AN ABSTRACT OF THE DISSERTATION OF

Seyed Alireza Mohseni for the degree of Doctor of Philosophy in Industrial Engineering presented on September 8, 2017.

Title: Optimal Experimental System Design.

Abstract approved:

_____

David Kim

Physical experimentation in various scientific and engineering areas continues to be a useful and often necessary approach that is applied in research, development, and for general problem solving. The experimental design process however, is iterative and often involves trial and error. Alternative designs are proposed and evaluated on cost and time requirements. They are also evaluated with respect to various statistical criteria such as statistical power for detecting effects if they exist. Selected experimental design alternatives then continue the design process and undergo additional modifications until the objectives of the experiment appear to be met, or are met as closely as time and cost constraints permit. The objective of this research is to develop a methodology for "Optimal Experimental System Design" where cost and time are integrated with experimental design selection so that optimal cost and time feasible experimental designs can be explored computationally.

Optimal Experimental System Design

by
Seyed Alireza Mohseni

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 8, 2017
Commencement June 2018

Doctor of Philosophy dissertation of Seyed Alireza Mohseni presented on September 8, 2017

APPROVED:

Major Professor, representing Industrial Engineering

Head of the School of Mechanical, Industrial & Manufacturing Engineering

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Seyed Alireza Mohseni, Author

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to my major advisor David Kim. Without his continuous guidance, help, and support throughout the past six years, it would not have been possible for me to reach this point. He taught me how to be a better student, a better scholar, and a better professional. I learned a lot from him and for all that, I am always grateful to him.

I also want to thank all my friends who became like family to me and always cared for me during these years.

Lastly, I would like to thank my parents, whom I miss more than I can put into words, for all their heartwarming love, support, and patience.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

1. Introduction

Physical experimentation in the sciences and engineering continues to be a useful and often necessary approach that is applied in research, development, and for general problem solving. In engineering, physical experimentation is utilized for multiple types of problem-solving such as identifying the key factors affecting suddenly degraded product quality, and identifying the process parameters that have the largest effects on process variability. Physical experiments are also frequently used in product and process design optimization. The development of analytically sound methods for the design of experiments and the analysis of experimental data has paralleled, supported, and encouraged the use of physical experimentation. These developments that started with the work of Sir Ronald A. Fisher in the 1920's, have led to the recognition of "Design of Experiments (DOE)" as a specific subject and accepted methodology, which in turn opens up experiments and their design to evaluation and critique. Assuming that after technical evaluation and critique, an experiment is deemed to be appropriately designed, conducted, and analyzed, the ultimate evaluation of the experiment is whether it met the needs of the experimenter. However, given the time and resources that were utilized to design and complete the experiment the following questions may be asked:

- Could more have been accomplished with the same experimental resources utilized (including time)?
- Could the same results have been obtained with fewer resources?
- Can a good experimental design that explicitly considers cost and time be produced by an algorithm?

These questions are in part addressed by developments that have occurred in the DOE field. One such development area is "optimal experimental design" that has focused on the evaluation and creation of experimental designs based on optimizing well-defined mathematical and statistical criteria. Developments in this area have led to algorithms for constructing experimental designs that are optimal with respect to specific mathematical criteria, given that a specific number of experiments will be conducted. However, these algorithms do not explicitly consider experimental cost and time. Other DOE developments

in the area of fractional factorial designs attempt to reduce the total number of experiments needed to make inferences with respect to particular factors and interactions. Such fractional designs can be used as experimental design alternatives in an iterative design procedure that "externally" evaluates whether a particular fractional design meets statistical, cost, and time criteria.

However, to date the explicit consideration of experimental cost and time has not been explicitly integrated with experimental design selection with the objective of identifying optimal cost and time feasible experimental designs algorithmically. Addressing this situation is the topic of the research presented. The objective of this research is to:

Develop a methodology for "Optimal Experimental System Design" where cost and time are integrated with experimental design selection so that optimal cost and time feasible experimental designs can be identified algorithmically.

To develop a better understanding of this research an "experimental design" will be defined, and terminology describing features of the design will be presented. This will then be followed by a presentation of an example to clarify the practical problem being addressed.

An experimental design is defined by the experimental factors included, the response to be measured, the total number of experiments conducted, and the manner (experimental order) in which the experiment is conducted. The experimental factors and response must be specified by the experimenter and are considered inputs to a design. The experimental design is then defined by the number of levels for each factor, the treatment combinations (factor combinations) to be tested, the number of replicates completed for each treatment combination, and the order and timing of how all the experiments are conducted. It is assumed that standard experimental procedures exist and will be used. Following the terminology used in Milliken and Johnson (2006), the treatment combinations and replicates conducted define the "*treatment structure*" of the experiment, and the experimental order is referred to as the "*design structure*". As will be seen in the following example, the experimental treatment and design structure are not independent given cost and time constraints.

1.1 Example Experiment

Consider an experiment investigating the fuel consumption of a specific type of vehicle engine. The two experimental factors being investigated are oil brand and fuel brand. There are multiple brands of oil and fuel, and ideally all brands can be tested and evaluated. Additionally, the oil and fuel brands can be separated by the primary types of additives they contain. These primary additives separate oil brands into two and fuel brands into three categories. Even though both oil and fuel are being examined, given a choice the experimenters are more interested in fuel differences. For ease of exposition it will be assumed that if multiple engines exist (experimental units) then the differences between engines is insignificant, and that resources only allow one test to be conducted at a time. A total experimental budget exists, and once the experiments are started they must be completed within a specific time period since the experimental facilities can be reserved for a set maximum time period. When the oil type is changed a significant amount of cost and time is required to purge the prior oil type from the engine. Furthermore, changing between oil types that use different primary additives takes additional time and cost. Fuel type changes require some purging but are much faster and can also be completed while oil is being purged. Both oil and fuel can be purchased as needed. Each fuel consumption experiment requires the same amount of time, which can be accurately estimated and includes setup and takedown but does not include the time required to change oil.

Given this description of the "experimental system" the following are some questions that may be asked with respect to the experimental design.

- How many levels of oil and fuel should be included? Should factor levels be defined by the primary additive type?
- How many replicates of each treatment combination will be run?
- Should any randomization restrictions such as blocking on oil type or primary oil additive be utilized in the design structure?
- If experiments are completed over multiple days, how can days be included as an experimental factor and how will this affect the design structure?

These are only some of the possible questions that may arise when designing the experiment and the answers to one question will affect answers to other questions. For

example, it is clear that if more fuel and oil type combinations are tested, then fewer replicates can be run since there are both cost and time constraints. However, if "blocking" (a specific type of randomization restriction) on primary oil additive is part of the design structure then more replicates of oil or fuel types may be included since fewer changes between oil brands with different primary additives will occur, and these types of oil changes incur the most cost and time. Figure 1.1 shows diagrams for two different experimental designs that test four different oil types and three different fuel types. In the randomized complete block design (RCBD) the treatment combinations within a randomly selected box outlined with thicker lines are run in a randomized order before treatment combinations in any other "block" are run. Both designs permit two replicates to be completed within the cost and time constraints, however the randomized complete block design or RCBD (the second design) will require less time and cost since only four oil changes (including the first oil tested) are needed. However, this savings in cost and time comes at the cost of not being able to distinguish oil effects from the effects of unknown nuisance factors. If the possible effects from unknown nuisance factors (confounded with oil effects) is judged to be minimal, then the trade-off for lower cost and less time will be worthwhile.

Completely Randomized Order					
		Oil Type			
		1	2	3	4
Fuel Type	1	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2
	2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2
	3	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2

Randomized Complete Block Design					
		Oil Type			
		1	2	3	4
Fuel Type	1	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2
	2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2
	3	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2	Rep1 Rep2

Figure 1.1. Completely randomized two-factor experiment and an RCBD.

The experimental design process is iterative and often involves trial and error. Alternative designs are proposed and evaluated on cost and time requirements. They are also evaluated with respect to various statistical criteria such as the effects and interactions for which statistical tests are available. This is also where criteria developed in the area of optimal experimental design can be utilized to distinguish experiments. Some experimental design alternatives may then move further on in the design process and undergo modifications until the objectives of the experimentation appear to be met, or are met as closely as time and cost constraints permit.

The objectives of this research are to explore how the search for a set of optimal candidate designs can be completed computationally. From the prior example, it can be seen that the number of possible designs for the fairly simple example experimental scenario is very large when the different possible treatment structures are combined with the many different possible design structures. Figure 1.2 shows two additional experimental designs that

results from a different design structure applied to the same treatment structure as found in Figure 1.1.

		Replicate 1				Replicate 2			
		Oil Type				Oil Type			
		1	2	3	4	1	2	3	4
Fuel Type	1	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	2	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	3	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2

		Replicate 1				Replicate 2			
		Oil Type				Oil Type			
		1	2	3	4	1	2	3	4
Fuel Type	1	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	2	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	3	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2

Figure 1.2. Alternative experimental designs resulting from different design structures.

The first experimental design shown in Figure 1.2 is one design referred to as a "split-plot design" (Montgomery 2013), where "blocking" on the oil factor is utilized, and this blocked experiment is completed twice (a randomization restriction on replicates is imposed). In the second design blocking on oil type occurs but the order in which oil types are tested is randomized. The randomization restrictions are shown in Figure 1.2 by the boxes with thicker borders. More general randomization restrictions are possible that add to the number of possible experimental designs. For example, in the top design in Figure 1.2, instead of a randomization restriction by replicate, and then by oil type, the restriction can be imposed so that the replicates for oil types 1 and 2, or oil types 3 and 4 are all completed in a randomized order. This design is depicted in Figure 1.3. In this example oil types 1 and 2 may use the same primary additive, as do oil types 3 and 4 so that this design avoids the longer times and higher costs of changing between oil types with different additives.

		Replicate 1				Replicate 2			
		Oil Type				Oil Type			
		1	2	3	4	1	2	3	4
Fuel Type	1	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	2	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2
	3	Rep1	Rep1	Rep1	Rep1	Rep2	Rep2	Rep2	Rep2

Figure 1.3. Another experimental design resulting from a different design structure.

If the designs in Figure 1.2 and Figure 1.3 are compared, the second design in Figure 2 is the least restrictive in the sense of imposed randomization restrictions. Another way to view this is that this experimental design has a design structure that is "closer" to a completely randomized design structure. The top design in Figure 1.2 ("split-plot") has the most restrictive design structure, and the design structure in Figure 1.3 falls between the two designs in Figure 1.2 with respect to randomization restrictions. Intuitively, a less restrictive design structure should have better statistical properties with respect to the main effects and interactions for which statistical tests are available. This turns out to be the case in this example, however fewer randomization restrictions comes with greater cost and time.

1.2 Research Requirements

In order to develop the methodology for "Optimal Experimental System Design", the research objective has been partitioned into the following three specific requirements/needs.

1. An experiment defined by its treatment and design structure should have a unique statistical model.
2. Mathematical models for the cost and time of an experiment as a function of its treatment and design structure.
3. An optimization procedure for searching the space of experimental design structures to find feasible experimental designs that make the best use of limited experimental resources for a given treatment structure.

The remainder of this thesis is organized as follows.

- Chapter 2 provides the required background information regarding different topics and concepts discussed in this research as well as a review of the most related research studies.
- Chapter 3 provides a more detailed overview of the research methodology and the organization of materials discussed in future chapters.
- Chapter 4 discusses the construction of a recursive procedure for enumerating all the possible design structures given the treatment structure of an experiment.
- Chapter 5 goes over the development of unique statistical models for experiments that represent the experimental design and treatment structures.
- Chapter 6 presents the analysis of variance (ANOVA) of experimental results. The ANOVA follows from the statistical model. The methods presented have been previously developed but can now be applied to statistical models that represent one and only one experiment.
- Chapter 7 presents the derivation of mathematical formulas for calculating the expected time and cost of an experiment given as functions of its design and treatment structures.
- Chapter 8 develops a procedure to measure the overall statistical effectiveness of different design structures through the statistical power of their corresponding tests of significance.
- Chapter 9 goes over a case study showing the usage of efficiency (derived in chapter 7) and effectiveness (derived in chapter 8) of design structures to make a selection of design structures most suited for an experiment.

2. Literature Review

This chapter is organized into two parts. The first part presents relevant background information that will be utilized throughout this thesis. Linear statistical models of experimental responses will be reviewed so that the need to extend these models is clarified. This will be followed by a review of randomization restrictions (changes to the design structure) and the "No-Name" approach of Lorenzen and Anderson (1993) for representing design structure in the statistical model. In the second part of this chapter, a review of relevant research that is the most closely related to optimal experimental system design is presented.

2.1 Background Material

Within the following two sections the two core concepts in the DEO field will be discussed. First, the statistical models representing an experiment, and second, the concept of randomization restriction which is vastly utilized in this study.

2.1.1 Models for Experiments

In the DOE field, it has been common practice to assume the experimental response is adequately represented with a mathematical model. The mathematical model then dictates what types of statistical analysis procedures may be applied to the experimental data. The most common "statistical model" of an experimental response is a linear model. In particular, a linear effects model is a representation of how the response variable is affected by the experimental factors and factor interactions in the experimental design. Linear statistical models are ubiquitous in textbooks and the research literature. Montgomery (2005) and Kuehl (2000) are two popular textbooks that use a linear effects models to represent experimental responses. In the prior fuel consumption example in section 1.1, a linear statistical effects model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

where y_{ijk} is the experimental response (fuel consumption) for the *k*th replication with fuel type *i* and oil type *j*. μ is an overall mean, α_i is the differential effect of fuel type *i* from the

overall mean, β_j is the differential effect of oil type j from the overall mean, $\alpha\beta_{ij}$ is the fuel and oil type interaction effect, and ϵ_{ijk} is the experimental "error" or noise term. ϵ_{ijk} is commonly assumed to be independent and normally distributed with a mean of zero, and fixed variance σ^2.

