tube contain the materials from each pig. Finally the swabs and blood samples need to be process to determine the degree of viral infection.

Set of 20 pigs with #s	→	Day 0, Randomize to treatments	→	Order of vaccination (T or C before denote Treatment or Control)	→	Challenge at Day 14	→	Order to Collect samples at Day 35
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20		Trt 3 5 6 7 9 12 14 15 16 19 Control 1 2 4 8 10 11 13 17 18 20		T15 C17 C11 T14 T19 C10 T7 C1 C8 C4 T12 C18 T16 C2 C20 T6 T5 C13 T3 T9				17 4 6 1 10 9 5 2 20 16 18 7 19 14 11 12 3 13 15 8

Figure 1. Process for assigning pigs to treatments and collecting samples

The swabs should be evaluated in a random order in the laboratory and the blood samples should be process in a random order in the laboratory. But there are some other constraints in the laboratory. The instrument used to evaluate the presence of viral titers from the blood samples can hold 10 tubes at a time, so it takes two instrument setup procedures to evaluate all 20 tubes. How does one decide which 10 tubes to include in the first setup (and the second setup). You could evaluate the 10 tubes from the control pigs in the first setup and the 10 tubes from the treated pigs in the second setup. That is not a good idea as that scheme would have confounded setup with treatment. So randomly select 5 tubes from control pigs and 5 tubes from treatment pigs and insert them into the instrument. There could be a positional effect of processing which can be somewhat accounted for by randomly assigning the 10 tubes to the 10 positions within the instrument. The samples and positions within the two setups of the instrument are displayed in Figure 2.

Samples and order for Setup 1	Samples and order for Setup 2
C12 T10 C5 T8 C6 C9 T11 T18 P19 T1	T13 C3 C7 T4 C16 T2 C14 T20 C15 T17
C# and T# denote pig number from control and treatment respectively	

Figure 2. Pigs and order within each of the setups for the blood analysis.

The second constraint is the viral titers from the swabs are evaluated by using what is called a 96 well plate where each plate consists of 8 rows and 12 columns. The rows are assigned a series of dilutions as 10^{-2} down to 10^{-9}. Three subsamples from four swabs are assigned to the columns of the plate with columns 1-3, 4-6, 7-9, and 10-12 containing the subsamples of swabs from four pigs respectively. The assignment of four pigs (two from each group) to the columns of a plate is shown in Figure 3, where T7 and T3 are two pigs from the treatment group and C17 and C11 are two pigs from the control group. Samples will be placed in each of the 96 wells (rectangles) of each plate.

	3 swabs from T7			3 swabs from C17			3 swabs from T3			3 swabs from C11		
10^{-2}												
10^{-3}												
10^{-4}												
10^{-5}												
10^{-6}												
10^{-7}												
10^{-8}												
10^{-9}												

Figure 3. The arrangement for plate 1 with subsamples of swabs assigned to three consecutive columns

Five plates are required to evaluate the titers from the swabs from the 20 pigs, so how should the swabs be assigned to the plates? In this case, I would randomly select swabs from two control pigs and two treated pigs for each plate. Within each plate I would randomly assign the four swabs to the four sets of three columns. Figure 4 contains the random assignment of pigs to columns of plates where two treatment pigs and two control pigs are evaluated on each plate.

Plate	Cols 1-3	Cols 4-6	Cols 7-9	Cols 10-12
1	T7	C17	T3	C11
2	C4	C13	T15	T6
3	C10	T12	C8	T19
4	T9	C18	C1	T16
5	C2	T14	T5	C20

Figure 4. The assignment of 4 pigs per plate along with the assigned columns

.

From the onset, this simple study seemed like an application of the completely randomized design, but the complete study consists of a process of several steps. At each step randomization was used to prevent possible biases from being introduced into the study. Fatigue is the most common source of bias in studies that require a repetitive process to apply treatments or collect samples or evaluate samples. For sure at no time should all of the control pigs be evaluated before evaluating any of the treated pigs. When evaluating the blood samples one must account for the limitations of the instrument. And when evaluating the swab samples one must account for the fact that the 96 well plates are being used and that only four swabs can be handled on each plate. A model generally used for the data from a study like this will be that of a completely randomized design with two treatments. The analysis will have a source of variation associated with the difference of the two treatments (signal) and a source of variation associated with the unexplained variability (noise) which is generally designated as ERROR or RESIDUAL. The variability between the two instrument runs to evaluate the titers from the blood samples will be included in the RESIDUAL for the blood titer analysis and the variability among the five plates to evaluate the titers from the swabs will be include in the RESIDUAL of the swab titer analysis. The five plates can be considered as blocks for the analysis of the data from the swabs and the two instrument setups can be considered as blocks for the blood titer analysis. The analysis of variances are in Tables 1 and 2 where the left hand side provides the analysis ignoring the blocking factors and the analysis on the right provides the analysis where the blocking factors are taken into account. The two runs of the instrument form a blocking factor for the blood titer analysis and the five plates form a blocking factor for the swab titer analysis. Failing to incorporate blocking factors into a model will provide an incorrect analysis and will usually provided comparisons of treatment effects where the estimated standard errors are too large, thus producing higher type I errors than desired or wider confidence intervals than are appropriate. From Table 1, σ^2_{1b} is the variance of the titer data from the blood without taking the setup process into account and σ^2_{2b} is the variance of the titer data from blood with taking setup into account. Generally, $\sigma^2_{2b} < \sigma^2_{1b}$ as most likely $\sigma^2_{setup} > 0$. An identical comparison can

be made for the titer data obtained from the swabs and 96 well plates. It is mighty important to be able to identify the extraneous blocking factors that just happen to occur because of the methods used to collect and evaluate the compounds in the samples. Most researchers are not going to volunteer this type of information, so it is up to the statistician, biostatistician, or data analyst to develop a very good understanding of the process that generates the data. This example demonstrates that within the same experiment there can be different blocking factors for different response variables. Tables 1 and 2 show the relationships between the variances of the responses with and without the blocking factors included in the analyses.

Table 1. Analysis of variance tables the titer data from the blood samples where $\sigma^2_{1b} = \sigma^2_{2b} + \frac{5}{9}\sigma^2_{setup}$

Without instrument setups as blocks			With Instrument setups as blocks		
Source	df	EMS	Source	df	EMS
Treatment	1	$\sigma^2_{1b} + \phi^2(trt)$	Treatment	1	$\sigma^2_{2b} + \phi^2(trt)$
Residual	18	σ^2_{1b}	Setup	1	$\sigma^2_{2b} + 10\sigma^2_{setup}$
			Residual	17	σ^2_{2b}

Table 2. Analysis of variance tables the titer data from the swab samples where $\sigma^2_{1s} = \sigma^2_{2s} + \frac{2}{9}\sigma^2_{plate}$

Without plates as blocks			With plates as blocks		
Source	df	EMS	Source	df	EMS
Treatment	1	$\sigma^2_{1s} + \phi^2(trt)$	Treatment	1	$\sigma^2_{2s} + \phi^2(trt)$
Residual	18	σ^2_{1s}	Plate	4	$\sigma^2_{2s} + 4\sigma^2_{plate}$
			Residual	14	σ^2_{2s}

A study or experiment is a process that consists of a series of steps as shown by the pig vaccine example. A new paradigm must be developed for the teaching of design of experiments where the initial design associated with the assigning treatments to the experimental units is not necessarily for the complete process as the complete experiment can involve many steps and principles of design of experiments may need to be used at many of these steps. The **new paradigm** should consist of teaching a set of basic tools that can be used at any step of the process of a study and then teach that even a simple study can consist of a set of steps (process) where one or more tools can be applied at each step. It is not sufficient to apply the tools to the steps of the process without connecting the steps together to provide a more appropriate model. Finally the methodology of the paradigm is complete when the tool or tools used at each step are combined into a model that appropriately describes the process used to generate a set of responses. Models or analyses that do not take into account all of the restrictions intrinsic to the process will provide a sub-optimal analysis that can possibly bestow very miss leading results. Knowing as much about the process as possible enables the statistician or biostatistician to become aware of the steps where randomization and/or a blocking structure should be incorporated so those factors can be included in a model that will more appropriately describe the data.

The approach to provide a basic change to the paradigm for designing experiments consists of three parts; (1) a set of tools for design of experiments, (2) an understanding of the process of the study to identify steps where tools can be applied (or should have been), and (3) knowledge of models that can be used to connect the structures constructed by using a tool at each step into an appropriate model. This presentation consists of describing a set of tools for the design of experiments and uses examples to demonstrate how to use the tools in a process to provide an appropriate design and analysis of a study. The tools are grouped into experimental unit concepts, design structures and treatment structures. There are four basic design structures that are demonstrated by using a two-way treatment structure with additional emphasis on the computation of the respective error terms. The basic design structures are the building blocks of complex designs as most complex designs are a combination of the basic design structures applied in different ways or at different steps of the process. Or to describe it another way, a complex design can be broken down into a collection of basic design structures that will enable a more appropriate model to be constructed. Identification of appropriate design structures that take into account the restrictions of the process is the basis of a suitable design for the experiment and resulting analysis. After the tools are presented, several examples are used to demonstrate the methodology. Design and treatment structures are discussed in much more detail along with the analyses in Milliken and Johnson (1989, 2001, 2009) and Milliken, et.al. (1998). Additional information on the analyses can be found in Littell, et.at. (1996, 2006).

TOOLS FOR DESIGN OF EXPERIMENTS

The design of experiments tool box contains several useful tools or groups of tools. The experimental unit tools are experimental unit, replication, and randomization. The other groups of tools discussed are treatment structure and design structure (Milliken and Johnson (2009)).

The experimental unit is the entity to which a treatment can be applied, but as will be demonstrated a study can involve more than one size of experimental unit.

The replication of a treatment or treatment combination is the entity to which treatments can be applied and observed independently of the other entities in the study or at that level of the study. Sometimes the experimental units and replications are identical, but as will be demonstrated that is not always the case.

Randomization is the process of assigning the levels of a factor to experimental units. Randomization is the insurance policy for preventing bias to occur. Biases can occur in many ways, such as subconsciously assigning the best pigs to the treatment and the worst pigs to the control. Randomization is also a good way to establish the order experimental units or parts of experimental units are measured during process through the laboratory, etc. As a general rule, if there is any doubt order can induce bias, then use randomization..

Design structures consist of methods of blocking the experimental units so that an effective experiment can be conducted. Examples of designs structures are completely randomized, randomized complete block, incomplete block, split-plot, strip-plot, etc.

Treatment structures consist of the combinations of levels of factors that are of interest to the researcher and are described in the protocol. Examples of design structures are one-way, two-way, n-way, Latin square, D-optimal, nested, a two-way with controls, etc.

A major assumption is that the factors describing **the blocking in the design structure must not interact with the factors defining the treatments in the treatment structure**. Most descriptions of blocking factors indicate blocking factors are those factors that are not of interest to the researcher but possibly help describe or account for some of the variability in the system. Describing variability in the system is not the only reason to select blocking factors, but those blocking factors in the design structure MUST NOT INTERACT with factors in the treatment structure. If there are some factors in the design structure that you think will help describe part of the variability in the system, but some of those factors will possibly interact with the factors in the treatment structure, then those possibly interacting factors must be moved from the design structure to the treatment structure so the appropriate interactions may be evaluated.

The total design of the experiment is determined by describing the treatment structure, describing the design structure, and describing the method of randomly assigning treatments from the treatment structure to the experimental units in the design structure as indicated by the schematic in Figure 5.

The design and treatment structures can be identical for two studies, but the resulting designs can be very different because of the method of randomization used to assign treatments from the treatment structure to the experimental units in the design structure. Different designs with the same design and treatment structures are described next and then the use of the set of tools is demonstrated by using a two-way treatment structure with sixteen treatments in different design structures all with thirty-two experimental units.