However, the "typical" or "standard" statistical models and, in particular, linear effects models only represent the treatment structure of an experiment. Since there may be multiple design structures associated with the same treatment structure, the experimental design structure must be considered in the interpretation of the results when the analysis of the experimental data is conducted. For example, it will be shown that a linear statistical effects model for a "randomized complete block design" is the same model used for a "two-factor completely randomized design".

2.1.1 Randomization Restrictions and the No-Name Approach

A commonly used approach to reduce the time and cost of an experiment is the use of randomization restrictions. As stated in various publications such as Ganjua and James (1998) or Simpson, Kowalski, and Landman (2004), this is achieved by restricting the randomization on the so called "hard-to-change" factors or factors that are expensive to change. Such "hard-to-change" factors are randomly set at one level, and then all the other treatment combinations at this "hard-to-change" factor level are run before changing the "hard-to-change" factor to another level. This pattern continues until all the experimental runs are carried out. "Split-Plot" designs are an example of a set of experimental designs that employ randomization restrictions. Federer and King (2007) discuss variations on Split-Plot and Split-Block experimental designs. The various randomization restrictions in these experiments restrict the order of experimental execution and thus represent different design structures. These randomization restrictions come at the cost of adding an extra source of variation that typically is not explicitly represented in the statistical model for an experiment. As will be discussed in Chapter 4 there are many ways beyond "Split-Plot" experiments and variations of Split-Plot experiments in which the experimental run order can be restricted. This is the experimental design space from which a specific design structure is selected.

Anderson (1970) was the first to incorporate a randomization restriction term (or factor) into the statistical model for an experiment that explicitly represents the experimental design structure. This new term shows up as an additional random factor with zero degrees of freedom that is nested within the restricted factor. With the experimental design structure incorporated into the statistical model the subsequent data analysis based on the statistical model makes the implications of the randomization restriction on statistical inference clear. Lorenzen (1984) proved the existence of the randomization restriction error term, and Lorenzen and Anderson (1993) utilize the concept as the basis of their "No-Name" approach to experimental design.

To demonstrate how this approach works suppose the fuel consumption experiment was conducted as a randomized complete block design (RCBD) with "blocking" on the oil type factor. The RCBD implies a restriction on the experimental order as described in the introduction. The response is modeled using the conventional linear statistical model (approach one), and the Lorenzen and Anderson approach with a randomization restriction term added to the model (approach two). The linear statistical models and expected mean squares are shown in Figure 2.1.

Randomized Complete Block Design					
		Oil Type			
		1	2	3	4
Fuel Type	1	Rep1	Rep1	Rep1	Rep1
	2	Rep1	Rep1	Rep1	Rep1
	3	Rep1	Rep1	Rep1	Rep1

Approach one:

$$Y_{ij} = \mu + \tau_i + \beta_j + \tau.\beta_{ij} + \varepsilon_{ij}$$

- $MS(Oil)$: $3\frac{\sum_{j=1}^{4}\beta_j^2}{3} + \sigma^2$
- $MS(Fuel)$: $4\frac{\sum_{i=1}^{3}\tau_i^2}{2} + \sigma^2$
- $MS(Oil.Fuel)$: $\frac{\sum_{j=1}^{4}\sum_{i=1}^{3}\beta\tau_{ij}^2}{3*2} + \sigma^2$
- $MS(E)$: σ^2 (Not estimable)

Approach two:

$$Y_{ijl} = \mu + \tau_i + \beta_j + \gamma_{l(j)} + \tau.\beta_{ij} + \varepsilon_{ijl}$$

- $MS(Oil)$: $3\frac{\sum_{j=1}^{4}\beta_j^2}{3} + 3\sigma_\gamma^2 + \sigma^2$
- $MS(Fuel)$: $4\frac{\sum_{i=1}^{3}\tau_i^2}{2} + \sigma^2$
- $MS(Oil.Fuel)$: $\frac{\sum_{j=1}^{4}\sum_{i=1}^{3}\beta\tau_{ij}^2}{3*2} + \sigma^2$
- $MS(E)$: σ^2 (Not estimable)

Figure 2.1. Analysis of a Randomized Complete Block Design using a linear effects model approach (one) and the Lorenzen and Anderson approach (two).

In Figure 2.1, μ is the overall mean, τ is fuel type fixed effect, β is the oil type fixed effect, and γ is the random effect due to the randomization restriction enforced on oil type.

In both approaches the error term in the model is not estimable since there are no replicates for the treatment combinations. If the interaction effect between the two factors is assumed to be nonexistent the interaction mean square serves as the error term. The statistical model and the expected mean squares constructed using approach one are identical to a two factor completely randomized design. This is because the statistical model does not represent the experimental design structure. There is no indication in the models that any oil type affects are confounded with possible unknown nuisance factors correlated with the non-randomized run order. This must be recognized by whoever completes the data analysis.

In the second approach, the design structure (blocking) is reflected in the statistical model so that the results of the randomization restriction on the expected mean squares can be seen. It is explicit (and follows from the model) that the effect of the oil type is confounded

with the randomization restriction error, and thus there will be no direct test for the significance of this effect. In both models, there is a direct test for the presence of a fuel type effect, which is correct since the order of fuel type is randomized.

Utilizing, analyzing, and evaluating different randomization restriction variations and combinations in an experiment will be a major component of the methodology developed in this research, and will be discussed in detail in future chapters.

2.2 Literature Review

2.2.1 The Need for Randomization

The need for randomization of experimental run order was first documented by Fisher (1926). Randomization, or the random assignment of treatment combinations to the experimental runs, is a fundamental principle of the statistically designed experiments. Experiments are run in randomized order to average out the effect of nuisance variation on the measurements to better detect the influence of factors of interest on a dependent variable. Through the construction of various test hypotheses for a simple comparative test Anscombe (1948) and Fisher (1966) discuss the importance of randomization. They discuss the statistical analysis validity and the need for unbiased mean treatment effect estimates, and a valid error variance estimate for comparing differences among treatments. Cox (2009) remarks that the possibility of including bias in treatment allocation, and in implementation and measurement of outcomes is serious in many contexts in which personal choice is even marginally involved. Randomization is stated as the most effective way of avoiding such biases.

2.2.2 Screening Designs and Optimal Designs

The need for reducing or minimization of the required number of experimental runs and generating efficient experimental designs is a longstanding research area in the DOE field. Some of the major topics in this area are 2^k factorial and 2^{k-p} fractional factorial designs, Plackett-Burman designs, and "optimal experimental design". These experimental design topics will be briefly discussed in this section.

2^k factorial designs, often referred to as screening experiments, are experimental designs where the main objective is to identify the factors (out of k total factors) that have the largest effects on the response variable. The number of factors in such experiments is relatively large number, and the premise of such experiments is the "80/20" rule (a large portion of the response is due to a small number of factors). In this category of designs, each factor is tested at two levels referred to as "high" (denoted with a + or 1) and "low"

(denoted by a - or -1) as a way to reduce the total number treatment combinations, and therefore the experimental runs to 2^k.

2^{k-p} fractional factorial designs are a category of experimental screening designs where a selected fraction of the 2^k treatment combinations are run. 2^{k-1} is a one-half fraction, 2^{k-2} is a one-quarter fraction and so on. The reduction in the number of experiments comes at the cost of confounded effects. However, the fraction of experimental runs selected can be done so that main effects and lower order interactions are confounded with higher order interactions. It is then assumed that interactions over a specified order are negligible so that the resulting effect estimates are attributed to the main effects or lower order interactions. Catalogs of pre-defined fractional factorial designs can be found in texts such as Montgomery (2005), Box (1978) and others, and provide experimenters with a selection of 2^{k-p} designs where main effects and lower order interactions are confounded with the highest order interaction effects possible for a particular k and p.

Although the categories of two-level factorial and fractional factorial designs reduce the number of experimental runs by a large factor, they require a total number of experimental runs that is a power-of-2. Plackett and Burman (1946) introduced two-level orthogonal experimental designs known as Plackett-Burman designs that require a total number of experimental runs that is a multiple-of-4. If the objective is to estimate main effects Plackett-Burman designs may accomplish this with fewer experimental runs than a two-level fractional factorial design. The downside of this category of designs is a complex confounding (or alias) structure between main effects and interactions. If it is reasonable to assume that all the interaction effects are negligible then the Plackett-Burman designs are very efficient for screening.

“Optimal Designs” are a class of two-level experimental designs that are optimal for any particular number of experimental runs with respect to some statistical criterion. Some of these statistical criteria such as D-Optimality seek to minimize the variance of the estimated coefficients of design variables and have been applied to screening experiments with a large number of factors. Other optimality criteria such as I-Optimality seek to minimize

the average prediction variance of the model over the design space, and for this reason are used in response surface optimization where the focus is on prediction.

There has been a large amount of optimal experimental design research. Optimal experimental designs are discussed in Goos and Jones (2011), Atkinson (2007), and Goos (2002), and contain many additional references. These references also present information about various heuristic and non-heuristic algorithms developed to computationally construct experimental designs based on a specified optimality criterion. Some of these algorithms are used in advanced computer software packages such as JMP (SAS Institute Inc 1989-2007).

Optimal experimental designs are generally known to reduce the experimental cost by obtaining similar "quality" estimates with fewer experimental runs. A non-optimal design requires a greater number of experimental runs to estimate effects with the same precision as an optimal design. However, these approaches do not explicitly consider experimental cost and time as constraints when constructing a design. The experimental designs initially considered throughout this research are balanced full factorial, and therefore D-Optimal. Full factorial designs are balanced designs with at least one experimental run for each possible treatment combination.

2.2.3 Related Research

Many experiments involve factors with level changes that are hard or expensive to implement. The constraints imposed by limited resources demand modifications to a completely randomized arrangement. Federer and King (2007) discuss the use of restricted randomization as a mean to reduce the overall cost of experiments and explain how restricting the order of runs forces the experimenter to have experimental units with different sizes at different levels of randomization. This opens up a class of more complicated designs such as variations of split-plot, split-block, and augmented split-plot designs. Simpson, Kowalski, and Landman (2004) use an augmented split-plot design for screening purposes in an automobile wind tunnel testing scenario to obtain the correct error terms (required with randomization restrictions) when practical limitations do not allow repeating the experiment with additional factors and/or factor levels. Jones and Goos

(2009) present the use of split-split plot designs in D-optimal factorial designs. Paniagua-Quiñones and Box (2007) propose the strip-strip-block design to reduce experimental costs in multi-staged industrial processes.

In this research, the enumeration of all possible experimental design structures for an experiment will be conducted. This enumeration will include many different variations of "split plot" designs that are considered in prior research. To the best of the author's knowledge there is no other published research that explicitly considers all possible balanced randomization restrictions, and the cost and time required to complete an experiment.

3. Research Methodology and Organization

In Chapter 1 it was stated that the development of the optimal experimental *system* design methodology could be could be organized into three requirements/needs. These are,

1. An experiment defined by its treatment and design structure should have a unique statistical model.
2. Mathematical models for the cost and time of an experiment as a function of its treatment and design structure.
3. An optimization procedure for searching the space of experimental design structures to find feasible experimental designs that make the best use of limited experimental resources for a given treatment structure.

Each of these requirements/needs have been met and the results for each will be presented in separate chapters. The separation of chapters is discussed in the following section.

3.1 Organization of Chapters

The results for requirement/need three are presented first in Chapter 4 since the description and notation of experimental design structures makes subsequent chapters easier to understand. In Chapter 4 a recursive procedure to enumerate all the possible design structures for a given treatment structure is developed and explained. Since the various design structures can be enumerated (for most practically sized experiments) enumeration constitutes the optimization procedure needed.

Chapter 5 presents the development of linear statistical models for experiments that represent the treatment and design structure. Each different experiment will now be represented with a unique statistical model so that the specifics of the ANOVA of experimental results, which can be obtained from the statistical model can be determined algorithmically.

Chapter 6 presents the algorithms used to determine how the ANOVA of experimental results should be conducted. Given a statistical model the test statistics for all “testable” factors and interactions will be determined. For those factors and interactions that are not testable, test statistics for conservative tests may be determined.

In Chapter 7 the mathematical models for calculating the expected time and expected cost of running an experiment are developed. The time and cost to complete all experiments in a design is a function of various cost and time parameters and the experimental treatment and design structures. The costs and time are computed with a recursive procedure developed in Chapter 7.

Chapter 8 will present the mathematical procedure utilized to make comparative evaluations of the statistical performance of different experimental design structures for a given treatment structure. This will then constitute the objective function in the optimization procedure to algorithmically find optimal experimental system designs.

The final chapter (Chapter 9) will present case studies of the optimal experimental *system* design methodology.

Before presenting the research details a starting point will be established for the development of an optimal experimental *system* design methodology by putting limits on the scope of the problem addressed. Additionally, relevant overarching assumptions with respect to the experimental environment will be clarified

3.2 Scope and Assumptions

To initiate this research, the specific type of experiments considered will be limited to balanced full factorial experiments. The hypothetical fuel consumption experiment presented earlier is an example of a balanced full factorial experiment. Additionally, the following assumptions will be made.

- Standard experimental procedures have been developed and will be utilized to conduct experiments. The time and cost of running experiments can be accurately estimated, as can the cost and time incurred when changing levels of experimental factors.
- Experiments are conducted one at a time in a pre-determined order.
- Differences in the experimental units (the smallest units to which a treatment is applied and independently observed) are negligible and considered part of the experimental error.

- Due to the presence of unknown and uncontrollable nuisance factors, the design structure distinguishes experiments. This excludes computer based experimentation where experimental order does not matter.
- The experimental factors and response are known. The factors are fixed and the response is a continuous quantitative variable.
- Time and cost constraints are known. Even if one or both of these are not known precisely, approximate values can be established and then modified to examine the effect on the resulting experimental designs.

4. Enumeration of Design Structures

The first step to find a set of most efficient and most statistically effective design structures for a given treatment structure, is to enumerate and evaluate all possible design structures. A design structure establishes the experimental run orders that are feasible. A fully randomized design structure is one where all possible experimental run orders are feasible. Other design structures are defined by randomization restrictions, which reduce the feasible experimental run orders to specific subsets of all possible experimental run orders. The design structures explored when searching for a set of optimal design structures, are "balanced" design structures (all treatment structures are also balanced). In a balanced design structure randomization restrictions partition factor levels into equal size "blocks". For example, if a factor has four levels then randomization restrictions with respect to this factor can partition the factor levels into four blocks corresponding to each factor level, or two blocks that each correspond to two factor levels.

Since a design structure applies to a specific treatment structure, notation will be defined for treatment structures to make the presentation more efficient. This is presented next.

4.1 Treatment Structure Representation

A treatment structure is defined by the following information:

- Factors,
- Factor type (fixed or random),
- Factor nesting,
- Number of levels for each factor,
- Number of replicates of each treatment combination.

The notation for a treatment structure will be a vector of elements, with each element consisting of a factor identifier and the number of factor levels separated by a colon. For example,

Factor name: *number of levels.*

To distinguish between fixed and random factors, the name of a random factor is put in parenthesis. For example,

(Factor name): *number of levels.*

A nested factor is shown with the factor/interaction it is nested within separated by an arrow. For example, factor B is a fixed factor that is nested within factor A,

B → A: *number of levels of factor B.*

The last element will be the replicates which will be shown as a factor:

(R): *number of replicates of each treatment combination*

Since replicates are a random factor, R is always shown within parenthesis. Some complete examples are presented next.

For an experiment with two fixed factors A and B with two levels and a random factor C with 4 levels, and 3 replicates per treatment combination, the treatment structure is expressed as

[A: 2, B: 2, (C): 4, (R): 3].

If factor C is nested within factor A, this treatment structure will be expressed as

[A: 2, B: 2, (C) → A: 4, (R): 3].

4.2 A Fully Randomized Design

Consider an experiment that has two factors, each with four levels, with two replicates per treatment combination. The treatment structure is specified as [A: 4, B: 4, (R): 2].

There are thirty-two total experimental runs. To specify an experimental run order for a fully randomized design structure, each experimental run can be assigned a random integer between one and thirty-two whose value is the run order. The underlined numbers in Figure 4.1. show an example of fully randomized run order for this experiment.

		Factor A			
		1	2	3	4
Factor B	1	Rep1 *5* Rep2 *28*	Rep1 *18* Rep2 *25*	Rep1 *10* Rep2 *30*	Rep1 *12* Rep2 *27*
	2	Rep1 *8* Rep2 *31*	Rep1 *16* Rep2 *21*	Rep1 *2* Rep2 *22*	Rep1 *15* Rep2 *19*
	3	Rep1 *1* Rep2 *23*	Rep1 *6* Rep2 *32*	Rep1 *7* Rep2 *20*	Rep1 *4* Rep2 *9*
	4	Rep1 *14* Rep2 *29*	Rep1 *3* Rep2 *26*	Rep1 *13* Rep2 *24*	Rep1 *11* Rep2 *17*

Figure 4.1. Design structure 1. A fully randomized design structure

4.3 Defining Randomization Restrictions

Enumeration of all possible design structures must consider all the parameters that are related to the specification of a randomization restriction, which in turn correspond to a specific design structure. The parameters can then be set to different values to generate different design structures.

4.3.1 The Restricted Factors - The First Parameter

The first parameter that specifies a design structure is the list of factors that have been restricted. In a fully randomized design structure this list is empty. Note that replicates can also be considered a factor and can be included in the list of restricted factors.

For example, suppose in the prior experiment a randomization on factor A is utilized to reduce experimental costs as shown in Figure 4.2.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 4.2. Design structure 2. Factor A is fully restricted

In this design structure, the experimenter chooses one of the four levels of factor A randomly and runs all eight experiments with that factor A level in a fully randomized order. Next another level of factor A is selected randomly and all eight experiments with that factor A level are run in a fully randomized order. This continues until all treatment combinations have been run twice.

Note that in this design structure the only restricted factor is A. However, the experimenter can choose to restrict the randomization on two factors, and, there are multiple ways that this can be done. One way is to restrict the replicates and factor A "at the same time" (to be discussed later). This design is depicted in Figure 4.3.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 4.3. Design structure 3. Randomization is restricted on replicates and factor A at the same time.

In this design, one of the eight thicker boxes is randomly selected, and then the four experiments within that box are run in random order. This continues until all thirty-two experiments are run.

4.3.2 The Randomization Restriction Size – The Second Parameter

The second parameter that specifies a design structure is the size of the randomization restriction on a factor. The randomization restriction size on a factor is defined as the number of factor levels that can be randomly selected within the restricted set of experimental runs. For example, in the last two design structures, the size of the factor A randomization restriction is one since all experiments within a "box" of restricted experimental runs are at the same factor A level. Randomization restriction sizes for a factor can be greater than one as long as this size evenly divides the total number of factor levels (i.e., a balanced randomization restriction results).

Figure 4.4 shows a design structure with a randomization restriction size of one on replicates and two on factor A. These restrictions occur at the same time.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 4.4. Design structure 4. Randomization is restricted on replicates and every two levels of factor A at the same time.

In this design structure one of the four bold boxes is selected randomly (although in this case replicate 1 for the same factor A levels precedes replicate 2), and then all experimental runs within that box are run in a fully randomized order. It is a less restricted design structure than the prior one (Figure 4.3).

4.3.3 The Randomization Restriction Layer – The Third Parameter

The third and final parameter needed to fully specify a design structure is the layer of randomization restrictions when there is more than one restricted factor. In the prior examples, both factor A and replicates were restricted at the "same time", which means that any combination of the restricted factors is randomly selected, and then all

experimental runs within the selected factor values are run in random order. If two or more factors are restricted at the same time it will be said that the randomization restrictions are at the same layer. In the prior examples, the randomization restrictions on replicates and factor A are both at layer one.

Additional design structures result from changing the randomization restriction order. Figure 4.5 shows a design structure with a randomization restriction of size one on replicates and factor A as in a prior example. However, in this example, the restriction on replicates is at layer one, and the restriction on factor A is at layer two.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 4.5. Design Structure 5. First, the randomization is restricted for replicates. Next, the randomization is restricted on factor A.

In this design structure, a replicate is selected first. All factor A level experimental runs at this replicate level must be completed before moving on to replicate 2. The order of completing the experiments within a replicate level are restricted by randomly selecting a factor A level and then completing all experiments at this factor A level in random order. This design structure is referred to as a "split-plot" experiment in Montgomery (2013).

In the next example, the randomization restriction size on factor A is two at layer one, and the randomization restriction size on factor B is two at layer two. This design structure is shown in Figure 4.6. In this design structure, a solid bordered box is selected randomly, which represents the factor A randomization restriction of size two. It is selected first because the factor A restriction is at layer one. After this selection, one of the two smaller boxes within a larger box represents the factor B randomization restriction at layer two and is selected randomly. All the experiments within that box are run in random order. After this, all of the experiments within the other smaller box are run in random order. After all

the 16 experimental runs within one solid bordered box are run, the experimental runs in the second box will be completed in a similar manner.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 4.6. Design Structure 6. The randomization is restricted on every two levels of factor A at the first layer. At the second layer, the randomization is restricted for every two levels of factor B.

4.4 Design Structures Representation

Notation, based on the three parameters described, will be developed to uniquely specify a balanced design structure. A design structure will be represented with a "set" of expressions, where each expression specifies a randomization restriction. A single such expression is,

$$(Restricted\ Factor,\ Restriction\ Size,\ Restriction\ Layer).$$

A design structure is specified by an expression for each factor with a randomization restriction. The set of such expressions for a completely randomized design structure is the empty set.

Below, the design structures in prior examples are shown with their notations. These design structures are applied to the same treatment structure. This treatment structure consists of factor A with 4 levels, factor B with 4 levels, and 2 replicates. The treatment structure notation is [A: 4, B: 4, (R): 2].

- Design structure 1 - Figure 4.1
  - {}
- Design structure 2 - Figure 4.2
  - {(Factor: A, Size = 1, Layer = 1)} = {(A,1,1)}

- Design structure 3 - Figure 4.3
  - {(Factor: A, Size = 1, Layer = 1), (Factor: R, Size =1, Layer =1)} = {(A,1,1), (R,1,1)}
- Design structure 4 - Figure 4.4
  - {(Factor: A, Size = 2, Layer = 1), (Factor: R, Size =1, Layer =1)} = {(A,2,1), (R,1,1)}
- Design structure 5 - Figure 4.5
  - {(Factor: R, Size = 1, Layer = 1), Factor: A, Size =1, Layer =2)} = {(R,1,1), (A,1,2)}
- Design structure 6 - Figure 4.6
  - {(Factor: A, Size = 2, Layer = 1), (Factor: B, Size =2, Layer =2)} = {(A,2,1), (B,2,2)}

4.5 Enumeration of Design Structures

The procedure for enumerating all possible balanced design structures is based on the following fact. *At each design structure layer, the randomization restrictions partition a treatment structure into multiple smaller equivalent treatment structures.*

For example, consider the [A: 4, B: 4, (R): 2] treatment structure with the design structure {(A,2,1), (R,1,1), (B,2,2)} (Figure 4.7). This design structure is defined as

- {Factor: A, Size = 2, Layer =1}
- {Factor: R, Size = 1, Layer =1}
- {Factor: B, Size = 2, Layer =2}

`		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 4.7. A two-layered design structure.

In this design structure, there are two layers of randomization restrictions. The first layer includes restrictions on factor A and the replicates. These randomization restrictions

partition the original treatment structure into four treatment structures of [A: 2, B: 4, (R): 1]. The second layer randomization restrictions are applied to the four “boxes” with the smaller [A: 2, B: 4, (R): 1] treatment structure. This randomization restriction, which is on factor B, partitions the [A: 2, B: 4, (R): 1] treatment structure into two smaller [A: 2, B: 2, (R): 1] treatment structures. Since there are no more layers of randomization restrictions, the experiments within each of the eight [A: 2, B: 2, (R): 1] treatment structures are run in random order.

The process of applying the randomization restrictions within a design structure at a lower layer to results from the prior layer, suggests a recursive procedure. At each iteration of this procedure, a new layer of randomization restrictions is applied to the treatment structures generated from the prior randomization restriction layer. This continues until all randomization restriction layers have been applied as indicated by the design structure parameters (factors, restriction sizes, and restriction layers).

Starting with original treatment structure, all possible factors and factor combinations restricted at the first layer are considered. Once a factor is selected to be restricted, all possible restriction sizes for that factor are considered. The restriction sizes for a factor with L levels is the set of divisors of L excluding the L itself. For example, for a factor with 6 levels, the possible restriction sizes are 1, 2, and 3 as shown in Figure 4.8.

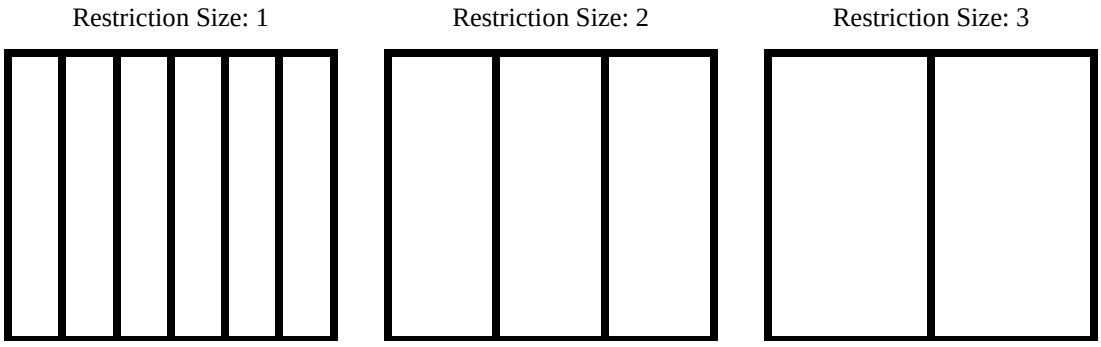


Figure 4.8. Possible randomization restriction sizes for a factor with 6 levels

At each step of the procedure, the resulting design structure is recorded and then the procedure moves to the next randomization restriction layer. Treatment structures are partitioned until they can no longer be partitioned into smaller treatment structures.

Pseudocode for this procedure is shown in Figure 4.10. The procedure starts with an empty set of design structures and adds newly created design structures to the set at each iteration, under the condition that it does not already exist in the set.

The “minLayer” and “maxLayer” variables assigned to each factor, are the minimum and maximum possible restriction layer for each factor. They are continuously updated at each iteration according to the status of current design and treatment structures.

```
EnumerateDesignStrutures( [factors] , str, [designStructures])
    For each factor i in [factors]                          //all possible factors
        For each m in {multiples of L_i} \ L_i              //all possible sizes
            For each o in range [ minLayer_i , maxLayer_i ] //all possible layers
                str += "factorName_i , m , o"               // for example "A,2,1"
                if str does not exist in [designStructures]
                        add str to [designStructures]
                str += ";"
                L_i = m
                minLayer_i = o+1
                maxLayer_i = o+1
                For all other factors
                    minLayer = o
                    maxLayer = o+1
                EnumerateDesignStrutures( [factors] , str)
```

Figure 4.9. Pseudocode for the recursive procedure of enumerating design structures, given a treatment structure (part one)

```
***//Initialization//***

For each factor i [factors]

    minLayer = 1
    maxLayer = 1

str = " ";

EnumerateDesignStrutures( [factors] , str, [])
```

Figure 4.10. Pseudocode for the recursive procedure of enumerating design structures, given a treatment structure (part 2)

It was stated that only balanced randomization restrictions will be considered. To clarify this, a design structure such as shown below in Figure 10 will be excluded when enumerating possible design structures.