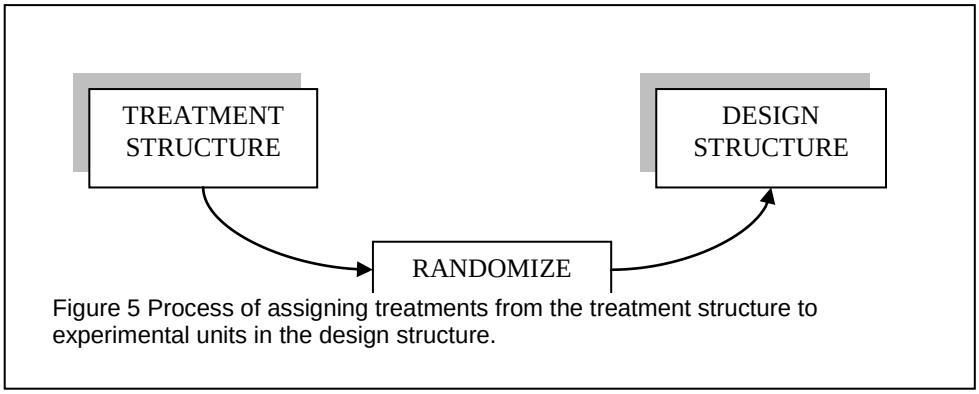

Figure 5 Process of assigning treatments from the treatment structure to experimental units in the design structure.

SIMPLE EXAMPLES WITH SAME DESIGN AND TREATMENT STRUCTURES

Consider an experiment with four treatments in the treatment structure and six blocks of size two as shown in Figure 6. The experiment is not complete until one specifies how to assign the treatments in the treatment structure to the experimental units in the design structure. The resulting design will be some sort of incomplete block as there are more treatments in the treatment structure than there are experimental units in the blocks of the design structure.

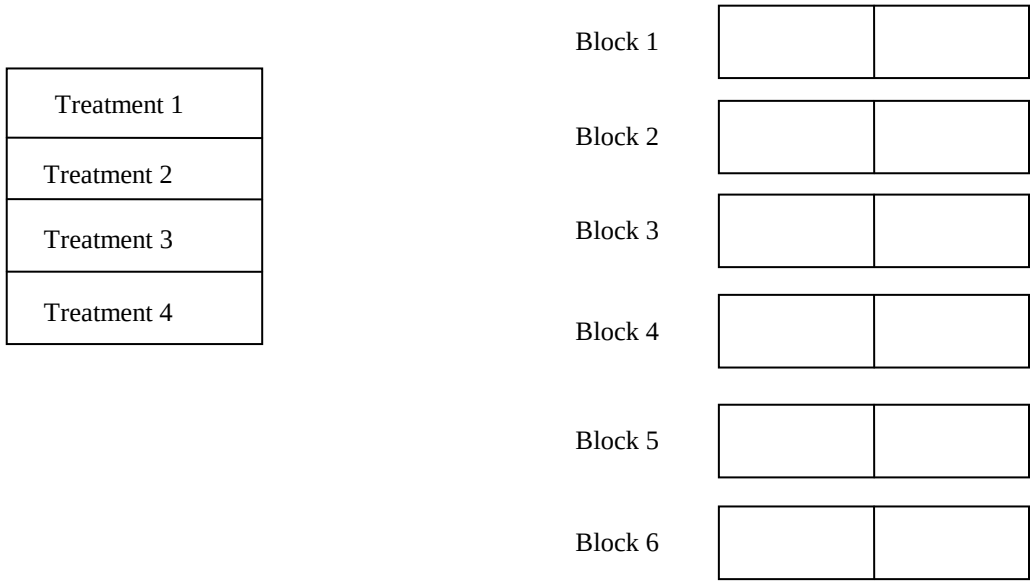

Figure 6 One-way Treatment Structure and 6 Blocks of size 2 Design Structure

There are three designs in Figure 7. The first one is constructed by assigning each pair of treatments to one of the blocks and then randomly assigns the two treatments to the two experimental units within the block**.** The design provides three replications of each of the treatments.

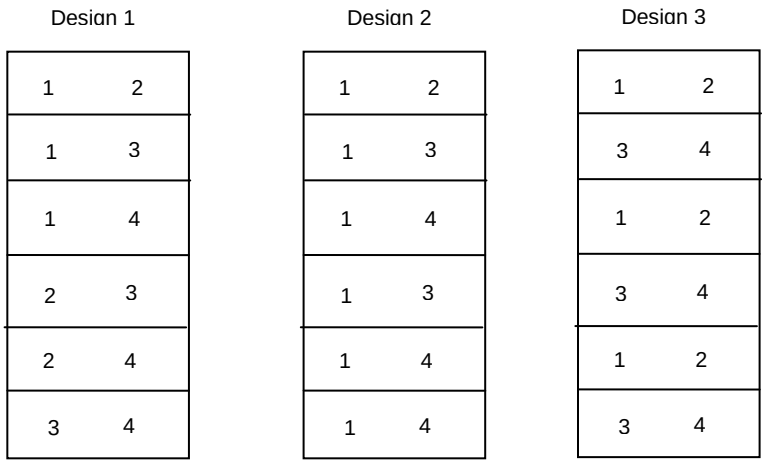

Figure 7 Three designs of experiments from identical treatment structures and identical design structures.

The second design is constructed by assigning treatment 1 to each block and treatments 2, 3 and 4 each to 2 blocks. There are six replications of treatment 1 and two replications of treatments 2, 3, and 4. Design 1 is optimal for comparing the equality of all four treatment means and design 2 is optimal for comparing the mean treatment 1 to the mean of treatment 2, to the mean of treatment 3 and to the mean of treatment 4. Design 3 consists of three blocks containing treatments 1 and 2 and three blocks containing treatments 3 and 4. The comparisons between the means

of treatments 1 and 2 and between the means of treatments 3 and 4 have smaller variances than the comparisons of any other pairs of means. The designs in Figure 7 provide the assignments of two treatments to each block. To finish the design randomly assign the two treatments to the two experimental units within a block.

Baking Bread Examples

The process of baking bread provides a series of examples to demonstrate the four basic design structures and two to four way treatment structures. The starting point is that there are four receipts for bread to be baked at four temperatures as shown in Figure 8.

TREATMENT STRUCTURE
4 RECIPES BY
4 TEMPERATURES

	Recipe 1	Recipe 2	Recipe 3	Recipe 4
Temperature 1	T1R1	T1R2	T1R3	T1R4
Temperature 2	T2R1	T2R2	T2R3	T2R4
Temperature 3	T3R1	T3R2	T3R3	T3R4
Temperature 4	T4R1	T4R2	T4R3	T4R4

Figure 8 Treatment structure for the bread baking examples

The baker wishes to observe two replications of each of the 16 treatment combinations, requiring 32 loaves of bread. The 32 loaves of bread are displayed in Figure 9. The process is to make up a batch of bread dough using one of the recipes and then baking a loaf of bread in an oven set to a specified temperature. The key to looking at different designs is to describe the order the loaves of bread are baked after the levels of temperature and recipe have been assigned as the laboratory only has one oven available to use for this study. The collection of the experimental units used to form the design structure and the two-way treatment structure are displayed in Figure 10. The following discussion demonstrates four different designs by describing how to assign the elements of the treatment structure to the loaves of bread.

Two Replications of Each of the 16 Treatment Combinations or need 16 loaves of bread

Experimental Units or Loaves 1-16

Experimental Units or loaves 17-32

Figure 9 32 loaves of bread to be baked

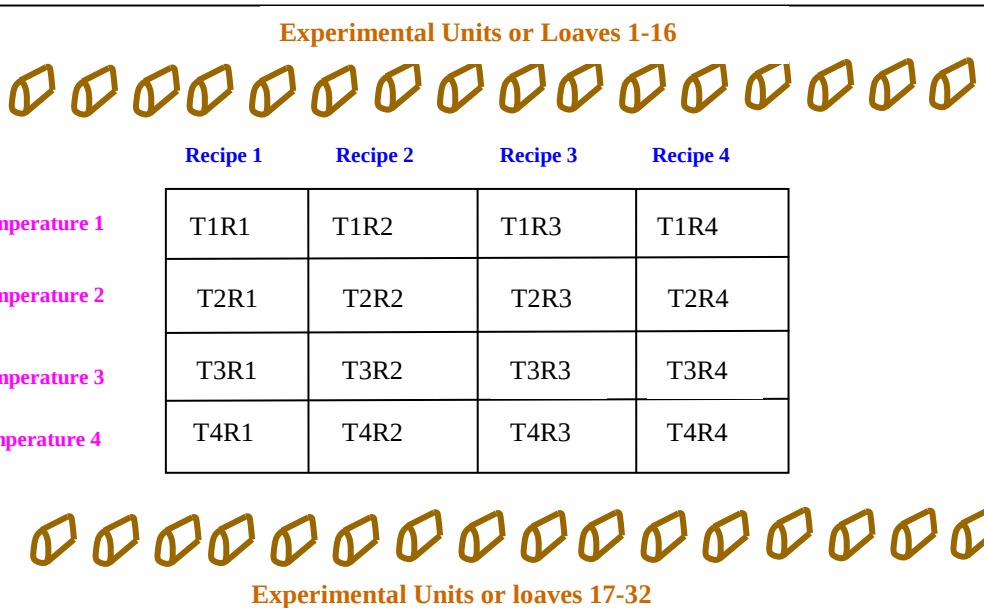

Figure 10 Design and Treatment Structures for the bread baking example.

COMPLETELY RANDOMIZED DESIGN STRUCTURE

The completely randomized design is constructed by completely at random assigning the 16 treatment combinations so that each treatment combination is observed twice. The loaves are numbered 1-32 indicating the order a loaf will be baked after it is assigned a recipe and temperature. Figure 11 is a graphic displaying the random assignment of each of the 16 treatment combinations to the 32 loaves of bread. For example the first loaf to be baked is with recipe 2 at temperature 3 (T3R2) and the last loaf to be baked is with recipe 1 at temperature 3. The process is to randomly order the combinations of recipes and temperatures. Using the random order, make a batch of dough with the specified recipe, form a loaf and put it into a container, and finally bake the loaf using the specified temperature. For discussion, the volume of the loaf is used as the response variable. The process requires that 32 batches of dough be made and each loaf is baked independently in an oven at a given temperature. That is, the completely randomized design requires 32 batches of dough and 32 bakes or uses of an oven.
A model that can be used to describe the loaf volume data is

$$V_{ijk} = \mu_{ij} + \varepsilon_{ijk} \; i = 1,2,3,4, \; j = 1,2,3,4, \text{ and } k = 1,2$$

where one distributional assumption is $\varepsilon_{ijk} \sim IID\ N(0, \sigma^2_{\text{residual}})$. The variance $\sigma^2_{\text{residual}}$ is the process variance which includes batch to batch, oven to oven, loaf to loaf and possibly day to day variances. The basic analysis of variance table is in Table 3 where there are 15 degrees of freedom for the treatment structure and there are 16 degrees of freedom for the design structure. The 16 degrees of freedom arise by computing the variation of the two replications of each of the treatment combinations and then pooling those degrees of freedom across the treatment combinations. In general, error terms are computed from variation of experimental units treated alike and pooled across the entities within which the experimental units were treated alike.