		Factor A				
		1	2	3	4	5
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Figure 4.11. An unbalanced design structure

The main reason for this exclusion is to avoid statistical analysis complications with these unbalanced restrictions. The statistical analyses of different design structures are discussed in later sections.

5. One Experiment, One Statistical Model

In this section, the statistical model for an experiment will be developed that explicitly represents the design structure defined by randomization restrictions, as well as the treatment structure. The "standard" linear statistical model used to represent an experiment as found in Montgomery (2013), Keuhl (2000), Barrentine (1999), and others represents only the experimental treatment structure (however nested factors which may also be considered as part of the design structure are represented in these statistical models). The statistical model developed here is a generalization of the statistical model for an experiment found in the "No Name" approach to experimental design developed by Lorenzen and Anderson (1993), Anderson (1970), Anderson and McClean (1974).

Once developed, the statistical model will serve as the foundation for the analysis of variance (ANOVA). This statistical model will define the effects and interactions that may be present in an experiment, and through existing algorithms the statistical model specifies their degrees of freedom, and expected mean squares. This information can then be used to develop the test statistics and the nature of the statistical tests (exact, conservative) available for each effect and interaction.

In the previous section, the process of generating all possible balanced design structures for a given treatment structure was presented. The possible design structures were generated through a recursive enumeration procedure where all possible balanced randomization restrictions are enumerated. A design structure applied to a particular treatment structure can be uniquely defined by the following three parameters.

- The factor or factors on which randomization restrictions are applied,
- The randomization restriction size for each factor,
- The restriction layer for each factor with a randomization restriction.

The design structure defined by these parameters determines which main effects or interaction effects are confounded with the effect of a randomization restriction.

To demonstrate the statistical model development, examples of different design structures for a given treatment structure and the step by step development of the corresponding statistical models will be presented.

5.1 Statistical Model Notation

The notation used to present the statistical model in this research is similar to the notation used by Lorenzen and Anderson in "Design of Experiments, A No-name approach" (1991). Factors and interactions are represented with upper-case Latin letters. The overall mean is represented with the Greek letter μ and the response variable is shown with the Latin letter "Y".

For instance, for the treatment structure [A: *I*, B: *J*, (R): *K*], the corresponding linear effect model is shown as below.

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk}$$

Where the subscript for each factor effect has a range from 1 to the number of level of that factor. In the prior model, the subscript are ranges are,

$$\begin{cases} i = 1, \dots, I \\ j = 1, \dots, J \\ k = 1, \dots, K \end{cases}$$

As it will be explained later in this section, it is sometimes preferred to include replicates, or R, as a factor in the model. In that case, the statistical model is presented as follows.

$$Y_{ijk} = \mu + R_k + A_i + AR_{ik} + B_j + AB_{ij} + BR_{jk} + ABR_{ijk} + \varepsilon_{ijk}$$

For models with nested factors, the nested relationship between factors is shown with subscripts of the nested factors followed by subscripts of the factor they are nested within in parentheses. For example, for the treatment structure of [A: *I*, B $\rightarrow$ *A: J*, (R): *K*] in which factor B is nested within factor A, the statistical model is

$$Y_{ijk} = \mu + R_k + A_i + AR_{ik} + B_{j(i)} + BR_{j(i)k} + \varepsilon_{ijk}$$

The subscript *j(i)* shows that factor B is nested within the levels of factor A.

5.2 Statistical Model Development

The construction of the statistical model starts with specifying terms for main effects, followed by specifying the terms for interaction effects, and concluding with specifying

the randomization restriction model terms that determine the design structure. The generalization of the Lorenzen and Anderson's statistical models is such that any type of balanced randomization restriction can be accommodated. The statistical models presented by Lorezen and Anderson have only been developed for randomization restrictions that are "*fully restrictive*" (a fully restricted factor has a restriction size of one), and only consider a randomization restriction for a single factor at each layer. Figure 5.1 shows an example of both fully restrictive and partially restrictive designs. The diagram on the left demonstrates a fully restrictive randomization restriction on factor A, while the diagram on the right represents a partially restrictive randomization restriction on factor A.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 5.1. The diagram on the left shows a fully restrictive randomization restriction on factor A. The diagram on the right shows a partially restrictive randomization restriction on factor A

The remainder of this section will go over a brief review of the Lorenzen and Anderson's method for representing randomization restrictions in the statistical model as well as the notation used for this term in the following statistical models in this research. This will be followed by a description of the statistical models for experiments with no randomization restrictions (a completely randomized experiment), and experiments with fully restricted factors. For these types of experiments, the Lorenzen and Anderson's method is applicable so that statistical models describing the treatment and design structure can be written. The generalization of the Lorenzen and Anderson's method will then be presented for developing statistical models of experiments with more complex randomization restrictions.

5.3 Randomization restriction terms

Lorenzen and Anderson (1993) developed the concept of adding a randomization restriction term to the statistical model to reflect the existence of randomization restrictions in the experiment. The randomization restriction term is a random factor with zero degrees of freedom, and is nested within the corresponding restricted factor or interaction. Inclusion of this random factor makes it possible to uniquely depict the experimental design structure in the statistical model.

Design structures that involve randomization restrictions restrict the experimental order with respect to a factor at each randomization restriction layer. Any effect of the restricted factor will be completely confounded with the effect of the randomization restriction factor (e.g., blocks). Including the randomization restriction terms in the statistical model results in "one experiment, one statistical model". Each experiment will have a unique statistical model from which the specifics of the ANOVA will be determined.

To identify all the randomization restriction terms needed for a given design structure, all the restricted factors and interactions must be identified. For each restricted factor or interaction, a new random factor will be added to the model (nested within the factor or interaction) to reflect the corresponding randomization restriction. The process of identifying the restricted factors and interactions proceeds layer by layer. At each randomization restriction layer, all restricted factors in the layer and prior restriction layers are identified. If there are no prior layers, the randomization restriction term will be directly nested within the restricted factor. If there are prior restriction layers, the randomization restriction term will be nested within the corresponding interaction between the factors in prior layers and the restricted factor being considered. In a statistical model, a randomization restriction term corresponding to a restricted factor or interaction, is shown as "*restricted-term-name_r*".

For example, consider the fully restricted design structure 2 in Figure 5.2. The treatment structure is represented as [A: 4, B: 4, (R): 2], and the design structure is (Factor: A, Size: 1, Layer: 1)}. This design structure only includes one randomization restriction layer and it involves factor A. Therefore, there is only one restricted factor for which a randomization

restriction term needs to be added to the model. The following random factor will be added to the model.

- *A_r* nested within *A*

5.4 Completely Randomized Designs and Designs with Fully Restricted Factors

If an experimental design structure does not involve any randomization restrictions, then the "standard" linear statistical model is taken as the statistical model for the experiment. The lack of design structure information is interpreted to mean that the design structure is a completely randomized design. If an experimental design structure includes a fully restrictive randomization restriction, then the additions to the standard linear statistical model developed in the Lorenzen and Anderson "No-Name" approach is applicable. For example, Figure 5.2 shows two different design structures for an experiment with two factors (A and B), each with 4 levels. There are 2 replicates per treatment combination. The corresponding treatment structure is represented as [A: 4, B: 4, (R): 2]. The design structure for the experiment on the left can be represented as an empty set of randomization restrictions. For the experiment on the right, the design structure is shown as a single-element set of randomization restrictions: {(Factor: A, Size: 1, Layer: 1)}.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Design structure 1

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Design structure 2

Figure 5.2. Design structure 1 is a fully randomized design structure, and design structure 2 is a fully restrictive design structure (on factor A) for the treatment structure [A: 4, B: 4, (R): 2].

The statistical model for design structure 1:

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 4 \\ k = 1, 2 \end{cases}$$

The statistical model for design structure 2:

$$Y_{ijk} = \mu + A_i + A\_r_{j(i)} + B_k + AB_{ik} + \varepsilon_{ijkl} \qquad \begin{cases} i = 1, \dots, 4 \\ j = 1 \\ k = 1, \dots, 4 \\ l = 1,2 \end{cases}$$

Figure 5.3 depicts a two-layered design structure with one randomization restriction at each layer (for the same treatment structure [A: 4, B: 4, (R): 2]). One randomization restriction is fully restrictive and applied to replicates at layer 1, and the other randomization restriction is also fully restrictive and applied to factor A at layer 2. The replicates are included in the statistical model as a random factor and there is no estimate of the error term. The design structure is given as {(Factor: R, Size: 1, Layer: 1), (Factor: A, Size: 1, Layer: 2)}.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Design structure 3

Figure 5.3. A "split-plot" design structure with notation {(R, 1, 1), (A, 1, 2)}

Therefore, there will be two random factors nested within the appropriate factor and interaction. The first layer contains one randomization restriction on the factor replicates (R) so that one randomization restriction term that is nested in R is added to the statistical model. If there were more fully restricted factors in layer one additional terms would be added in a similar manner. There is one randomization restriction in the second layer on factor A. Therefore, a randomization restriction term is added and it is nested in the interaction between A and R. If more combinations of restricted factors between layer one

and two existed, additional randomization restriction terms would be added in a similar manner. In this example, the following random effects will be added to the model:

- *R_r* nested within *R*
- *AR_r* nested within *AR*

The complete model is:

$$Y_{ijklm} = \mu + R_i + R\_r_{j(i)} + A_k + AR_{ki} + AR\_r_{l(ki)} + B_m + AB_{km} + BR_{mi} + ABR_{kmi} + \varepsilon_{ijklm}$$

$$\begin{cases} i = 1,2 \\ j = 1 \\ k = 1, \dots, 4 \\ l = 1 \\ m = 1, \dots, 4 \end{cases}$$

This is referred to as a "split-plot" design in Montgomery (2013). The two models shown in Figure 5.4. are the statistical models presented in Montgomery (2013) for the design structure 3 and a fully randomized three factor factorial design. As it can be seen, the two models are identical and show no indication of any differences between the two design structures.

Fully randomized design structure	*Split-Plot design structure*
$Y_{ijk} = \mu + R_i + A_k + AR_{ki} + B_m$ $+AB_{km} + BR_{mi} + ABR_{kmi} + \varepsilon_{ijk}$	$Y_{ijk} = \mu + R_i + A_k + AR_{ki} + B_m$ $+AB_{km} + BR_{mi} + ABR_{kmi} + \varepsilon_{ijk}$

Figure 5.4. The presentation of the identical statistical models for two different design structures in Montgomery (2013)

5.5 Designs with Partially Restricted Factors

The category of design structures that adds a level of complexity when constructing the statistical model are designs with partially restricted factors and/or interactions. The example in Figure 5.5 is a two-layered design structure specified as:

- {(Factor: R, Size: 1, Layer: 1), (Factor: A, Size: 2, Layer: 2), (Factor: B, Size: 2, Layer: 2)}

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Design structure 4

Figure 5.5. A partially restricted design structure with notation {(R, 1, 1), (A, 2, 2), (B, 2, 2)}

This design structure includes second layer partial restrictions on factors A and B. The inclusion of these partial restrictions on factors A and B in the statistical model requires a modification to the experimental treatment structure. Partially restricted factors will be modeled as two nested factors. In this example the following nested structure of factors A and B results.

- $A^{(2)}$ with two levels, and $A^{(1)}$ nested within $A^{(2)}$ with two levels.
- $B^{(2)}$ with two levels, and $B^{(1)}$ nested within $B^{(2)}$ with two levels.

In this example each partially restricted factor A is modeled as two factors $A^{(1)}$ and $A^{(2)}$ with $A^{(1)}$ nested within $A^{(2)}$. A similar modeling structure is used for partially restricted factor B. The modified treatment structure of the experiment is written as [$A^{(2)}$: 2, $A^{(1)} \rightarrow A^{(2)}$: 2, $B^{(2)}$: 2, $B^{(1)} \rightarrow B^{(2)}$: 2, (R): 2].

With this modified treatment structure the design structure given by {(Factor: R, Size: 1, Layer: 1), (Factor: A, Size: 2, Layer: 2), (Factor: B, Size: 2, Layer: 2)} can be represented in the statistical model. The first layer restricted factor is the replicates factor *R* and this factor is fully restricted and thus a randomization restriction term (nested in R) is added to the statistical model. In the second layer the two partially restricted factors A and B with restriction size of two are represented as restrictions on factors $A^{(2)}$ and $B^{(2)}$. Since there is a randomization restriction on replicates in layer one, the randomization restrictions on factors $A^{(2)}$ and $B^{(2)}$ occur within each level of R. Therefore a randomization restriction term nested within the $A^{(2)}B^{(2)}R$ interaction is added to the statistical model. In summary, the following randomization restriction terms are added to the model.

- R_r nested within R
- $A^{(2)}B^{(2)}R\_r$ nested within $A^{(2)}B^{(2)}R$

The complete model is:

$$Y_{ijklmn} = \mu + R_i + R_{r_{j(i)}} + A^{(2)}{}_k + A^{(1)}{}_{l(k)} + B^{(2)}{}_m + B^{(1)}{}_{n(m)} + A^{(2)}B^{(2)}{}_{km} + A^{(2)}R_{ki} + RB^{(2)}{}_{im}$$
$$+A^{(2)}B^{(2)}R_{ikm} + A^{(2)}B^{(2)}R_{r_{o(ikm)}} + A^{(1)}B^{(2)}{}_{l(k)m} + A^{(2)}B^{(1)}{}_{kn(m)} + A^{(1)}B^{(2)}R_{l(k)mi}$$
$$+A^{(2)}B^{(1)}R_{kn(m)i} + A^{(1)}B^{(1)}{}_{l(k)n(m)} + A^{(1)}B^{(1)}R_{l(k)n(m)i} + \varepsilon_{ijklmno}$$

$$\begin{cases} i = 1,2 \\ j = 1 \\ k = 1,2 \\ l = 1,2 \\ m = 1,2 \\ n = 1,2 \end{cases}$$

5.6 Statistical Model Construction from the Specified Treatment and Design Structure

5.6.1 Modified Treatment Structures for Partially Restricted Factors

The modification of treatments structures so that partial factor restrictions can be represented in a statistical model can be generalized to factors with multiple layers of partial restrictions. For example, a single factor experiment (factor A) with eight levels can be partially restricted in two layers as indicated by the design structure.