RANDOMIZED COMPLETE BLOCK DESIGN STRUCTURE

The randomized complete block design structure is constructed by randomly assigning each of the 16 treatment combinations to loaves 1-16 (dashed lines) and then randomly assigning each of the 16 treatment combinations to loaves 17-32 (solid lines) as shown in Figure 11. Thus, loaves 1-16 form block 1 and loaves 17-32 form block 2. One reason to form these blocks is that maybe only 16 loaves of bread can be baked during one day, thus requiring two days of baking so the loaves baked on day 1 form block1 and the loaves baked on day 2 form block 2. Most importantly, one wants to make sure there should not be an interaction between the levels of the treatment structure and the levels of the factors used to form blocks. A model that can be used to describe the loaf volume data in this blocked design is

$V_{ijk} = \mu_{ij} + b_k + \varepsilon_{ijk}$ $i = 1,2,3,4,$ $j = 1,2,3,4,$ and $k = 1,2$

where one set of distributional assumptions are $\varepsilon_{ijk} \sim IID\ N(0, \sigma^2_{\text{residual\_1}})$ $b_k \sim IID\ N(0, \sigma^2_{\text{block}})$. The variance $\sigma^2_{\text{residual\_1}}$ is the process variance which includes batch to batch, oven to oven, and loaf to loaf variances. The variance σ^2_{block} is the block to block or day to day variance which has been removed from $\sigma^2_{\text{residual}}$ which means $\sigma^2_{\text{residual\_1}} < \sigma^2_{\text{residual}}$ Table 4 contains the analysis of variance table for the two-way treatment structure in a randomized complete block design structure. There are still 15 degrees of freedom associated with the treatment structure, but the 16 degrees of freedom associated with the design structure are assigned as 1 to Blocks and 15 to Residual_1. The Residual_1 sum of squares is computed as the treatment (16 treatment combinations) by block (two blocks) interaction.

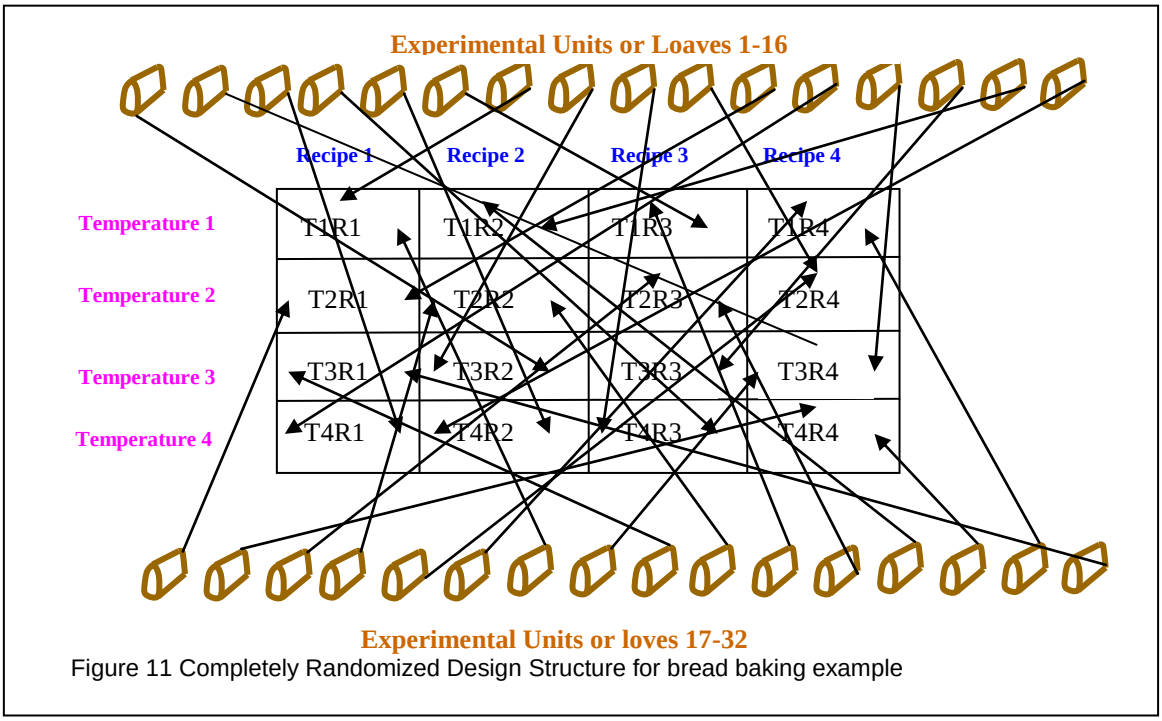

Figure 11 Completely Randomized Design Structure for bread baking example

Table 3. Analysis of variance table for the completely randomized design for the bread baking example.

Source	df	Expected Mean Squares
Temperature	3	$\sigma^2_{\text{residual}} + \phi^2(T)$
Recipe	3	$\sigma^2_{\text{residual}} + \phi^2(R)$
Temperature*Recipe	9	$\sigma^2_{\text{residual}} + \phi^2(TR)$
Residual	16	$\sigma^2_{\text{residual}}$

The process requires that 16 batches of dough be made where the loaves independently each day. That is, the randomized complete block design requires 32 batches of dough and 32 bakes or uses of an oven.

SPLIT-PLOT DESIGN STRUCTURE

The ovens being used for the baking study are large enough that more than one loaf could be baked at a time. In this case assume that four loaves of bread could be baked within an oven at the same time. If four loaves are baked within the same oven at the same time, the resulting responses become correlated. This correlation can occur

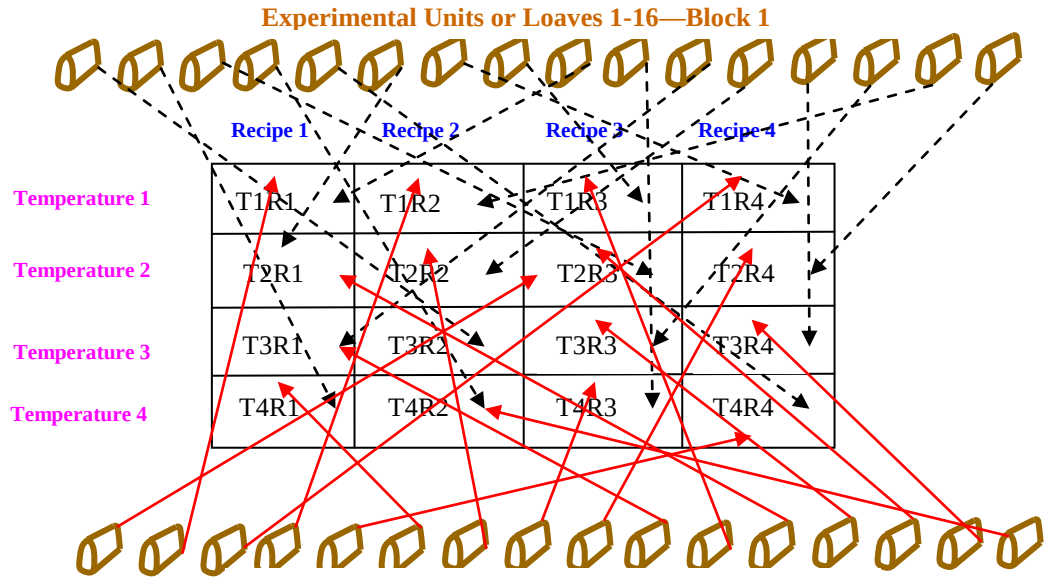

Figure 12 Randomized Complete Block Design Structure for the Bread Baking Example.

Table 4. Analysis of variance table for the randomized complete block design for the bread baking example.

Source	df	Expected Mean Squares
Blocks	1	$\sigma^2_{\text{residual\_1}} + 2\sigma^2_{block}$
Temperatures	3	$\sigma^2_{\text{residual\_1}} + \phi^2(T)$
Recipes	3	$\sigma^2_{\text{residual\_1}} + \phi^2(R)$
Temperature*Recipe	9	$\sigma^2_{\text{residual\_1}} + \phi^2(TR)$
Residual	15	$\sigma^2_{\text{residual\_1}}$

for several reasons as a) the temperature was not set exactly to the specified temperature, b) the amount of time the loaves were to be baked were all the same, etc. So how does one carry out such a study where the data are useful?

The process is to take the first four loaves and put them into the first oven where each loaf is made from one of the four recipes. In this case the oven becomes a block of size four which contains 4 treatment combinations where all 4 are assigned the same temperature. Figures 13 and 14 demonstrate the random assignment of the treatment combinations to the loaves. In Figure 13 the levels of temperature are assigned to the rows which form an oven. Within each oven one loaf from each recipe is inserted where the loaf is from its own batch of dough. The process generates an incomplete block design within each day which is similar to design 3 in Figure 7, which also generates two sizes of experimental units; the oven to which the levels of temperature are assigned and the loaf to which the levels of recipe are applied. Each of the experimental units has an associated variance component, so there is a variance component due to batches and a variance component due to ovens. Only part of the lines are presented in Figures 13 and 14 as there should be a set of lines assigning the levels of recipe to temperature 2, a set of lines assigning the levels of recipe to temperature 3 and a set of lines assigning the levels of recipe to temperature 4. Within each of the ovens, each loaf is obtained from its own batch of dough.

This split-plot design involves making 32 batches of dough (one for each loaf) and baking them in 8 ovens or 8 bakes, so this design structure takes one-fourth the numbers of bakes as the completely randomized and randomized complete block design structures. A model that can be used to describe the loaf volume data in this blocked design is

$$V_{ijk} = \mu_{ij} + b_k + o_{ik} + \varepsilon_{ijk} \; i = 1,2,3,4, \; j = 1,2,3,4, \text{ and } k = 1,2$$

or

$$V_{ijk} = \mu + T_i + b_k + o_{ik} + R_j + (TR)_{ij} + \varepsilon_{ijk}$$

where one set of distributional assumptions are $\varepsilon_{ijk} \sim IID\ N(0, \sigma^2_{\text{residual\_2}})$ $b_k \sim IID\ N(0, \sigma^2_{\text{block}})$ and $o_{ik} \sim IID\ N(0, \sigma^2_{\text{oven}})$ The variance $\sigma^2_{\text{residual\_2}}$ is the process variance which includes batch to batch and loaf to loaf variances. The variance σ^2_{block} is the block to block or day to day variance which has been removed from $\sigma^2_{\text{residual}}$ and σ^2_{oven} denotes the oven to oven variation.

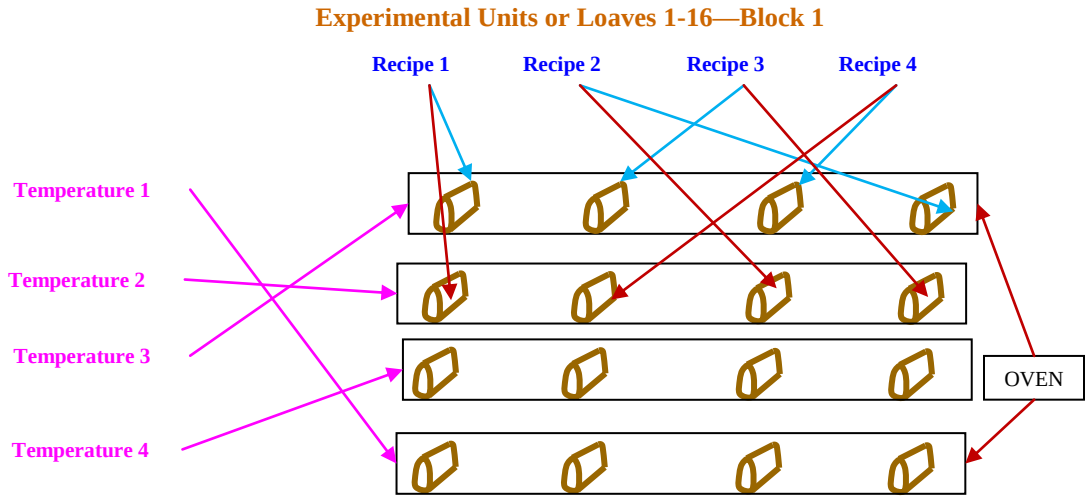

Figure 13 Block 1 of the split-plot design structure for the bread baking example where recipes are re-randomized within each temperature or oven.

Table 5 contains the analysis of variance table for the two-way treatment structure in a split-plot design structure. There are still 15 degrees of freedom associated with the treatment structure, but the 16 degrees of freedom associated with the design structure are assigned as 1 to Blocks, 3 to ovens nested within the levels of temperature and 12 to Residual_2. The Ovens(Temperature) sum of squares is computed by the block by temperature interaction. The Residual_2 sum of squares is computed as the recipe by block interaction within a level of temperature pooled across the levels of temperature.