- {(Factor: A, Size: 4, Layer: 1), (Factor: A, Size: 2, Layer: 2)}

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 5.6. A single factor experiment with 2 layers of randomization restrictions

In this experiment one group of four levels of factor A must be selected, and all experiments within the selected group of four levels must be completed before any experiment from the second group of four levels is conducted. However, in this experiment each group of four

levels is separated into two groups of two factor levels where the second layer restriction implies that all experiments in one group of two levels must be completed before any experiments in the second group of two levels is conducted. The modification of the treatment structure results in the following three nested factors.

- $A^{(4)}$ with two levels
- $A^{(2)}$ with two levels, nested within $A^{(4)}$
- $A^{(1)}$ with two levels, nested within $A^{(2)}$

5.6.2 Interactions

With all model factors identified so that randomization restrictions can be represented in the statistical model, listing the interactions is a straightforward process. For every factor pair, except for factors that are nested within each other, a two-factor interaction results. Note that these two-factor interactions are said to be *nested within* the two corresponding main effects.

Similarly, three-factor interactions result from combinations of three factors none of the three factors is nested within another. These interactions are nested within all the corresponding interacting effects. This process continues until all the interactions have been identified.

For design structures 1 and 2 shown in Figure 5.2, *A* and *B* are the two factors in the statistical model. In design structure 3 (Figure 5.3) factor *R* separately identified and included in the model due to the randomization restriction on the replicates. Although possibly redundant, for consistency replicates is identified as a random factor in all models.

For the design structures 1, 2, and 3 (Figure 5.2 and Figure 5.3) the factors and interactions are as follows.

- Factors:
  - *A*
  - *B*
  - *R*
- Two-factor interactions:
  - *AB* nested within *A* and *B*

    - AR nested within A and R
    - BR nested within B and R

- Three-factor interaction effects:
    - ABR nested within A, B, R, AB, AR, BR

For design structure 4 (Figure 5.5) the factors and interactions are as follows.

- Factors:
    - $A^{(2)}$
    - $A^{(1)}$ nested within $A^{(2)}$
    - $B^{(2)}$
    - $B^{(1)}$ nested within $B^{(2))}$
    - R

- Two-factor interaction:
    - $A^{(2)}B^{(2)}$ nested within $A^{(2)}$ and $B^{(2)}$
    - $A^{(2)}R$ nested within $A^{(2)}$ and R
    - $B^{(2)}R$ nested within $B^{(2)}$ and R
    - $A^{(2)}B^{(1)}$ nested within $B^{(1)}$, $A^{(2)}$, and $B^{(2)}$ (or $B^{(1)}$ and $A^{(2)}B^{(2)}$)
    - $A^{(1)}B^{(2)}$ nested within $A^{(1)}$, $A^{(2)}$, and $B^{(2)}$ (or $A^{(1)}$ and $A^{(2)}B^{(2)}$)
    - $A^{(1)}R$ nested within $A^{(1)}$, $A^{(2)}$, and R (or $A^{(1)}$ and $A^{(2)}R$)
    - $B^{(1)}R$ nested within $B^{(1)}$, $B^{(2)}$, and R (or $B^{(1)}$ and $B^{(2)}R$)
    - $A^{(1)}B^{(1)}$ nested within $A^{(1)}$, $A^{(2)}$, $B^{(1)}$, and $B^{(2)}$ (or $A^{(2)}B^{(1)}$ and $A^{(1)}B^{(2)}$)

- Three-factor interaction:
    - $A^{(2)}B^{(2)}R$ nested within $A^{(2)}$, $B^{(2)}$, and R (or $A^{(2)}B^{(2)}$, $A^{(2)}R$, and $B^{(2)}R$)
    - $A^{(2)}B^{(1)}R$ nested within $A^{(2)}$, $B^{(1)}$, $B^{(2)}$, and R (or $A^{(2)}B^{(1)}$, $B^{(1)}R$, and $A^{(2)}B^{(2)}R$)
    - $A^{(1)}B^{(2)}R$ nested within $A^{(1)}$, $A^{(2)}$, $B^{(2)}$, and R (or $A^{(1)}B^{(2)}$, $A^{(1)}R$, and $A^{(2)}B^{(2)}R$)
    - $A^{(1)}B^{(1)}R$ nested within $A^{(1)}$, $A^{(2)}$, $B^{(1)}$, $B^{(2)}$, and R (or $A^{(2)}B^{(2)}R$, $A^{(2)}B^{(1)}R$, and $A^{(1)}B^{(2)}R$)

The use of *nesting* terminology for the interaction effects might seem to be an over complication but it will become useful later in illustrating the factor structure diagrams of experiments.

The treatment structures for each experimental design includes R as a random effect. Some authors consider replicates and interactions with replicates as noise and aggregate all of these terms into the experimental error terms. In the approach taken here replicates are included in the statistical model for purposes of identifying any randomization restrictions. As will be discussed in the next section, randomization restrictions of replicates result in

multiple error terms in an ANOVA. A Split-Plot design is an example of an experimental design with multiple error terms, although in Montgomery (2013) the randomization restriction and confounding of replicates with a randomization restriction factor is not explicitly recognized.

6. Analysis of Design Structures

In the previous chapter the development of unique linear statistical models for experiments with particular design and treatment structures was discussed. It was shown how the inclusion of randomization restriction terms in the statistical model can distinguish fully randomized, partially restricted, and fully restricted design structures. It was also shown how a modification of treatment structure can be used to represent partially restrictive randomization restrictions in the statistical model.

In the first part of this chapter, the derivation of different components needed for an analysis of variance (ANOVA) is explained. The second part of this chapter goes over various examples describing the analysis of different design structures using ANOVA.

6.1 The Analysis of Variance (ANOVA)

The analysis of variance is the general approach to test for statistically significant differences among different factor level and interaction mean responses. ANOVA is valid under the following assumptions.

- The dependent, or response, variable is normally distributed for each level of the factor being tested (i.e., the error term in the statistical model is normal).
- The error term variance is constant.
- The observations, or measurements, are independent from one another.

The normality assumption of the error term in the statistical model gives rise to the F-distribution as the sampling distribution of ANOVA test statistics (ratios of mean squares). However, in the absence of the normality assumption test statistics remain ratios of sample variance (the mean squares) and thus still provide information on significant factors and interactions. Additionally, the expected values of the mean squares that are used to determine the mean square ratios used as test statistics (for the significance of specific factors and interactions) does not depend on the normality assumption. The ANOVA also does not include any assumptions about a specific relationship between the response variable and the factor levels and is thus generally applicable to many types of experimental data. An experimental design principle that is directly related to the assumption of constant error term variance is that of randomization and is a fundamental concept in this research.

Two fundamental components of an ANOVA, which can be derived from the statistical model are,

- The degrees of freedom for factors and interactions,
- The expected value of mean squares for factors and interactions.

The derivation of these two components from the statistical model is presented in the next two sections.

6.1.1 Degrees of Freedom

The number of degrees of freedom is defined as the number of independent observations used to calculate a statistic. In the analysis of variance, these statistics are the sum of squares values calculated to measure the amount of variability of the response that is explained by a particular source of variation.

The degrees of freedom for main effects is calculated by subtracting one from the number of levels. For nested factors, the number of degrees of freedom is calculated by multiplying the number of levels minus with the number of levels of the factor they are nested within. For an interaction the number of degrees of freedom is calculated as the product of the degrees of freedoms of each interacting factor. Since the assumed interactions and number of levels for each factor are parameters of the statistical model, all needed degrees of freedom can be computed given the model.

6.1.2 Expected Mean Squares

The expected value of the mean squares provide a breakdown of the fixed effects and variance components sum estimated by a mean square statistic. Once the expected mean squares are derived, the test statistics and the type of statistical test available for each effect can be determined.

There are various procedures for obtaining the expected mean. All of these procedures more or less follow the same set of underlying mathematical rules found in in the literature (Montgomery 2013). All of the procedures use the information contained in the statistical

model to find the expected value for the mean squares of each factor or interaction present in the statistical model.

Hasse Diagrams, or the factor structure diagrams, are visual tools that facilitate the illustration of the treatment and design structures of an experiment. Lohr (1995) used Hasse diagrams as a mean for summarizing the structure of statistical models as well as obtaining the expected mean squares for different factors and interactions in the model. Hasse diagrams are used in this chapter to illustrate the process of obtaining the expected mean squares.

Constructing the Hasse diagrams is a two-step process. In the first step, the treatment and design structures of the experiment are used to build the factor structure diagram. A factor structure diagram is a hierarchical diagram that depicts all of the factors and interactions in the model along with nested relationships. It also shows the number of levels, degrees of freedom, and type (fixed or random) of each factor and interaction.

In the second step, the factor structure diagram is used to obtain the expected mean squares for different factors and interactions in the model. This step follows specific accepted rules for deriving the expected mean squares.

6.1.3 Factor Structure Diagrams

A factor structure diagram is constructed according to the following four-step procedure:

- Start with *M* (representing the model) at the top and *E* (representing the error) at the bottom. All other factors and interactions are then drawn between *M* and *E*. The main effects that are not nested within other effects will be crossed at the first level. The two-factor interactions and the effects that are nested within the other main effects are drawn at the second level. The three-factor interactions are placed at the third level, and so on.
- Random and mixed effects are put in parentheses.
- The number of levels of each main and interaction effect is entered as a superscript (enter 1 for *M*).
- The number of degrees of freedom of each main and interaction effect is entered as a subscript (enter 1 for *M*).

To demonstrate this procedure, consider design structure 1, shown in Figure 5.2. It is a completely randomized design for the treatment structure [A: 4, B: 4, (R): 2]. The factor structure for this experiment is shown in Figure 6.1. Note that the error term is always assumed to be random, and therefore, shown inside parenthesis.

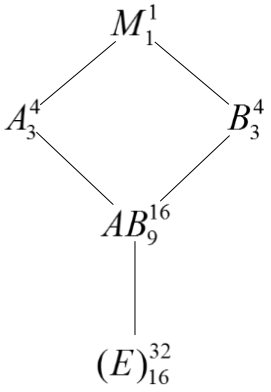


Figure 6.1. Factor structure diagram for design structure 1 (a fully randomized design).

If replicates are included as a random factor *R* in the model the factor structure diagram is more complicated as shown in Figure 6.2. Replicates will be identified as a random factor in the statistical models to facilitate modeling different types of randomization restrictions.

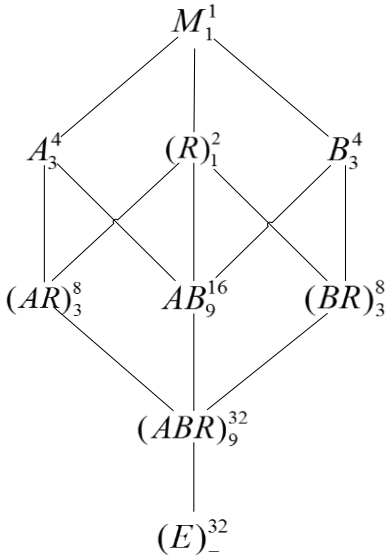


Figure 6.2. The factor structure diagram for design structure 1 (a fully randomized design) with replicates explicitly included as a random factor.

This diagram is equivalent to the prior one in the sense that aggregating all the sources of variability involving replicates (*R*, *AR*, *BR*, and *ABR*) results in the same number of degrees of freedom as the *E* term in Figure 6.1.

Figure 6.3 shows the factor structure diagram for design structure 2 shown in Figure 5.2. This design is fully restrictive on factor A.

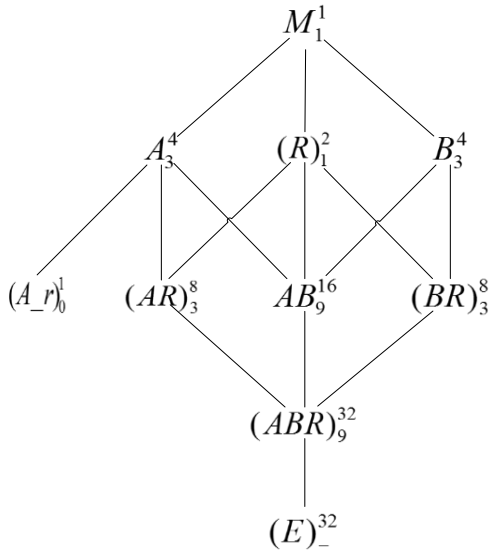


Figure 6.3. Factor structure diagram for design structure 2 represented as {(A, 1, 1)}

Figure 6.3 is very similar Figure 6.2. This similarity is because both experiments have the same treatment structure. The only difference between the two diagrams is the *A_r* random factor that is nested within *A* (with zero degrees of freedom) to model the randomization restriction in the design structure.

Figure 6.4 shows the factor structure diagram for design structure 3 in Figure 5.3. Since the experimental treatment structure is the same as the previous two experiments, the overall structure of this diagram is similar to the last two diagrams.