One important feature of the split-plot design is that it is the basis of the repeated measures design. The repeated measures design is constructed identically to that of a split-plot design, except the levels of the sub-plot factor cannot be randomized to the experimental units. Time is a really good example of a repeated measurement where the measurements at time 1 must occur before those at time 2, etc. Another example is position where the researcher is interested in what happens on the top, middle and bottom shelf of a freezer. You cannot randomize those positions, thus they are called repeated measurements. The topic of repeated measurements will not be discussed in detail, but there are numerous examples in Milliken and Johnson, 2009 and 2002 and Little,et.al. 2006. The main difference between the analyses of a split-plot and a repeated measure is the modeling of the covariance among of the repeated measurements which is not much of a concern for the split-plot.

STRIP-PLOT DESIGN STRUCTURE

The baker had another problem with the above designs in that she was required to make up a batch of dough from which only one loaf could be extracted. It does not take any more time to make a batch of dough that is large enough so more than one loaf could be extracted. The process was then modified so that each batch of dough was made to be large enough that four loaves of bread could be formed and put into pans.

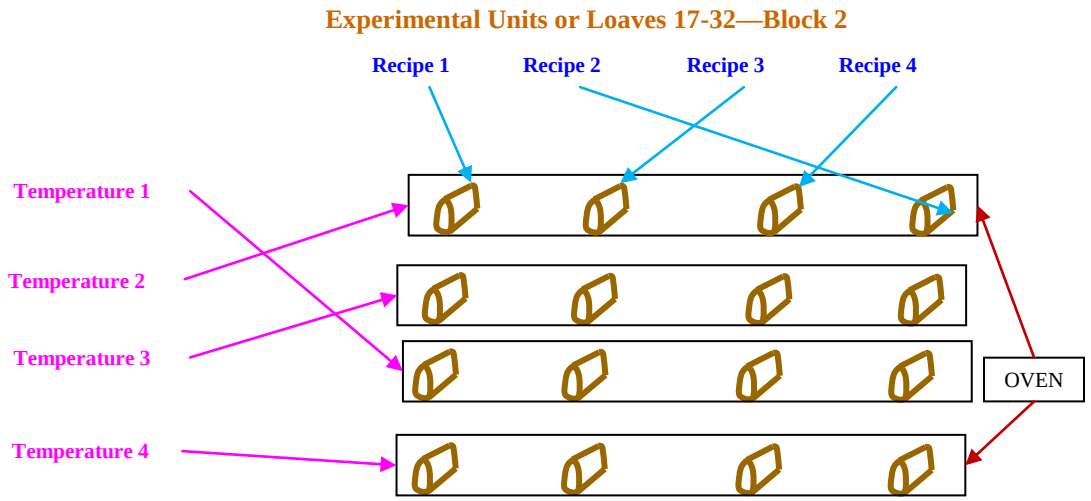

Figure 14 Block 2 of the split-plot design structure for the bread baking example where recipes are re-randomized within each temperature or oven.

Table 5. Analysis of variance table for the split-plot design structure for the bread baking example.

Source	df	Expected Mean Squares
Blocks	1	$\sigma^2_{\text{residual\_2}} + 4\sigma^2_{oven} + 2\sigma^2_{block}$
Temperatures	3	$\sigma^2_{\text{residual\_2}} + 4\sigma^2_{oven} + \phi^2(T)$
Ovens(Temperature)	3	$\sigma^2_{\text{residual\_2}} + 4\sigma^2_{oven}$
Recipes	3	$\sigma^2_{\text{residual\_2}} + \phi^2(R)$
Temperature*Recipe	9	$\sigma^2_{\text{residual\_2}} + \phi^2(TR)$
Residual	12	$\sigma^2_{\text{residual\_2}}$

The four loaves of bread are not independent as if she makes a little error in the ingredients, then that error is experienced by all four loaves. This process also generates another size of experimental unit which is called the batch. Figures 15 and 16 demonstrate the process of assigning the levels of temperature to the ovens (denoted by the rows) within each block and for assigning the levels of recipe to the batches (denoted by the columns). Within each day 16 loaves of bread are arranged into a 4 by 4 rectangle forming rows and columns. The levels of temperature are randomly assigned to the rows of the rectangle and the levels of recipe are randomly assigned to the columns of the rectangle.

The oven is the experimental unit for the levels of temperature since the levels of temperature are randomly assigned to the ovens (rows) within a block. The batch of dough (column) is the experimental unit for the levels of recipe since the levels of recipe are randomly assigned to the batches or columns of the rectangle. An interaction among the levels of recipe and the levels of temperature are computed as comparisons within ovens and within batches. Thus the experimental unit for a temperature by recipe interaction comparison is the loaf of bread. The strip-plot design structure involves three different experimental units, the oven (row), the batch (column) and the loaf (cell) and each experimental unit has an associated error term and variance component. This strip-plot design involves making 8 batches of dough (one for each column) and 8 baking of the loaves(one for each row), so this design structure takes one-fourth the number of bakes and one-fourth the number of batches as the completely randomized and randomized complete block design structures.
A model that can be used to describe the loaf volume data in this blocked design is

$$V_{ijk} = \mu_{ij} + b_k + ba_{jk} + o_{ik} + \varepsilon_{ijk} \; i = 1,2,3,4, \; j = 1,2,3,4, \text{ and } k = 1,2$$

or

$$V_{ijk} = \mu + b_k + T_i + o_{ik} + R_j + ba_{jk} + (TR)_{ij} + \varepsilon_{ijk}$$

where one set of distributional assumptions are $\varepsilon_{ijk} \sim IID\ N(0, \sigma^2_{\text{loaf}})$, $b_k \sim IID\ N(0, \sigma^2_{\text{block}})$ $ba_{jk} \sim IID\ N(0, \sigma^2_{batch})$ and $o_{ik} \sim IID\ N(0, \sigma^2_{\text{oven}})$ The variance σ^2_{loaf} is that part of the process variance that is left over after the variation due to blocks (day), ovens and batches have been accounted for by the analysis. The variance σ^2_{block} is the block to block or day to day variance, σ^2_{loaf} denotes the loaf to loaf variation, σ^2_{oven} denotes the

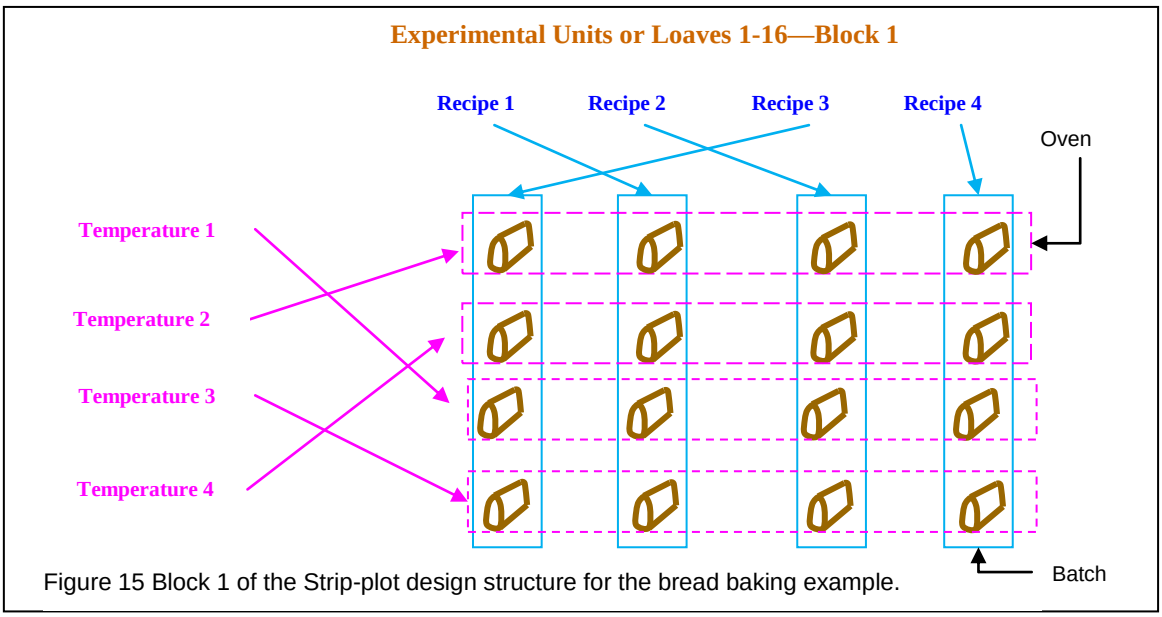

Figure 15 Block 1 of the Strip-plot design structure for the bread baking example.

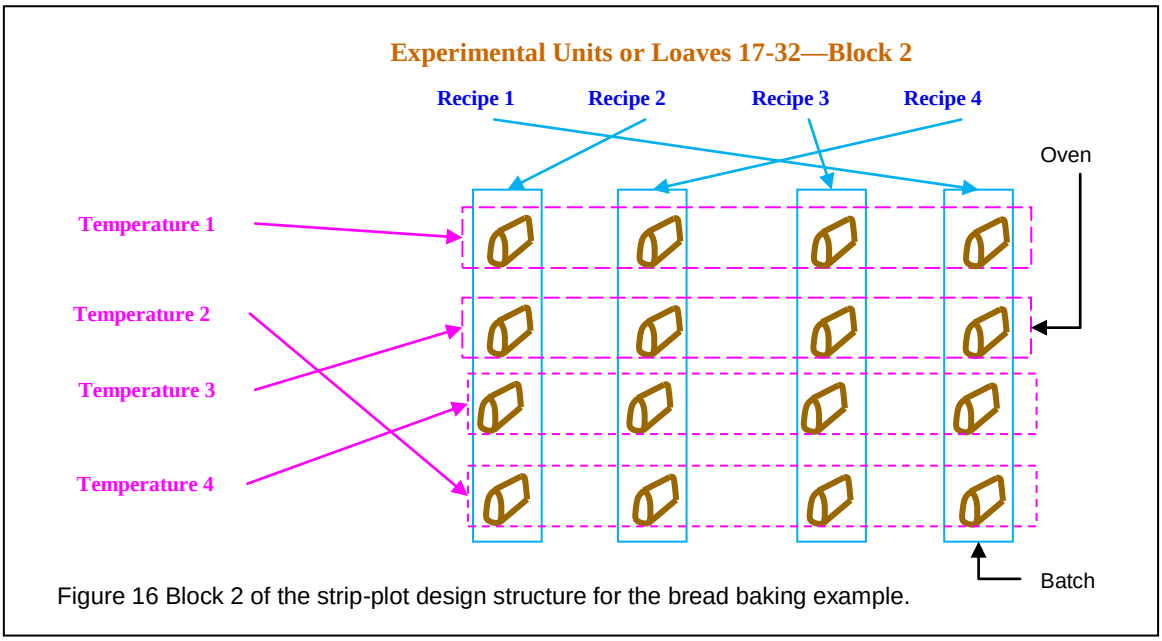

Figure 16 Block 2 of the strip-plot design structure for the bread baking example.

oven to oven variance, and σ^2_{batch} denotes the batch to batch variance. The total variance in the system or process is $\sigma^2_{loaf} + \sigma^2_{oven} + \sigma^2_{batch} + \sigma^2_{block}$.

Most analyses involve making comparisons among the temperature, recipe and temperature by recipe means. There are many methods for doing those comparisons, but they are not discussed here. The new edition of the book by Milliken and Johnson (2009) provides a great discussion of the multiple comparison and multiple testing methods.

Table 6. Analysis of variance table for the split-plot design structure for the bread baking example.