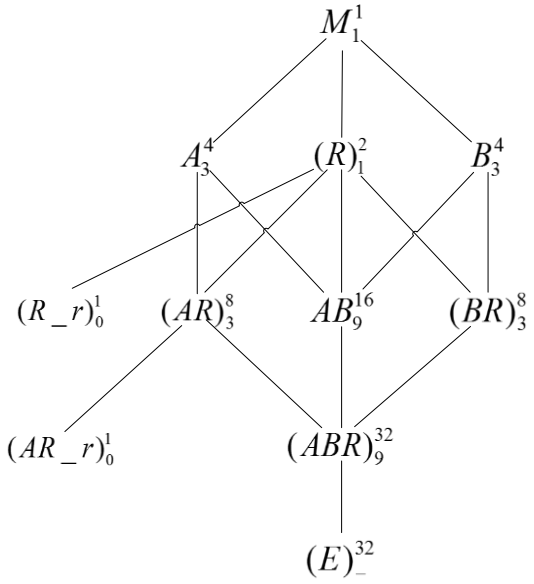


Figure 6.4. Factor structure diagram for design structure 3 represented as {(R, 1, 1), (A, 1,2)}

The factor structure diagram for design structure 4 in Figure 5.5, is depicted in Figure 6.5.

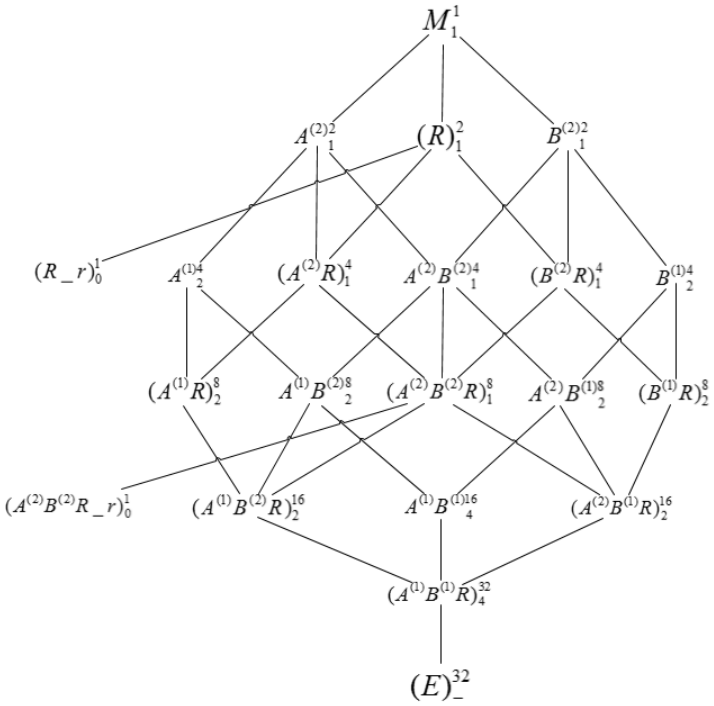


Figure 6.5. Factor structure diagram for design structure 4 represented as {(R, 1, 1), (A, 2,2), (B, 2,2)}

Due to the difference in treatment structure to model the partial randomization restrictions, Figure 6.5 is visually very different than earlier diagrams.

6.1.4 Obtaining the Expected Mean Squares

The factor structure diagrams are used to obtain the expected mean squares for different factors and interactions represented in the statistical model. For any factor *F* in an orthogonal design, this can be achieved according to the following steps:

- Step 1. If *F* is random then include
  - $\frac{Total\ number\ of\ observations}{superscript\ of\ F} * \sigma_F^2$

  If *F* is fixed then include
  - $\frac{Total\ number\ of\ observations}{superscript\ of\ F} * \Phi(F)$
- Step 2. For every random effect R below F, if every fixed factor found in *R* is also in *F*, include the term
  - $\frac{Total\ number\ of\ observations}{superscript\ of\ R} * \sigma_R^2$

Where $\Phi(F)$ is the quadratic effect function of the fixed factor *F*. For example, for factor A with L_A levels,

$$\Phi(A) = \frac{\sum_{i=1}^{L_A} \mathrm{A}_i{}^2}{L_A - 1}.$$

The notation of $\Phi(F)$ will be used through the rest of this chapter to represent the quadratic effect function for fixed factors. The next four figures show the expected mean squares for different factors and interactions in design structures 1 through 4.

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2_{AR}$
B	3	$\sigma^2 + 8*\Phi(B) + 4*\sigma^2_{BR}$
R	1	$\sigma^2 + 16*\sigma^2_{R}$
AB	9	$\sigma^2 + 2*\Phi(AB) + \sigma^2_{ABR}$
AR	3	$\sigma^2 + 4*\sigma^2_{AR}$
BR	3	$\sigma^2 + 4*\sigma^2_{BR}$
ABR	9	$\sigma^2 + \sigma^2_{ABR}$
Error	0	σ^2
Total	31	

Figure 6.6. Degrees of freedom and the expected mean squares for design structure 1

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR + 8*\sigma^2A\_r$
A_r	0	$\sigma^2 + 8*\sigma^2A\_r$
B	3	$\sigma^2 + 8*\Phi(B) + 4*\sigma^2BR$
R	1	$\sigma^2 + 16*\sigma^2R$
AB	9	$\sigma^2 + 2*\Phi(AB) + \sigma^2ABR$
AR	3	$\sigma^2 + 4*\sigma^2AR$
BR	3	$\sigma^2 + 4*\sigma^2BR$
ABR	9	$\sigma^2 + \sigma^2ABR$
Error	0	σ^2
Total	31	

Figure 6.7. Degrees of freedom and the expected mean squares for design structure 2

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR + 4*\sigma^2AR\_r$
B	3	$\sigma^2 + 8*\Phi(B) + 4*\sigma^2BR$
R	1	$\sigma^2 + 16*\sigma^2R + 16*\sigma^2R\_r + 4*\sigma^2AR\_r$
R_r	0	$\sigma^2 + 16*\sigma^2R\_r$
AB	9	$\sigma^2 + 2*\Phi(AB) + \sigma^2ABR$
AR	3	$\sigma^2 + 4*\sigma^2AR + 4*\sigma^2AR\_r$
AR_r	0	$\sigma^2 + 4*\sigma^2AR\_r$
BR	3	$\sigma^2 + 4*\sigma^2BR$
ABR	9	$\sigma^2 + \sigma^2ABR$
Error	0	σ^2
Total	31	

Figure 6.8. Degrees of freedom and the expected mean squares for design structure 3

Effect	DF	Expected Mean Square
A(2)	1	$\sigma^2 + 16*\Phi(A^{(2)}) + 8*\sigma^2_A{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
A(1)	2	$\sigma^2 + 8*\Phi(A^{(1)}) + 4*\sigma^2_A{}^{(1)}_R$
B(2)	1	$\sigma^2 + 16*\Phi(B^{(2)}) + 8*\sigma^2_B{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
B(1)	2	$\sigma^2 + 8*\Phi(B^{(1)}) + 4*\sigma^2_B{}^{(1)}_R$
R	1	$\sigma^2 + 16*\sigma^2_R + 16*\sigma^2_{R\_r} + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
R_r	0	$\sigma^2 + 16*\sigma^2_{R\_r}$
A(2)B(2)	1	$\sigma^2 + 8*\Phi(A^{(2)}B^{(2)}) + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
A(2)B(1)	2	$\sigma^2 + 4*\Phi(A^{(2)}B^{(1)}) + 2*\sigma^2_A{}^{(2)}_B{}^{(1)}_R$
A(2)R	1	$\sigma^2 + 8*\sigma^2_A{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
A(1)B(2)	2	$\sigma^2 + 4*\Phi(A^{(1)}B^{(2)}) + 2*\sigma^2_A{}^{(1)}_B{}^{(2)}_R$
A(1)B(1)	4	$\sigma^2 + 2*\Phi(A^{(1)}B^{(1)}) + \sigma^2_A{}^{(1)}_B{}^{(1)}_R$
A(1)R	2	$\sigma^2 + 4*\sigma^2_A{}^{(1)}_R$
B(2)R	1	$\sigma^2 + 8*\sigma^2_B{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
B(1)R	2	$\sigma^2 + 4*\sigma^2_B{}^{(1)}_R$
A(2)B(2)R	1	$\sigma^2 + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_R + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
A(2)B(2)R_r	0	$\sigma^2 + 4*\sigma^2_A{}^{(2)}_B{}^{(2)}_{R\_r}$
A(2)B(1)R	2	$\sigma^2 + 2*\sigma^2_A{}^{(2)}_B{}^{(1)}_R$
A(1)B(2)R	2	$\sigma^2 + 2*\sigma^2_A{}^{(1)}_B{}^{(2)}_R$
A(1)B(1)R	4	$\sigma^2 + \sigma^2_A{}^{(1)}_B{}^{(1)}_R$
Error	0	σ^2
Total	31	

Figure 6.9. Degrees of freedom and the expected mean squares for design structure 4

Some authors aggregate the variance components involving replicates (e.g. $\sigma^2(R)$, $\sigma^2(AR)$, or $\sigma^2(ABR)$) into a single error term. In other cases such as "split plot" experiments this may result in multiple "error" terms

An experiment with replicates and no randomization restrictions includes a single error term. The expected mean square of this error term only involves σ^2. Figure 6.10 shows how the error term for design structure 1, shown in Figure 5.2, is obtained by aggregating variance components involving replicates.

- Removing the variance components involving *R*.

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A)$
B	3	$\sigma^2 + 8*\Phi(B)$
R	1	σ^2
AB	9	$\sigma^2 + 2*\Phi(AB)$
AR	3	σ^2
BR	3	σ^2
ABR	9	σ^2
Error	0	σ^2
Total	31	

- Aggregating the effects with similar variance components into separate error terms.

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A)$
B	3	$\sigma^2 + 8*\Phi(B)$
AB	9	$\sigma^2 + 2*\Phi(AB)$
Error	16	σ^2
Total	31	

Figure 6.10. Aggregating the error terms for design structure 1

Another example is design structure 3 depicted in Figure 5.3. This is the "split plot" experiment in Montgomery (2013). Assuming any variance components involving

replicates are negligible the expected mean squares are shown in Figure 6.11. Note that there are three different “error terms” in the model.

- Removing the variance components involving *R*.

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR\_r$
B	3	$\sigma^2 + 8*\Phi(B)$
R	1	$\sigma^2 + 16*\sigma^2R\_r + 4*\sigma^2AR\_r$
R_r	0	$\sigma^2 + 16*\sigma^2R\_r$
AB	9	$\sigma^2 + 2*\Phi(AB)$
AR	3	$\sigma^2 + 4*\sigma^2AR\_r$
AR_r	0	$\sigma^2 + 4*\sigma^2AR\_r$
BR	3	σ^2
ABR	9	σ^2
Error	0	σ^2
Total	31	

- Aggregating the effects with similar variance components into separate error terms.

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR\_r$
B	3	$\sigma^2 + 8*\Phi(B)$
R_r	0	$\sigma^2 + 16*\sigma^2R\_r$
AB	9	$\sigma^2 + 2*\Phi(AB)$
AR_r	0	$\sigma^2 + 4*\sigma^2AR\_r$
Error1	1	$\sigma^2 + 16*\sigma^2R\_r + 4*\sigma^2AR\_r$
Error2	3	$\sigma^2 + 4*\sigma^2AR\_r$
Error3	12	σ^2
Total	31	

Figure 6.11. Aggregating the error terms for design structure 3

For most experimental designs in this research replicates will be included as a random factor and no assumption regarding the magnitude of variance components involving replicates will be made.

6.2 Examples of the Analysis Procedure to Different Experimental Designs

6.2.1 The Analysis of Completely Randomized Design Structures (CRD)

Consider an experiment with the treatment structure of [A: 4, B: 4, (R): 2] and a completely randomized design structure. The experiment diagram, statistical model, and the expected mean squares are shown in Figure 6.12.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Effect	**DF**	**Expected Mean Square**
A	3	$\sigma^2 + 8*\Phi(A)$
B	3	$\sigma^2 + 8*\Phi(B)$
AB	9	$\sigma^2 + 2*\Phi(AB)$
Error	16	σ^2
Total	31	

Statistical model:

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 4 \\ k = 1, 2 \end{cases}$$

Figure 6.12. The experiment diagram, expected mean squares, and the statistical model for the treatment structure [A: 4, B: 4, (R): 2] and a fully randomized design structure.

The A_i and B_j shown in the statistical model are factor A and factor B effects respectively. The factor effect is defined as the possible departure of the overall mean response for a level of the fixed factor. For example, for the fixed factor A,

$$Mean\ response\ at\ level\ "i"of\ factor\ A = \mu_i = \mu + A_i \qquad (6.1)$$

Since μ is the mean response over all factor A levels, the following holds.

$$\frac{\sum_{i=1}^{4} \mu_i}{4} = \mu \qquad (6.2)$$

Combining (6.1) and (6.2) results in the following.

$$\sum_{i=1}^{4} A_i = 0 \qquad (6.3)$$

This also holds for factor B,

$$\sum_{j=1}^{4} B_j = 0 \qquad (6.4)$$

and the interaction effects.

$$\sum_{i=1}^{4} AB_{ij} = 0 \qquad and \qquad \sum_{j=1}^{4} AB_{ij} = 0 \qquad (6.5)$$

The appropriate hypothesis to test a factor A effect on mean response is

$$\begin{cases} H_0\colon \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_1\colon Otherwise. \end{cases} \qquad (6.6)$$

Using equation (6.1), this hypothesis is equivalent to

$$\begin{cases} H_0\colon A_1 = A_2 = A_3 = A_4 = 0 \\ H_1\colon Otherwise. \end{cases} \qquad (6.7)$$

Rejecting the null hypothesis in (6.7) is interpreted as an evidence for a significant factor A effect on the response variable. Similar hypotheses are used for a factor B and AB interaction effect.