Source	df	Expected Mean Squares
Blocks	1	$\sigma^2_{loaf} + 4\sigma^2_{oven} + 2\sigma^2_{block}$
Temperatures	3	$\sigma^2_{loaf} + 4\sigma^2_{oven} + \phi^2(T)$
Ovens(Temperature)	3	$\sigma^2_{loaf} + 4\sigma^2_{oven}$
Recipes	3	$\sigma^2_{loaf} + 4\sigma^2_{batch} + \phi^2(R)$
Batches(Recipes)	3	$\sigma^2_{loaf} + 4\sigma^2_{batch}$
Temperature*Recipe	9	$\sigma^2_{loaf} + \phi^2(TR)$
Residual	9	σ^2_{loaf}

Table 7 contains a comparison of the error terms associated with each of the designs and the sources of variability accounted for by the term. As the design becomes more complex, going from the CR or RCB to the strip-plot design structures the degrees of freedom associated with the design structure are partitioned into more and more terms, but the terms become targeted an separate. It is easily seen by observing the sources of variation associated with Residual, Residual_1, Residual_2 and Residual_3. There are fewer degrees of freedom associated with the residual terms as the design becomes more complex, but the possible reduction in the magnitude of the residual variance very often more than accounts for the reduction in degrees of freedom when constructing confidence intervals and carrying out tests of hypotheses. The other benefit of the complex design structure is the reduction in resources required to carry out the experiments, the strip-plot design structure is the most efficient of the four design structures.

Table 7 Sources of variability attributed to each of the error terms for the designs associated with the bread baking designs.

Design Structure	Error Term	Source of variance
Completely Randomized	Residual	Day, batch, oven, loaf
Randomized Complete Block	Residual_1	Batch, oven, loaf
Slit-plot RCB whole-plot Design Structure	Block	Block or Day
	Oven(Temperature)	Oven
	Residual_2	Batch, loaf
Strip-plot Design Structure	Block	Block or Day
	Oven(Temperature)	Oven
	Batch(Recipe)	Batch
	Error(loaf)	loaf

FOUR-WAY TREATMENT STRUCTURE

One thing that needs to be stressed while looking through the tools in your tool box is that most any treatment structure can be fit with most any design structure. To demonstrate, assume the four recipes are constructed from combinations of two types of flour (whole wheat and white) with two types of yeast (fresh and stored) and the temperatures are constructed from combinations of two temperatures (335 and 365 degrees F) and two types of ovens (conventional and confection). This treatment structure consists of four factors each at two levels. Figure 17 is a graphical display of the treatment structure where FLOUR has two levels, YEAST has two levels, TEMPERATURE has two levels and STOVE type has two levels. This four-way treatment structure can be applied in any of the above design structures. Table 8 contains the analysis of variance table for applying this four-way treatment structure in a strip-plot design structure where the four combinations of TEMP by STOVE are assigned to the rows and the four

combinations of FLOUR by YEAST are assigned to the columns as in Figures 15 and 16. The analysis of variance table is in Table 8 where there are 15 degrees of freedom associated with the treatment structure (4 factors each with two levels providing 15 main effects and interactions) and 16 degrees of freedom associated with the design structure assigned to blocking, error for ovens, error for batches and error for loaves.

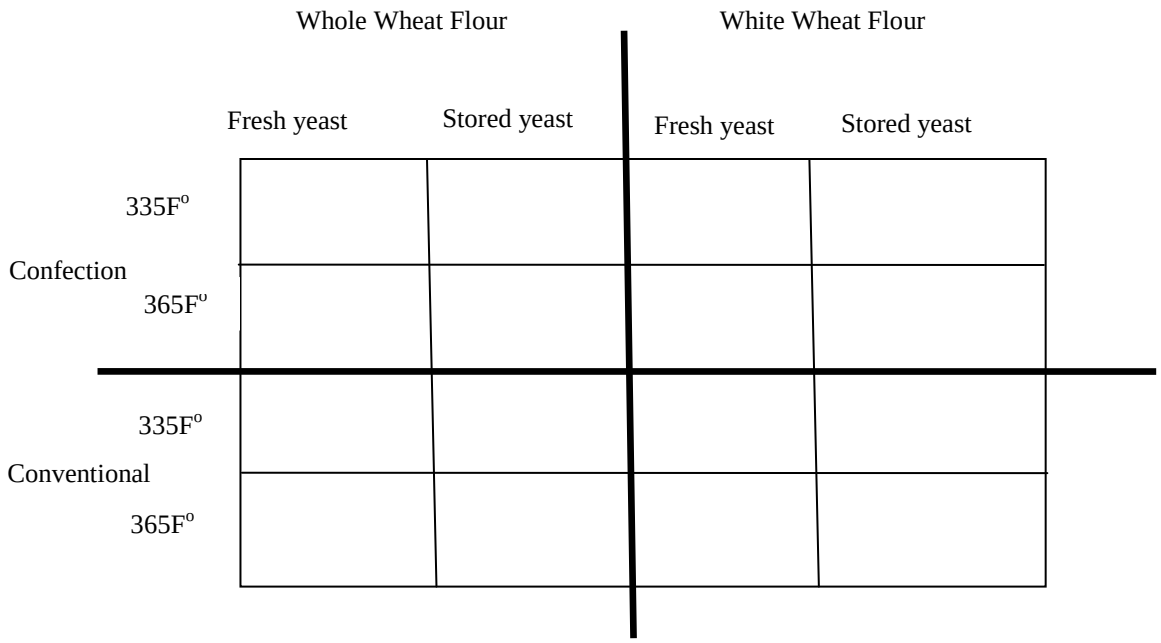

Figure 17 Temperatures by recipe represented as a four-way treatment structure

The treatment structure has a more complex composition than that for the initial bread making examples with just Temperature and Recipe. Now with the four-way treatment structure, the analysis of variance table is greatly expanded as each entry has only one associated degree of freedom. On the other hand the design structure stays the same as is displayed in Table 8. A very important tool in the tool box is demonstrated in that the treatment structure can change (simpler or complex) within the same design structure. For the four-way treatment structure, it is by far easier to work with temperatures and recipes than it is to work with the four-way treatment structure. This structural effect is shown by the first column in the table with the different shading patterns for the temperature analysis, for the recipe analysis and for the recepie by temperature analysis.

SUMMARY OF DESIGN TOOLS IN TOOL BOX

The first point to remember is there is no substitute for knowledge of the process used to generate the data and there is no substitute for the knowledge of how the tools in the tool box work. To this point three types of tools have been described, 1) Experimental Unit tools, 2) Treatment Structure and 3) Design Structure that can be used to help design a study or help determine how an existing experiment was conducted so the process of the total design of the experiment can be incorporated into an appropriate analysis. It is important to notice that the experimental unit tools are used to connect the elements in the treatment structure to the elements in the design structure and the combination of all three are used to construct an appropriate model. The bread making experiment and the pig vaccination study provide really good examples of how important it is to be able to use the tools. The next set of examples are presented in an attempt to help you understand the structures or concepts and how to use the tools to enable you identify more appropriate designs and thus provide more appropriate analyses of the resulting data. The tools are more easily used when one can help design the study, but it is much more important to have an understanding of the tools when someone brings data and you have to determine the steps in the process of running the experiment in an attempt to salvage information from the collected data.

Table 8. Analysis of variance table for 4-way treatment structure in a strip-plot design structure for the bread making study with shaded errors relating to the original two-way treatment structure.

	Source	df	EMS
	BLOCK	1	$\sigma^2_{loaf} + 4\sigma^2_{oven} + 2\sigma^2_{block}$
Temperatures	**STOVE**	1	$\sigma^2_{loaf} + 4\sigma^2_{oven} + \phi^2(S)$
	TEMP	1	$\sigma^2_{loaf} + 4\sigma^2_{oven} + \phi^2(T)$
	STOVE*TEMP	1	$\sigma^2_{loaf} + 4\sigma^2_{oven} + \phi^2(ST)$
	OVEN(STOVE TEMP)	3	$\sigma^2_{loaf} + 4\sigma^2_{oven}$
Recipes	**FLOUR**	1	$\sigma^2_{loaf} + 4\sigma^2_{batch} + \phi^2(F)$
	YEAST	1	$\sigma^2_{loaf} + 4\sigma^2_{batch} + \phi^2(Y)$
	FLOUR*YEAST	1	$\sigma^2_{loaf} + 4\sigma^2_{batch} + \phi^2(FY)$
	BATCH(FLOUR YEAST)	3	$\sigma^2_{loaf} + 4\sigma^2_{batch}$
Temperatures*Recipes	**STOVE*FLOUR**	1	$\sigma^2_{loaf} + \phi^2(SF)$
	STOVE*YEAST	1	$\sigma^2_{loaf} + \phi^2(SY)$
	STOVE*FLOUR*YEAST	1	$\sigma^2_{loaf} + \phi^2(SFY)$
	TEMP*FLOUR	1	$\sigma^2_{loaf} + \phi^2(TF)$
	TEMP*YEAST	1	$\sigma^2_{loaf} + \phi^2(TY)$
	TEMP*FLOUR*YEAST	1	$\sigma^2_{loaf} + \phi^2(TFY)$
	STOVE*TEMP*FLOUR	1	$\sigma^2_{loaf} + \phi^2(STF)$
	STOVE*TEMP*YEAST	1	$\sigma^2_{loaf} + \phi^2(STY)$
	STOVE*TEMP*FLOUR*YEAST	1	$\sigma^2_{loaf} + \phi^2(STFY)$
	LOAF ERROR	9	σ^2_{loaf}

SCHOOL BASED PHYSICAL ACTIVITY STUDY

A hot topic in elementary schools is to try to influence school obesity by using a randomized control trial. The process is to select a set of schools and assign an intervention to half of the schools and then compare the effect of the intervention by some measure of physical activity (denoted by PA) or some measure of obesity (such as body mass index, denoted by BMI). This type of design seems like a simple process, so let's start. The first step is to identify the types of students one wants to evaluate, which in this case are third, fourth and fifth graders, and to identify the population of schools in the area of interest. The population of schools are stratified into four groups based on ethnicity (more than 50% white and less than 50%) and income (low income where there are at least 50% are receiving assistance and high income there at least 50% are not receiving assistance), as shown in Figure 18. Next, randomly select four schools from each stratum of schools to be included in the study. Within each stratum, randomly select two schools to receive the intervention program and the other two will not receive the intervention program. As a side note, a research group may not stratify the schools prior to selection and uses income and ethnicity variables as covariates in the modeling process. The group has devices that the student can wear to measure the intensity and duration of activity where they wear the device for one week. They only have enough devices so that students at four schools can be measured during the same time. It will take 4 weeks to obtain the data. The physical activity is to be measured at three time points, baseline, end of the first semester and the end of the second semester. The baseline data can be treated for modeling in one of two ways. Baseline could be one of the levels of time or baseline could be considered as a covariate. Either way, the baseline data must be incorporated

into the analysis. The baseline for this example is considered as a level of time where time has three levels, baseline, 1st semester and 2nd semester. Within each school there are students from three grade levels, 3rd, 4th and 5th as well there are both female and male students.

The process started with the selection of 16 schools as shown in Figure 18. The second step of the process is to randomly select four schools to be evaluated during week 1 with sets of 4 more schools to be evaluated during weeks 2, 3 and 4. Without some careful planning as to which schools are evaluated during each week there can be a great loss of information. It was decided that income level was the least important factor so schools were to be assigned to weeks so that income was confounded with week. The process is to randomly select either a low or a high income stratum. Next, randomly select one of the intervention schools from the two schools in the low ethnicity group and randomly select one of the intervention schools from the two schools in the high ethnicity. Repeat the process for the non-intervention schools. At this point, four schools with the same income classification have been selected to have the PA of the students evaluated during week one. Repeat the process for each of the three other weeks providing a structure as shown in Figure 19. The process of using the devices to obtain the PA of the students generates a blocking factor and the levels of income are completely confounded with those blocks. All of the students within a set of four schools will be given a device and will be trained to use the device the first day. The devices will be returned at the end of the fifth day where the data will that is stored on the device will be down loaded to a computer file.

The discussion now revolves around the identification of the factors in the treatment structure and the factors of the design structure. There are schools that are classified in two ways, by income (high and low) and by ethnicity (high and low). Two schools within each stratum will be assigned to a level of intervention (yes, no). The groups of students within each school are classified as to being in a particular grade (3rd, 4th, or 5th) and the students are classified as to their gender, (male or female). The students within a school are evaluated three times during the year.

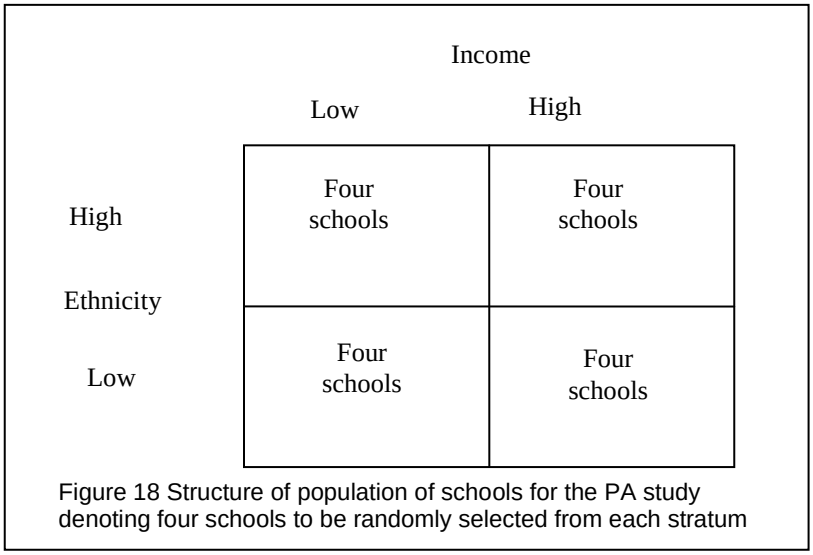

Figure 18 Structure of population of schools for the PA study denoting four schools to be randomly selected from each stratum

Within each school there is a strip-plot structure and a split-plot structure. Figure 20 shows that within a school there are three grades and these schools are measured three times. This generates a strip-plot design structure and then within each of the grades there are male and female students, which are a split-plot structure of the grade. There are 5 experimental units. The first one is the set of four schools that are evaluated for PA during the same week. The second experimental unit is a single school that is from a level of ethnicity assigned to the intervention or non-intervention. Then there are three grades per school which are split-plots (or subgroups) of the schools, there are three times (time intervals) which are split-plots of the schools where the grades and times form a strip-plot on each school, and there are males and females within each class which are split-plots of the grade and strip-plots of the time intervals. The treatment structure has 6 factors with income by ethnicity by intervention by grade by sex by time. The design structure has 7 different sizes of experimental units with (1) four schools assigned to levels of income, (2) a single school assigned to levels of ethnicity and intervention, (3) a group of students assigned to a grade, (4) time interval measured within each school, (5) a time interval of a group of students within a grade, (6) male and female students are within a grade, and (7) a time interval of a female or male student on which the value of PA is determined.

Some could decide that time is a repeated measurement on the student instead of the school. The reality is that it is probably somewhere in between as the student is at home (away from school) for about as much time as they are in the school setting during each day. Moving time to a repeated measurement on the student would provide a different analysis and its construction is left to the reader.

The above describes the woven structures of this study so the next step is to break the structures down and provide an analysis for each. The overall analysis can be obtained by putting the analyses of the parts together to get one.

Level of Income and week	Ethnicity=low Intervention	Ethnicity=high Intervention	Ethnicity=low No Intervention	Ethnicity=low No Intervention
Week 1 Income =low	1	10	3	7
Week 2 income=high	9	4	8	12
Week 3 Income=low	16	2	13	11
Week 4 Income= low	6	15	5	14

Figure 19. Assignment of schools (renumbered from 1-16) to income*ethnicity*intervention combinations.

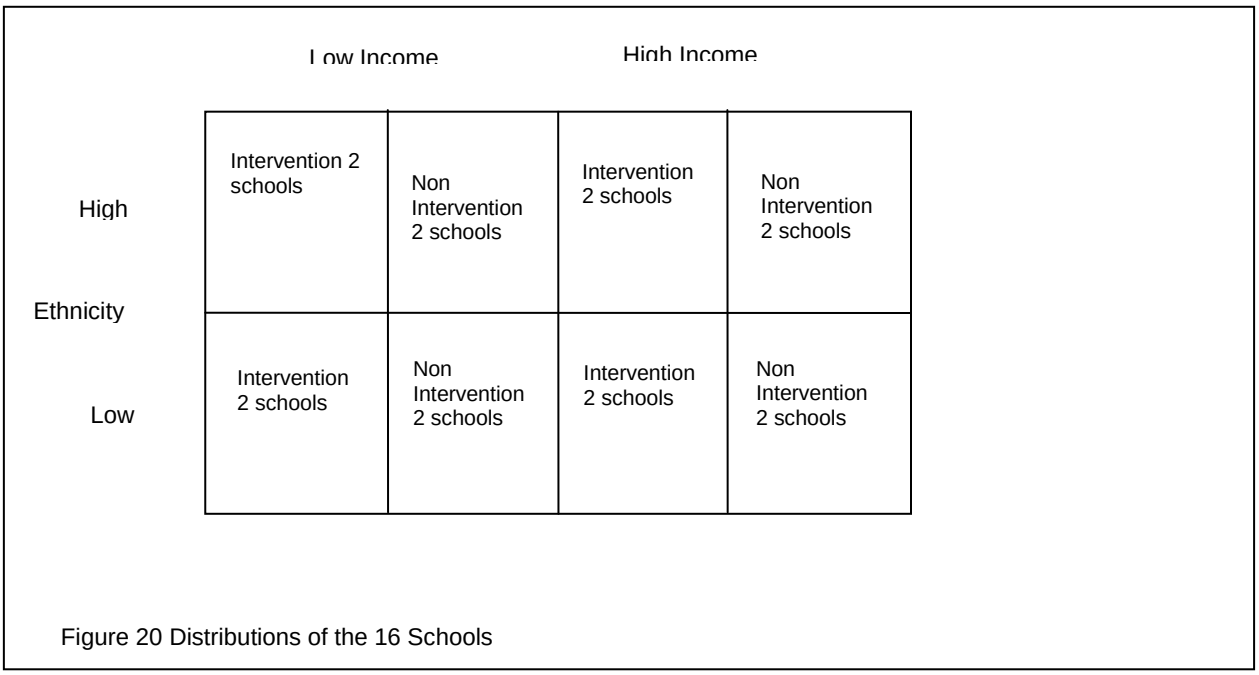

Figure 20 Distributions of the 16 Schools

The structure of the schools is a three-way treatment structure in a split-plot design structure where the four schools assigned to an income is the large experimental unit and the individual schools are the small size experimental unit as displayed in Figure 20. The corresponding split-plot design structure analysis of a measurement on each school is displayed in Table 10 where there are two errors, one for a group of 4 schools grouped together to obtain the PA measurements during a single week and the single school from an level of ethnicity and intervention that is measured within a week.

Table 10. Four school and one school analyses in a split-plot design structure

Source	df	Tester
Income(In)	1	Error(4 Schools)
Error(4 Schools)	2	
Ethnicity(E)	1	Error(school)
Intervention(I)	1	Error(school)
Ethnicity*Intervention	1	Error(school)
Ethnicity*Income	1	Error(school)
Intervention*Income	1	Error(school)
Ethnicity*Intervention*Income	1	Error(school)
Error(school)	6	

Ignoring the time intervals (left hand side of Figure 21), the grades form split-plots of the schools so a split-plot analysis is needed with grade and the three factors assigned to the schools as shown in the first two columns in Table 11. Ignoring the grades (middle of Figure 21), the three times are split-plots of the schools and the split-plot analysis with the three factors assigned to schools are in the two center columns of Table 11. Then Grade and Time form a strip-plot on the schools (right side of Figure 21), so the Grade by Time interaction analysis is in the last two columns of Table 11.

Table 11. Grade and time for a strip plot within each school, but individually they are subplots within each school

Grade analysis a subplot of school		Time analysis a subplot of school		Grade*time analysis a strip-plot of school	
Source	df	Source	df	Source	df
G	1	T	1	T*G	1
In* G	1	In*T	1	In*T*G	1
E* G	1	E*T	1	E*T*G	1
I* G	1	I*T	1	I*T*G	1
I*E* G	1	I*E*T	1	I*E*T*G	1
In*E* G	1	In*E*T	1	In*E*T*G	1
I*In* G	1	I*In*T	1	I*In*T*G	1
I*E*In* G	1	I*E*In*T	1	I*E*In* T*G	1
Error(grade)		Error(time)		Error(grade*time)	

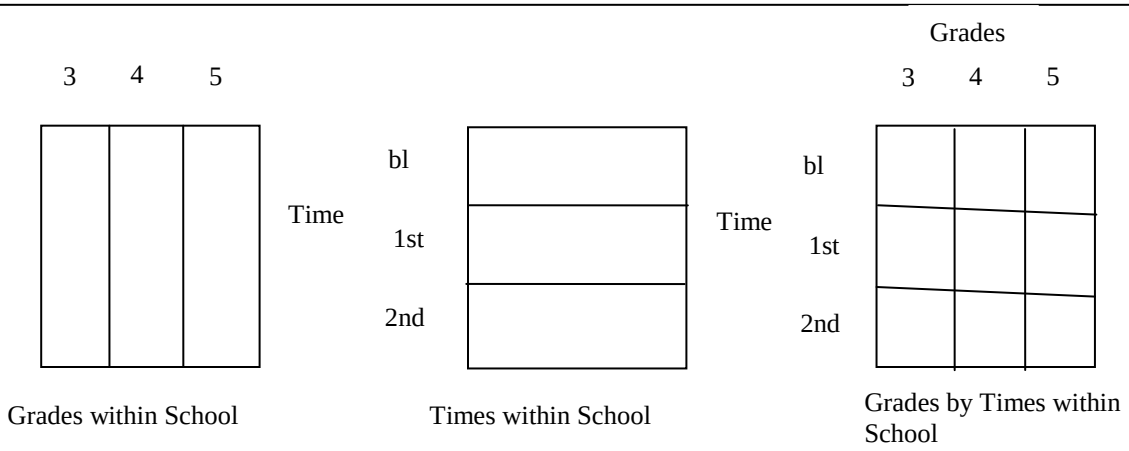

Figure 21 Split-plot for grades and time and strip-plot within a school

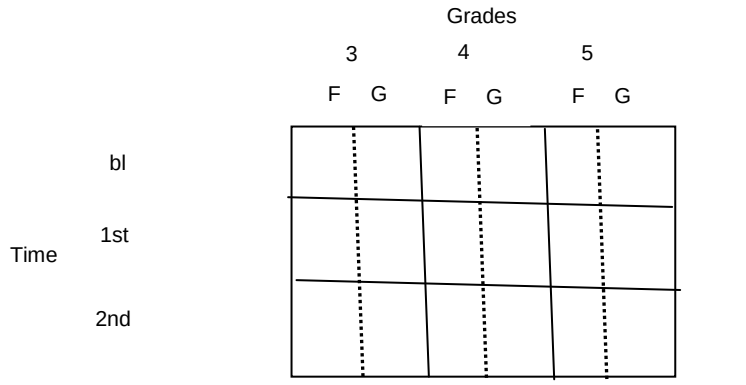

Figure 22 Gender is sub-plot of grade and strip plot of time within grade

The students are split-plots of each of the grades and the levels of gender are assigned to the students (see Figure 22). The split-plot analysis for students is left hand side two columns in Table 12. The levels of time are strip-plots on the students within the grades. The time by gender analysis is in the right two most columns in Table 12 where the error term is the variability of the time intervals within a student.