The test statistics for these hypotheses are constructed by taking the ratio of two mean squares where the expected value of the numerator mean square differs from the expected value of the denominator mean squares by the variance component or fixed effect of interest. If such a ratio these test ratios are referred to as direct or exact tests.

For example, from Figure 6.12, the expected value of factor A mean squares is

$$E(MS_A) = \sigma^2 + 8 * \Phi(A) = \sigma^2 + 8 * \frac{\sum_{i=1}^{4} {A_i}^2}{(4-1)} \qquad (6.8)$$

And the expected value of the error mean square is

$$E(MS_E) = \sigma^2 \qquad (6.9)$$

The difference between these two expected mean squares isolates the term $\frac{\sum_{i=1}^{4} {A_i}^2}{(4-1)}$ which is the variability explained by different factor A levels.

Therefore, the factor A mean square divided by the error mean square can be used as an exact test for a significant factor A effect. Similarly, the factor B and interaction effects in this replicated fully randomized design can be directly tested and are estimable.

6.2.2 The Analysis of Fully Restrictive Design Structures

Consider the experiment in Figure 6.13. This experiment has the same treatment structure [A: 4, B: 4, (R): 2] as the example shown in Figure 6.12, but with a fully restrictive design structure {(A, 1, 1)}.

		Factor A			
		1	2	3	4
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 8*\sigma^2A\_r$
A_r	0	$\sigma^2 + 8*\sigma^2A\_r$
B	3	$\sigma^2 + 8*\Phi(B)$
AB	9	$\sigma^2 + 2*\Phi(AB)$
Error	16	σ^2
Total	31	

Statistical model:

$$Y_{ijk} = \mu + A_i + A\_r_{j(i)} + B_k + AB_{ik} + \varepsilon_{ijkl} \qquad \begin{cases} i = 1, \dots, 4 \\ j = 1 \\ k = 1, \dots, 4 \\ l = 1,2 \end{cases}$$

Figure 6.13. The experiment diagram, expected mean squares, and the statistical model for the treatment structure [A: 4, B: 4, (R): 2] and a fully restricted design structure {(A, 1, 1)}

In the statistical model shown in Figure 6.13, the A_r term represents the randomization restriction on A. Note that the number of degrees of freedom for this term is zero. Therefore, the mean squares value cannot be calculated for this term and its direct effect on the response cannot be tested. It is also assumed that there are no interaction effects between this term and other effects in the model.

Adding the variance component for this term to the expected mean squares of factor A changes the nature of the statistical test for a factor A effect. The test ratio for factor A is now constructed as

$$\frac{E(MS_A)}{E(MS_E)} = \frac{\sigma^2 + 8 * \frac{\sum_{i=1}^{4} {A_i}^2}{(4-1)} + 8 * {\sigma_{A\_r}}^2}{\sigma^2} \quad (6.10)$$

The difference between the numerator and the denominator of this ratio is $8 * \frac{\sum_{i=1}^{4} {A_i}^2}{(4-1)} + 8 * {\sigma_{A\_r}}^2$ which is a combination of the factor A effect, and the randomization restriction effect. If the result of this test turns out to be insignificant, then neither source of variability has a significant effect on the response variable. However, if the result is significant, no conclusion can be made as to which one of the variability sources are have a significant effect. These types of tests are referred to as conservative tests (Lorenzen and Anderson, 1993).

Note that both factor B and AB interaction effects can still have direct tests.

As another example, consider the experiment illustrated in Figure 6.14. This is an example from a very commonly used groups of experimental design structures known as “split-plot” designs. To show how the two layers of randomization restriction are affecting the statistical model, two tables of expected mean squares are shown in Figure 6.14. In the table on the left of Figure 6.14 replicates are treated as a random factor and the experiment is modeled as a three-factor factorial design (A and B fixed, R random).

In the table on the right all the sources of variability involving the replicates are removed from the model and are aggregated into corresponding error terms. Error_1 represents the variability explained by the blocks of replicates. Error_2 is the variability explained by the interaction between replicates and factor A. The rest of the variabilities explained by replicates are aggregated into Error_3.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR + 4*\sigma^2AR\_r$
B	3	$\sigma^2 + 8*\Phi(B) + 4*\sigma^2BR$
R	1	$\sigma^2 + 16*\sigma^2(R) + 16*\sigma^2R\_r + 4*\sigma^2AR\_r$
R_r	0	$\sigma^2 + 16*\sigma^2R\_r$
AB	9	$\sigma^2 + 2*\Phi(AB) + \sigma^2ABR$
AR	3	$\sigma^2 + 4*\sigma^2AR + 4*\sigma^2AR\_r$
AR_r	0	$\sigma^2 + 4*\sigma^2AR\_r$
BR	3	$\sigma^2 + 4*\sigma^2BR$
ABR	9	$\sigma^2 + \sigma^2ABR$
Error	0	σ^2
Total	31	

Effect	DF	Expected Mean Square
A	3	$\sigma^2 + 8*\Phi(A) + 4*\sigma^2AR\_r$
B	3	$\sigma^2 + 8*\Phi(B)$
R_r	0	$\sigma^2 + 16*\sigma^2R\_r$
AB	9	$\sigma^2 + 2*\Phi(AB)$
AR_r	0	$\sigma^2 + 4*\sigma^2AR\_r$
Error1	1	$\sigma^2 + 16*\sigma^2R\_r + 4*\sigma^2AR\_r$
Error2	3	$\sigma^2 + 4*\sigma^2AR\_r$
Error3	12	σ^2
Total	31	

Statistical model:

$$Y_{ijk} = e1_i + e1\_r_{j(i)} + A_k + e2_{ik} + e2\_r_{l(ik)} + B_m + AB_{km} + \varepsilon_{ijklm} \qquad \begin{cases} i = 1,2 \\ j = 1 \\ k = 1, \dots, 4 \\ l = 1 \\ m = 1, \dots, 4 \end{cases}$$

Figure 6.14. The diagram, expected mean squares, and the statistical model of a Split-Plot design structure with the treatment structure of [A: 4, B: 4, (R): 2].

The interesting fact about this design structure is that, even though there are two layers of randomization restrictions in the experiment, all factor and interaction effects (A, B, and AB) in the model are estimable and have exact tests. However, these effects are being estimated at different levels of precision. The denominator of the exact test formed for factor A (or the whole-plot factor) is the mean square of Error_2 (or the whole-plot error) with only three degrees of freedom. The denominator of the exact tests formed for factor B and AB interaction effect (referred to as sub-plot effects) is the mean square of Error_3 (or the sub-plot error) with twelve degrees of freedom.

6.2.3 The Analysis of Partially Restrictive Design Structures

Consider the experiment depicted in Figure 6.15. This is a single factor experiment with the treatment structure of [A: 4, (R): 8] and the design structure of {(A, 2, 1)}.

In this design structure, the four levels of factor A are equally split into two fully randomized boxes, or blocks. Due to the restricted randomization, the only valid comparisons among the levels of factor A are the ones made between levels from the same block and not cross blocks. Thus, to perform any analyses on factor A, the total variability explained by the levels of this factor is broken down into smaller portions in such a way that estimable portions can be separated from the ones that are confounded with the effects of the randomization restriction.

Factor A			
1	2	3	4
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2

Figure 6.15. The diagram of an experiment with treatment structure of [A: 4, (R): 8] and design structure {(A, 2, 1)}

6.2.4 Splitting a Source of Variability

The sources of variability in an analysis of variance are estimated and tested for significance. The variability is calculated using sum of squares formulas. For example, Equation (6.11) shows the sum of squares formula for factor A in a single-factor experiment with the treatment structure of [A: a, (R): r].

$$SS_A = \sum_{i=1}^{i=a} (\overline{y_{i.}} - \overline{y_{..}})^2 \qquad (6.11)$$

Where $\overline{y_{i.}}$ is the mean response at level *i* of factor A and $\overline{y_{..}}$ is the overall mean response. According to this formula, the result of this calculation is the total *between-level,* or *between-treatment,* variation of the levels of factor A.

There are times, however, that the experimenter is interested in some preplanned comparisons between specific treatment means. These comparisons are expressed with customized hypotheses constructed to answer specific questions. Following is an example of customized hypothesis.

$$\begin{cases} H_0\colon\ \mu_1 + \mu_2 = 2 * \mu_3 \\ H_1\colon\ \mu_1 + \mu_2 \neq 2 * \mu_3 \end{cases} \qquad (\ 6.12\ )$$

The comparison of different treatment means can be expressed using *contrasts*. Contrasts are linear combination of parameters whose coefficients add up to zero. Considering a factor with "a" levels, the corresponding contrasts are of the form

$$C = \sum_{i=1}^{a} c_i \mu_i \qquad where\colon\ \sum_{i=1}^{a} c_i = 0 \qquad (\ 6.13\ )$$

For example, in the hypothesis (6.12), $c_1 = 1$, $c_2 = 1$, $c_3 = -2$, anc $c_i = 0$ for *i = 4, ..., a*.

A useful special case of contrasts are the *orthogonal contrasts*. The relationship between two vectors is defined as orthogonal if the inner product of them is zero. For two contrasts with coefficients $\{c_i\}$ and $\{d_i\}$, orthogonality is simply achieved if the following condition holds.

$$\sum_{i=1}^{a} c_i d_i = 0 \qquad (\ 6.14\ )$$

Since the covariance between two variables is formed as an inner product, orthogonal contrasts are independent from each other. Because of this independence, each orthogonal contrast provides different information than the others. For a factor with "a" levels, the set of *a - 1* orthogonal contrasts partition the total sum of squares, explained by those levels,

into *a - 1* independent single-degree-of-freedom components with all of the information that can be captured about the corresponding effect from the experiment.

Splitting the sources of variability into smaller independent components in forms of contrasts can be utilized as a solution to the analysis of partially restrictive design structures. Consider the example shown in Figure 6.15. In order to separate the confounded portions of variability, explained by factor A, from other variability, a set of three orthogonal contrasts is defined. The following shows these contrasts along with their corresponding test hypothesis.

$$
\begin{aligned}
C_1 = \mu_1 + \mu_2 - \mu_3 - \mu_4 \quad &=> \quad \begin{cases} H_0\colon\ \mu_1 + \mu_2 = \mu_3 + \mu_4 \\ H_1\colon\ \mu_1 + \mu_2 \neq \mu_3 + \mu_4 \end{cases} \quad => \quad \begin{cases} H_0\colon\ A_1 + A_2 = A_3 + A_4 \\ H_1\colon\ A_1 + A_2 \neq A_3 + A_4 \end{cases} \\
C_2 = \mu_1 - \mu_2 \quad &=> \quad \begin{cases} H_0\colon\ \mu_1 = \mu_2 \\ H_1\colon\ \mu_1 \neq \mu_2 \end{cases} \quad => \quad \begin{cases} H_0\colon\ A_1 = A_2 \\ H_1\colon\ A_1 \neq A_2 \end{cases} \\
C_3 = \mu_3 - \mu_4 \quad &=> \quad \begin{cases} H_0\colon\ \mu_3 = \mu_4 \\ H_1\colon\ \mu_3 \neq \mu_4 \end{cases} \quad => \quad \begin{cases} H_0\colon\ A_3 = A_4 \\ H_1\colon\ A_3 \neq A_4 \end{cases}
\end{aligned}
\qquad (\ 6.15\ )
$$

The sum squares formulas for these three contrasts is shown in (6.16).

$$
\begin{aligned}
SS_{C_1} &= \frac{(\bar{y}_{1.} + \bar{y}_{2.} - \bar{y}_{3.} - \bar{y}_{4.})^2}{\frac{4}{8}} \\
SS_{C_2} &= \frac{(\bar{y}_{1.} - \bar{y}_{2.})^2}{\frac{2}{8}} \\
SS_{C_3} &= \frac{(\bar{y}_{3.} - \bar{y}_{4.})^2}{\frac{2}{8}}
\end{aligned}
\qquad (\ 6.16\ )
$$

Where the $\bar{y}_{i.}$ is the average response of column *i* in Figure 6.15.

In an experiment with no randomization restrictions, the three test hypotheses in (6.15) can be used to evaluate the test hypothesis in (6.6). The sum of three sum of squares values in (6.16) is also equal to the total sum of squares value in (6.11).

However, because of the randomization restriction on factor A, the first contrast, C_1, which is making a comparison between the restricted factors, is confounded with the effect of the randomization restriction and therefore, will be a conservative test. The other two contrasts, C_2 and C_3, which are making within-block comparisons between the levels of factor A have exact tests.

The breakdown of the factor A variation into three orthogonal contrasts can be modeled with a new treatment structure. Consider the experiment shown in Figure 6.16. It shows the experimental diagram and the statistical model of a fully randomized design structure for the treatment structure of [$A^{(2)}$: 2, $A^{(1)} \rightarrow A^{(2)}$: 2, (R): 8]. The arrow in the treatment structure indicates that factor $A^{(1)}$ is nested within factor $A^{(2)}$.

$A^{(2)}$ 1		$A^{(2)}$ 2	
$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2
R1	R1	R1	R1
R2	R2	R2	R2
R1	R1	R1	R1
R2	R2	R2	R2
R1	R1	R1	R1
R2	R2	R2	R2
R1	R1	R1	R1
R2	R2	R2	R2

Statistical model:

$$Y_{ijk} = \mu + {A^{(2)}}_i + {A^{(1)}}_{j(i)} + \varepsilon_{ijk}$$

$$\begin{cases} i = 1, 2 \\ j = 1, 2 \\ k = 1, \dots, 8 \end{cases}$$

Figure 6.16. The diagram and the statistical model of a fully randomized experiment with the treatment structure of [$A^{(2)}$: 2, $A^{(1)}$ -> $A^{(2)}$: 2, (R): 8]

In the statistical model in Figure 6.16, ${A^{(2)}}_i$ is the effect of $A^{(2)}$ and the ${A^{(1)}}_{i(j)}$ is the effect of $A^{(1)}$ nested within $A^{(2)}$. The modified test hypotheses for factors $A^{(2)}$ and $A^{(1)}$ are shown in (6.17).