Table 12. Gender is sub-plot of a grade and gender within a grade and time are strip-plots of grades			
Gender(S) is sub-plot of grade		Gender and time are strip-plot of grades	
S	1	S*T	1
In* S	1	In* S*T	1
E* S	1	E* S*T	1
I* S	1	I* S*T	1
I*E* S	1	I*E* S*T	1
In*E* S	1	In*E* S*T	1
I*In* S	1	I*In* S*T	1
I*E*In* S	1	I*E*In* S*T	1
G*S	1	G* S*T	1
In* G*S	1	In* G* S*T	1
E* G*S	1	E* G* S*T	1
I* G*S	1	I* G* S*T	1
I*E* G*S	1	I*E* G* S*T	1
In*E* G*S	1	In*E* G* S*T	1
I*In* G*S	1	I*In* G* S*T	1
I*E*In* G*S	1	I*E*In* G* S*T	1
Error(student)		Error(student*time)	

This complex design involving six factors arranged in many ways can be broken down into its sup-parts all of which are tools in the tool box discussed above. This total design consists of a completely randomized, split-plot and strip-plot design structures.

CHOCOLATE CANDY DISSOLVING EXPERIMENT

Persons dissolve substances in their mouths at different rates and different types of substances dissolve at different rates. A study was devised where 24 persons that were able to eat chocolate were randomly selected from a class of 49 students. Two brands of large chocolate chips were used in the study where the brands were designated as A and B. Twelve students were randomly assigned to each of the brands. The process of measurement is to put the chocolate chip in your mouth and dissolve the chip as fast as possible by putting it on your tongue and rubbing it across the top of your mouth as fast as you can. A researcher was on hand to provide the specified chip to the person and use a stopwatch to determine the time required to dissolve their chip. The data are in Table 13 and a graph of the distributions for each treatment or brand is in Figure 23. The treatment structure is a one-way with two brands of chocolate chips in a completely randomized design structure. A simple one-way analysis of variance model can be used to compare the means of the two brands. A one-way analysis for comparing the means of the two types of chips is in Table 14 where the item of interest is the estimated standard between the difference of the two means. The difference between the two means is -3.09 sec with estimated standard error of 1.80 sec. The significance level for testing the equality of the means (assuming equal variances) is 0.1002, indicating the difference between the mean times of the two brands is not significantly different at a 0.05 type I error rate.

Table 13. Data for candy dissolving time where dissolve_t is number of seconds.									
Obs	Brand	dissolve_t	Brand	dissolve_t	Obs	Brand	dissolve_t	Brand	dissolve_t
1	A	45.0	B	39.1	7	A	28.4	B	39.5
2	A	39.5	B	36.5	8	A	32.7	B	30.1
3	A	31.8	B	33.2	9	A	28.5	B	32.8
4	A	31.0	B	33.8	10	A	29.7	B	37.0
5	A	33.2	B	32.2	11	A	25.3	B	33.9
6	A	31.9	B	34.5	12	A	26.6	B	38.1

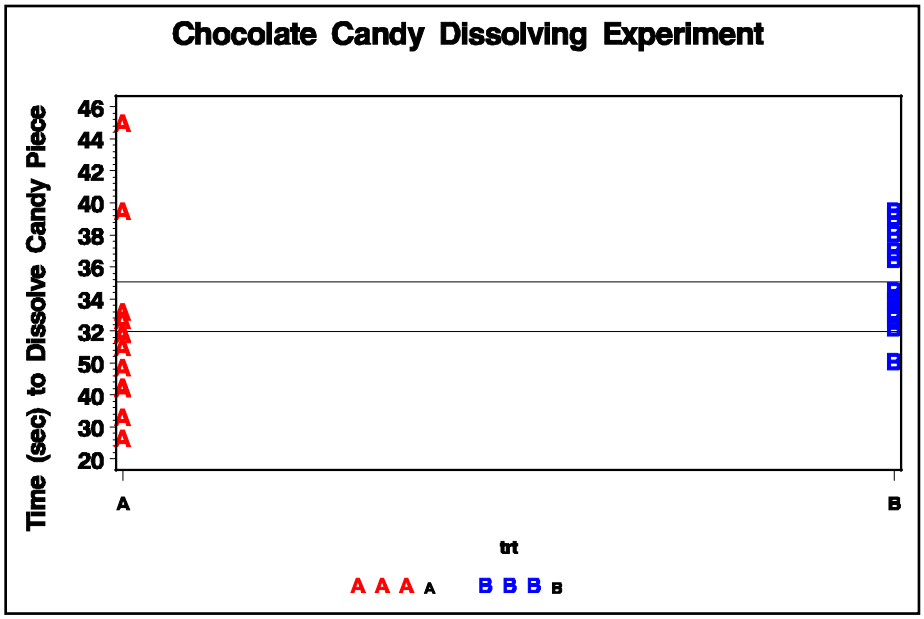

Figure 23 Distributions of times to dissolve type A and B chocolate chips

Table 14 Analysis of variance of time to dissolve two brands of chocolate chips ignoring order

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	57.3504167	57.3504167	2.94	0.1002
Error	22	428.4758333	19.4761742		
Corrected Total	23	485.8262500			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Brand	1	57.35041667	57.35041667	2.94	0.1002

Brand	dissolve_t LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
A	31.9666667	0.1002
B	35.0583333	

Parameter	Estimate	Standard Error	t Value	Pr > |t|
A-B	-3.09166667	1.80167396	-1.72	0.1002

At this point one should step back and really study how the experiment was conducted. The process was to randomly assign the brands to the 24 persons. There was only one stopwatch so the 24 persons had to be evaluated one at a time. The 24 persons were randomly assigning an order and they dissolved their chocolate chip in the order assigned as shown in Table 15.In addition to the order of evaluation, the elapsed time from the beginning of the study was controlled so that one person was evaluated every 3 minutes. A complication was discovered in that the chocolate chips were brought to the class room in a cooler and they were taken out of the cooler as the start of the trial and not replaced. So the chips gradually warmed up to room temperature during the study. It was thought that warmer chocolate chips would dissolve faster than cooler chips. So the data were re-analyzed there the start time was used as a covariate to see if in fact the time to dissolve the chip did in fact depend on the temperature. Table 16 contains the analysis of covariance table for comparing the brand means adjusted for the time chips were out of the cooler. The significance level associated with the slope (start_t) is 0.0031, indicating there is a linear relationship between the time to dissolve the chips and the elapsed time from the beginning of the trial. Figure 24

contains the graph of the two lines for the brands and the respective data points. The least squares means are the evaluation of the two regression lines at the mean starting time (34.5 min) and the difference between line A minus line B at 34.5 minutes is -3.45 sec with an estimated standard error of 1.49 sec. The significance level for testing the difference between the means to be zero is 0.0311, indicating there is a significance difference between the mean times to dissolve the brands of chocolate chips.

Table 15 Data for candy dissolving time where dissolve_t is number of seconds order is the order the person evaluated his/her chocolate chip and start_t is the elapsed time from the start of the study until the person started their part of the process..

Order	Brand	dissolve_t	start_t	Order	Brand	dissolve_t	start_t
1	A	45.0	0	13	B	39.5	36
2	B	39.1	3	14	A	28.4	39
3	A	39.5	6	15	B	30.1	42
4	A	31.8	9	16	A	32.7	45
5	B	36.5	12	17	B	32.8	48
6	A	31.0	15	18	A	28.5	51
7	A	33.2	18	19	A	29.7	54
8	B	33.2	21	20	B	37.0	57
9	A	31.9	24	21	B	33.9	60
10	B	33.8	27	22	B	38.1	63
11	B	32.2	30	23	A	25.3	66
12	B	34.5	33	24	A	26.6	69

Table 16 Analysis of covariance of the time to dissolve the chocolate chip using the elapsed time (start_t) as a covariate.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	206.0883974	103.0441987	7.74	0.0030
Error	21	279.7378526	13.3208501		
Corrected Total	23	485.8262500			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Brand	1	71.1348300	71.1348300	5.34	0.0311
start_t	1	148.7379808	148.7379808	11.17	0.0031

Brand	dissolve_t LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
A	31.7863782	0.0311
B	35.2386218	

Parameter	Estimate	Standard Error	t Value	Pr > |t|
A-B	-3.45224359	1.49391626	-2.31	0.0311

There is a very important message to be learned. If we understand exactly how the study was conducted including the conditions under which the samples were obtained there is a greater chance we can do a better job in getting closer to an appropriate analysis of the resulting data.

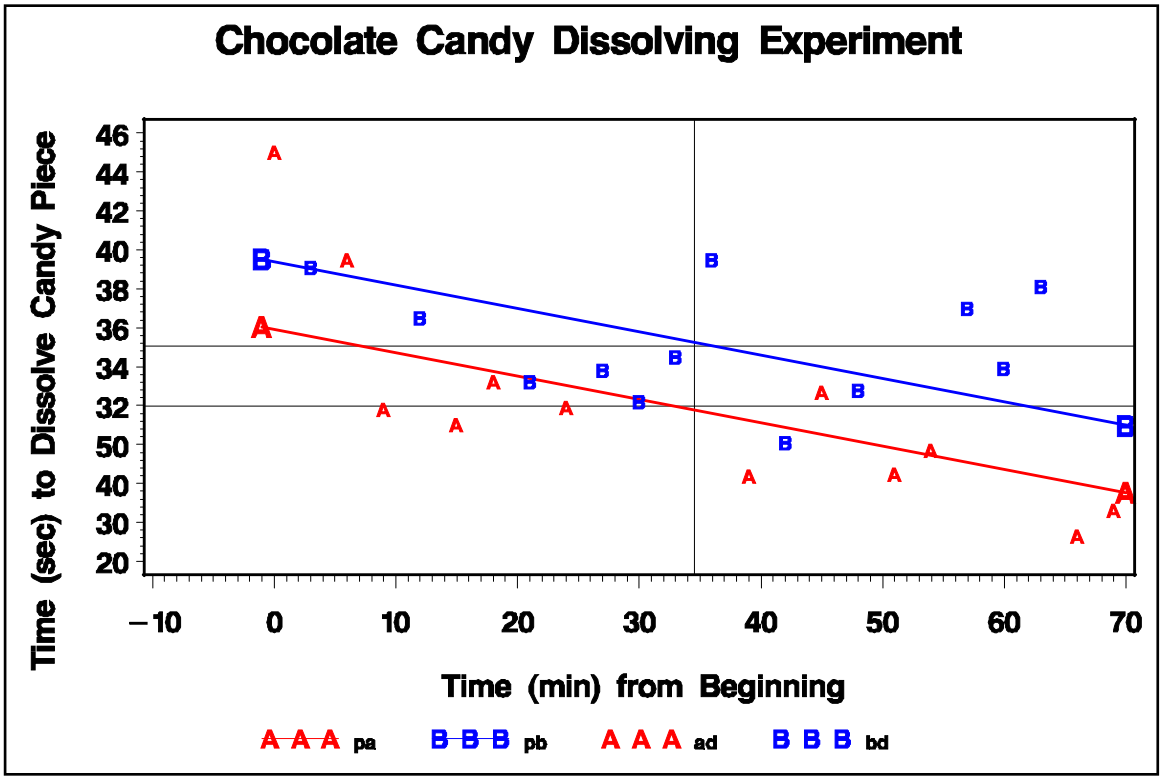

Figure 24 Graph of time to dissolve a chocolate chip vs. elapsed time for types A and B

SEMICONDUCTORS

The construction of semiconductors or computer chips is a multiple step process where at each step one is either adding a layer of material or removing a layer of material. This study involves looking at the process during six steps toward the end of construction. The six steps are denoted by 1, 2, 3, 4, 5, and 6. The unit of construction is called a wafer. During many steps the sets of wafers are treated together and during some of the steps wafers are treated individually. In this study six wafers from a set of 24 wafers were being used to evaluate the effects of small changes in the process. In this case the six wafers are treated together during steps 1 and 2. During step 3, the wafers are treated individually and during steps 4, 5 and 6 they are treated as a group. During step three, three of the wafers were randomly assigned to the current voltage and three of the wafers were randomly assigned to a lower voltage. The graphic in Figure 25 denote which wafers (chips) were assigned to treatment a and which were assigned to treatment b during step 3. Figure 26 is a display of a region of interest on the wafer there are three line sizes, two row treatments, four column treatments where two rows are nested within the levels of trt, 1 or 2. The wafers or chips are assumed to represent a population of wafers that are yet to be produced, so the chips are considered as random effects.

The construction of an appropriate analysis can be done is a straight forward manner if the design is broken down into its parts. Figure 25 represents the application of the levels of the treatment at step 3 providing a one-way treatment structure in a completely randomized design structure and the entries in the analysis of variance table are in Table 17.

Table 17. The chip level of the analysis of the computer chip data	
Source	df
Trt3	1
Chips(trt3)=Error(chip)	4

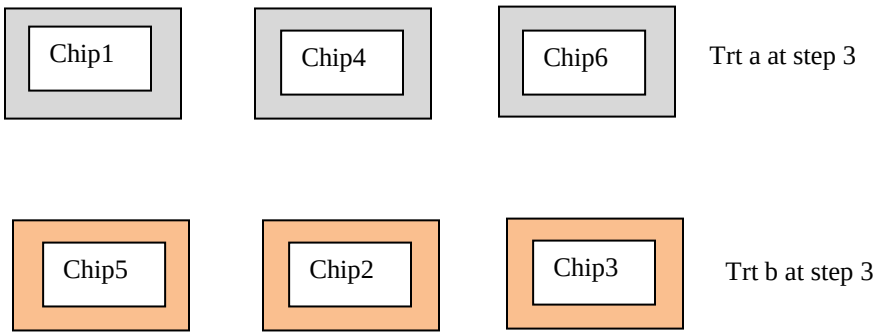

Figure 25 Random assignments of chips to the levels of step three treatment

The levels of the row treatments are applied to the rows of each of the chips. Thus the chips are divided in half as shown in Figure 26 and the levels of the row treatment are assigned to the rows. The chips are blocks for the row treatments, but since the blocks are stratified by the application of the levels of TRT3, the resulting design is that of a split-plot. The half chip analysis is provided in Table 18. Likewise each chip is divided into four columns where two columns are assigned to column treatment 1 and two are assigned to column treatment 2. Within the two columns of treatment 1 there are two other levels of the column treatment and another two levels of the column treatment are applied to the levels of treatment 2. Thus this part of the treatment structure is nested as the levels of Column treatment are nested within the levels of treatment. The chips are blocks, but since blocks are cofounded with TRT3, the design structure is a split-plot. The ¼ chip analysis is for a split-plot design structure with a two level nested treatment structure, which is displayed in Table 19. The sub-plot treatments are the levels of the column treatments nested within the levels of the square treatments. The fixed effects are square treatments, column treatments nested within the square treatments and the interaction with the third step treatment. The error term is constructed by pooling the chip by square treatment interaction pooled across the levels of the third step treatment pooled with the chip by column treatment nested within square treatment pooled across the levels of the third step treatment.

Table 18. The row analysis where the entry Chip consists of the 5 df from Table 1.	
Chip	5
Row	1
Row*TRT3	1
Error(1/2 chip)=row*chip(trt3)	4

Table 19. One fourth chip analysis for a split-plot design structure with a nexted treatment structure. The df for Chip are from Table 1	
Source	df
Chip	5
T	1
C(T)	2
TRT3*T	1
TRT3*C(T)	2
Error(1/4 chip)	12

The levels of the row treatment and the levels of the column treatments nested within the square treatment form a strip-plot over the chip. The analysis of variance will include fixed effects that involve a term from the row interaction of a term from the column and their interactions with the third step treatment. The analysis of variance table for the 1/8th part of a chip is in Table 20. Remember this part of the analysis only involves the interaction of the rows and

columns.

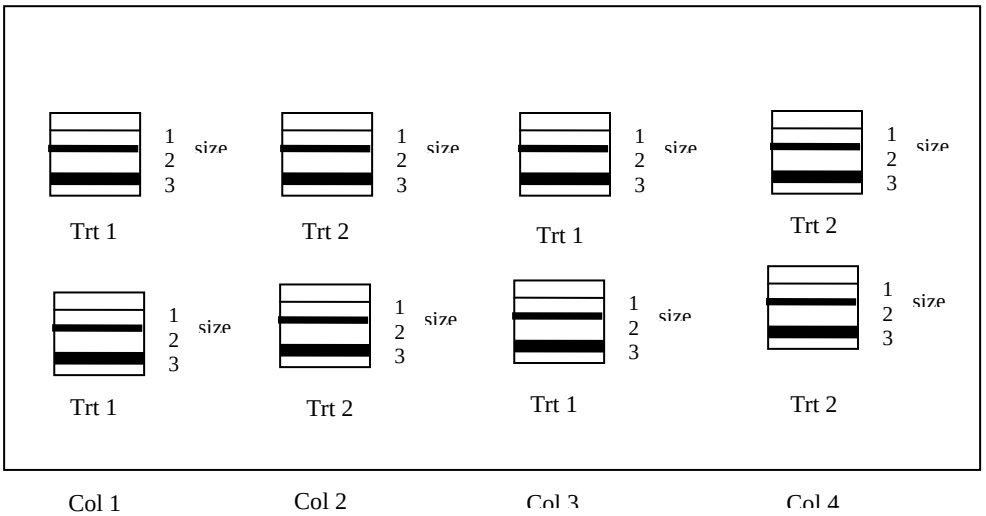

Figure 26 Structure within each of the chips with rows, treatments, columns within treatments and sizes within cells

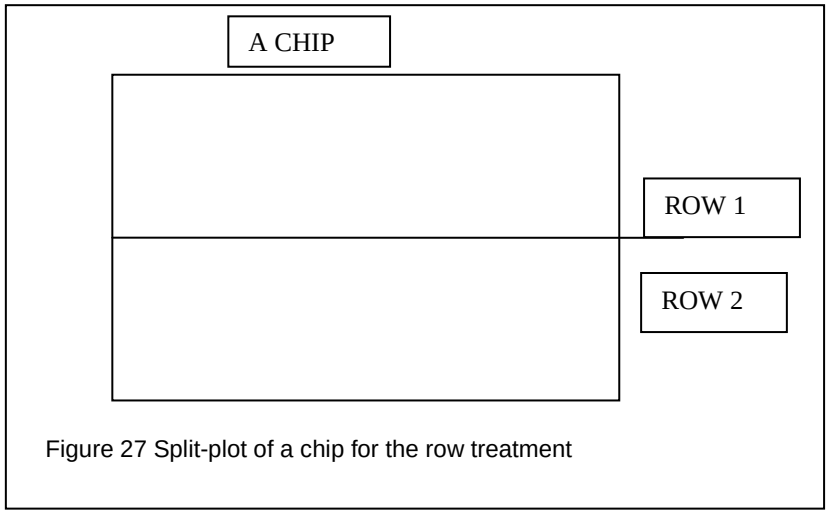

Figure 27 Split-plot of a chip for the row treatment

Table 20. The strip-plot analysis for the row and column treatments for the chips	
Source	df
R*T	1
R*C(T)	2
TRT3*R*T	1
TRT3*R*C(T)	2
Error(1/8 th chip)	12

Finally, there are three widths of circuits within each square which provide a split-plot with the squares. The remaining analysis involves the size and the interaction of all of the treatment structure effects in all of the other analyses. The circuit part of the analysis is in Table 21.

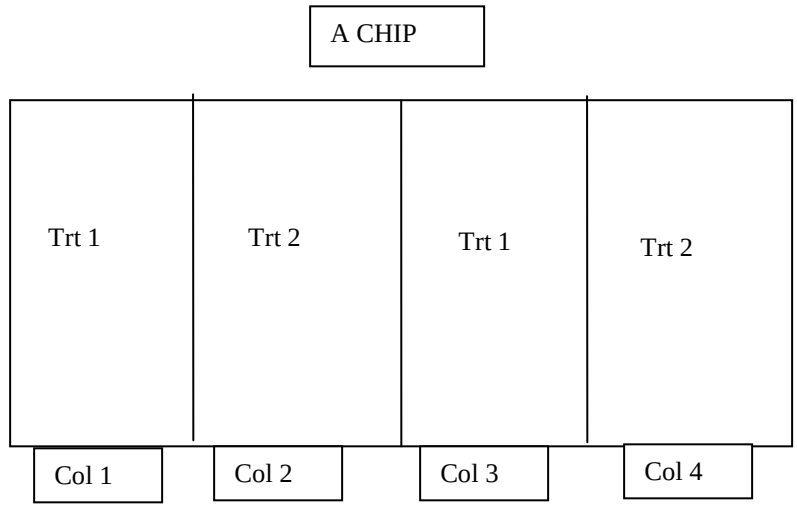

Figure 28 Split-plot design structure for column treatments nested within treatments

Table 21. The circuit level the analysis of variance which includes the levels of size (S) and the interactions of size with all of the other effects in the treatment structure

Source	df
S (size)	2
TRT3*S	2
Row*S	2
Row*TRT3*S	2
T*S	2
C(T)*S	4
TRT3*T*S	2
TRT3*C(T)*S	4
R*T*S	2
R*C(T)*S	4
TRT3*R*T*S	2
TRT3*R*C(T)*S	4
Error(circuit)	64

The SAS® system code that can be used to carry out the complete analysis to combine the above analyses of variance tables is

```
proc mixed ;
class t3 cp r co t s;
model y=    t3
      r r*t3
      t co(t) t*t3 t3*co(t)
      r*t r*co(t) r*t*t3 r*t3*co(t)
      s s*t3 s*r s*r*t3 s*t s*co(t) s*t*t3 s*t3*co(t)
      s*r*t s*r*co(t) s*r*t*t3 s*r*t3*co(t) ;
```

```
random cp(t3) cp*r(t3) cp*t3*co(t) cp*r*co(t t3);
```

where t3, cp, r, co,t and s denote the levels of the third step treatment, the chips, the rows within a chip the columns of a chip nested within the treatments and the size of the circuits respectively. The terms in the random statement define the respective error terms with the exception of the residual.

CONCLUSION

The teaching of design of experiments by presenting the material for one design at a time as if it is its own entity, such as split-plot or Latin square, misses out on the application of designs to the real world where experiments are conducted in several steps. Making sure the students understand these steps exist and how to use the tools of design of experiments to connect these steps in order for an appropriate analysis to be constructed is a recipe for successfully educating students about how designs of experiments work or are needed in the real world. In our teaching of the design of experiments we have failed to inform our students of these importance connections and that additional design features can be generated while the experiment is being executed. As a final point, we must stress to the researcher that all of the information is important to improve the quality of the data and resulting analyses.

REFERENCES

Milliken, G. A., Shi, X., Mendicino, M., and Vasudev, P.K. 1998. “Strip-plot design for two step processes”, *Quality and Reliability Engineering International*, 14: 1-15.

Milliken, G. A. and D. E. Johnson. 1992. Analysis of Messy Data, Volume 1: Designed Experiments. Chapman & Hall/CRC, New York.

Milliken, G. A. and D. E. Johnson. 1989. *Analysis of Messy Data, Vol. 2: Nonreplicated Experiments.* Chapman & Hall/CRC, New York.

Milliken, G. A. and D. E. Johnson. 2001. *Analysis of Messy Data, Vol. 3: Analysis of Covariance*. Chapman & Hall/CRC, New York.

Milliken, G. A. and D. E. Johnson. 2009. *Analysis of Messy Data, Vol. 1: Designed Experiments, 2nd Edition*. Chapman & Hall/CRC, New York.

Littell, R, G.A. Milliken, W. Stroup, R. Wolfinger. 1996. *SAS System for Mixed Models*. SAS Institute Inc. Cary NC., 1996.

Littell, R, G.A. Milliken, W. Stroup, R. Wolfinger, O. Schabenberger. 2006. *SAS for Mixed Models, 2nd Edition*. SAS Institute Inc. Cary NC.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: George A. Milliken
Enterprise: Milliken Associates, Inc.
Address: 1401 Deep Creek Lane
City, State ZIP: Manhattan, KS 66502 USA
Work Phone: 785 539 9469
Fax: 785 776 0630
E-mail: Milliken@ksu.edu
Web: stat911.com