Test hypothesis for factor $A^{(2)}$:

$$\begin{cases} H_0\colon \mu^{(2)}{}_1 = \mu^{(2)}{}_2 \\ H_1\colon \mu^{(2)}{}_1 \neq \mu^{(2)}{}_2 \end{cases} \quad ==> \quad \begin{cases} H_0\colon A^{(2)}{}_1 = A^{(2)}{}_2 \\ H_1\colon A^{(2)}{}_1 \neq A^{(2)}{}_2 \end{cases}$$

Test Hypothesis for factor $A^{(1)}$:

$$for\ each\ level\ of\ A^{(2)}\colon \begin{cases} H_0\colon \mu^{(1)}{}_1 = \mu^{(1)}{}_2 \\ H_1\colon \mu^{(2)}{}_1 \neq \mu^{(2)}{}_2 \end{cases} \quad ==> \quad \begin{cases} H_0\colon A^{(1)}{}_{1(1)} = A^{(1)}{}_{2(1)} \\ H_1\colon A^{(1)}{}_{1(1)} \neq A^{(1)}{}_{2(1)} \end{cases} \quad \begin{cases} H_0\colon A^{(2)}{}_{1(2)} = A^{(2)}{}_{2(2)} \\ H_1\colon A^{(2)}{}_{1(2)} \neq A^{(2)}{}_{2(2)} \end{cases} \qquad (6.17)$$

The sum of squares formulas for $A^{(2)}$ and $A^{(1)}$ is shown in (6.18).

$$\begin{aligned} SS_{A^{(2)}} &= \frac{1}{16}((y_{1.} + y_{2.})^2 + (y_{3.} + y_{4.})^2) - \frac{y_{..}{}^2}{32} \\ SS_{A^{(1)}} &= \frac{1}{8}(y_{1.}{}^2 + y_{2.}{}^2 + y_{3.}{}^2 + y_{4.}{}^2) - \frac{1}{16}((y_{1.} + y_{2.})^2 + (y_{3.} + y_{4.})^2) \end{aligned} \qquad (6.18)$$

It can be shown that the analysis for this statistical model is equivalent to using the three orthogonal contrasts in (6.15). First, note that the following assumptions hold for $A^{(2)}$ and $A^{(1)}$.

$$\begin{aligned} \sum_{i=1}^{2} A^{(2)}{}_i &= 0 \\ \sum_{j=1}^{2} A^{(1)}{}_{j(1)} &= 0 \\ \sum_{j=1}^{2} A^{(1)}{}_{j(2)} &= 0 \end{aligned} \qquad (6.19)$$

Second, based on the treatment structure in Figure 6.15 and regardless of the randomization restrictions, the mean response for each column of observations is as follows.

$$\mu_i = \mu + A_i \qquad (6.20)$$

Where A_i is the i^{th} treatment effect of factor A. For the treatment structure in Figure 6.16, the mean response of column of observations is

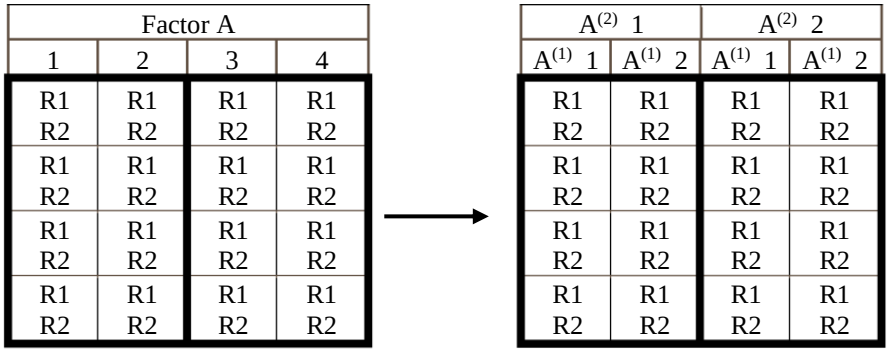
$$\mu_{ij} = \mu + A^{(2)}{}_i + A^{(1)}{}_{j(i)} \qquad (6.21)$$

Combining (6.20) and (6.21) gives the following relationships between the treatment effects.

$$\begin{aligned}
\mu_{column\ 1} - \mu &= A_1 = A^{(2)}{}_1 + A^{(1)}{}_{1(1)} \\
\mu_{column\ 2} - \mu &= A_2 = A^{(2)}{}_1 + A^{(1)}{}_{2(1)} \\
\mu_{column\ 3} - \mu &= A_3 = A^{(2)}{}_2 + A^{(1)}{}_{1(2)} \\
\mu_{column\ 4} - \mu &= A_4 = A^{(2)}{}_2 + A^{(1)}{}_{2(2)}
\end{aligned} \qquad (6.22)$$

Using (6.19) and (6.22), it can be shown that the test hypotheses in (6.15) are equivalent to the test hypotheses in (6.17). As a result, it is possible to change the treatment structure of the experiment, by breaking down a source of variability into two or more independent sources of variability and maintain the information captured by the original treatment structure. This is particularly useful when the design structure of an experiment involves partially restrictive randomization restrictions such as the one shown in Figure 6.15.

Figure 6.17 depicts the experimental diagram, expected mean squares, and the statistical model of the experiment in Figure 6.15. The treatment structure [A: 4, (R): 8] is transformed to [$A^{(2)}$: 2, $A^{(1)} \rightarrow A^{(2)}$: 2, (R): 8].



Factor A			
1	2	3	4
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2

$A^{(2)}$ 1		$A^{(2)}$ 2	
$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2
R1 R2	R1 R2	R1 R2	R1 R2

Statistical model:

$$Y_{ijk} = \mu + A^{(2)}{}_i + A^{(2)}\_r_{j(i)} + A^{(1)}{}_{k(i)} + \varepsilon_{ijkl}$$

$$\begin{cases} i = 1, 2 \\ j = 1 \\ k = 1, 2 \\ l = 1, \dots, 8 \end{cases}$$

Effect	DF	Expected Mean Square
A(2)	1	$\sigma^2 + 16*\Phi(A^{(2)}) + 8*\sigma^2{}_{A^{(2)}\_r}$
A(2)_r	0	$\sigma^2 + 8*\sigma^2{}_{A^{(2)}\_r}$
A(1)	2	$\sigma^2 + 8*\Phi(A^{(1)})$
Error	28	σ^2
Total	31	

Figure 6.17. Transformation of treatment structure [A: 4, (R): 8] to [$A^{(2)}$: 2, $A^{(1)}$ -> $A^{(2)}$: 2, (R): 8] in order to analyze the partially restrictive design structure {(A, 2, 1)}.

In the statistical model $A^{(2)}{}_i$ is the effect of $A^{(2)}$, $A^{(1)}{}_{k(i)}$ is the effect of $A^{(1)}$, and $A^{(2)}\_r$ is the term for the randomization restriction on $A^{(2)}$. From the statistical model and expected mean squares, it can be seen that the any factor $A^{(2)}$ effect is confounded with the effect of the partially restrictive randomization restriction, and therefore there will be no exact test for that effect. However, $A^{(1)}$ has an exact test and effect, $A^{(1)}{}_{k(i)}$, which is a part of factor A's treatment effect, can be estimated.

This type of transformation of treatment structures can be generalized to larger experiments with more complicated design structures. Figure 6.18 and Figure 6.19 show the transformation of two other single factor experiments with partially restrictive randomization restrictions of {(A, 3, 1)} and {(A, 4, 1), (A, 2, 2)}, respectively.

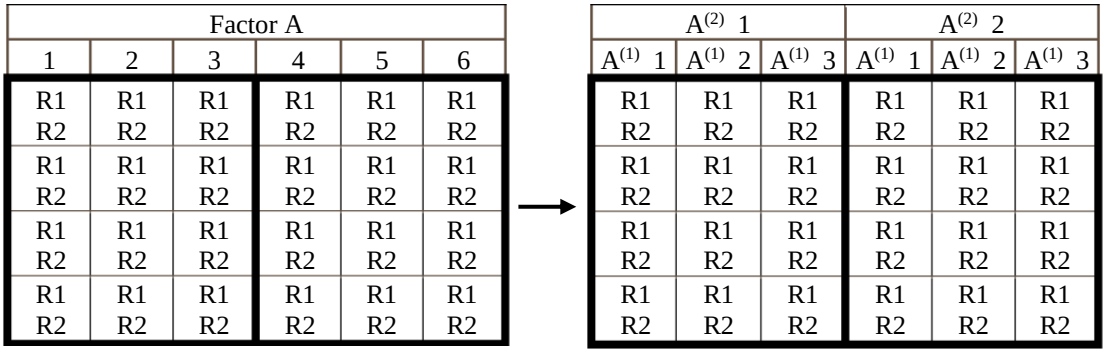


Figure 6.18. Transforming the treatment structure [A: 6, (R): 8] to [$A^{(2)}$: 2, $A^{(1)}$ -> $A^{(2)}$: 3, (R): 8] for the analysis of the design structure {(A, 3, 1)}.

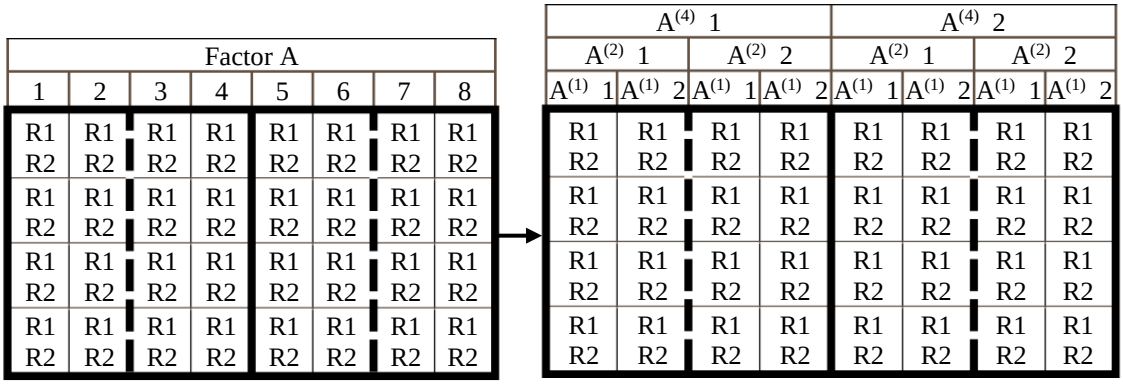


Figure 6.19. Transforming the treatment structure [A: 8, (R): 8] to [$A^{(4)}$: 2, $A^{(2)}$ -> $A^{(4)}$: 2, $A^{(1)}$ -> $A^{(2)}$: 2, (R): 8] for the analysis of the design structure {(A, 4, 1), (A, 2, 2)}.

The design structure in Figure 6.19 is a two-layered design structure. In the first layer, a randomization restriction of size four is applied on factor A. At the second layer, another randomization restriction of size two is applied on A. As a result, factor A is split into a two-stage nested structure of $A^{(4)}$, $A^{(2)}$, and $A^{(1)}$.

Partially restrictive randomization restrictions can be applied to more than one factor in a design structure. Consider the experiment in Figure 6.20. The design structure used is a two-layered design structure represented as {(R, 1, 1), (A, 2, 2), (B, 2, 2)}. To separate the effect of randomization restrictions from other effects, the original treatment structure [A: 4, B: 4, (R): 2] is transformed into [$A^{(2)}$: 2, $A^{(1)}$ → $A^{(2)}$: 2, $B^{(2)}$: 2, $B^{(1)}$ → $B^{(2)}$: 2, (R): 2].

		Replicate 1 Factor A				Replicate 2 Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

		Replicate 1				Replicate 2			
		$A^{(2)}$ 1		$A^{(2)}$ 2		$A^{(2)}$ 1		$A^{(2)}$ 2	
		$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2
$B^{(2)}$ 1	$B^{(1)}$ 1	R1	R1	R1	R1	R2	R2	R2	R2
	$B^{(1)}$ 2	R1	R1	R1	R1	R2	R2	R2	R2
$B^{(2)}$ 2	$B^{(1)}$ 1	R1	R1	R1	R1	R2	R2	R2	R2
	$B^{(1)}$ 2	R1	R1	R1	R1	R2	R2	R2	R2

Effect	**DF**	**Expected Mean Square**
A(2)	1	$\sigma^2 + 16*\Phi(A^{(2)}) + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
A(1)	2	$\sigma^2 + 8*\Phi(A^{(1)})$
B(2)	1	$\sigma^2 + 16*\Phi(B^{(2)}) + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
B(1)	2	$\sigma^2 + 8*\Phi(B^{(1)})$
R_r	0	$\sigma^2 + 16*\sigma^2_{R\_r}$
A(2)B(2)	1	$\sigma^2 + 8*\Phi(A^{(2)}B^{(2)}) + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
A(2)B(1)	2	$\sigma^2 + 4*\Phi(A^{(2)}B^{(1)})$
A(1)B(2)	2	$\sigma^2 + 4*\Phi(A^{(1)}B^{(2)})$
A(1)B(1)	4	$\sigma^2 + 2*\Phi(A^{(1)}B^{(1)})$
A(2)B(2)R_r	0	$\sigma^2 + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
Error1	1	$\sigma^2 + 16*\sigma^2_{R\_r} + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
Erorr2	3	$\sigma^2 + 4*\sigma^2_{A^{(2)}B^{(2)}R\_r}$
Error3	12	σ^2
Total	31	

Statistical model:

$$Y_{ijk} = e1_i + e1\_r_{j(i)} + A^{(2)}{}_k + A^{(1)}{}_{l(k)} + B^{(2)}{}_m + B^{(1)}{}_{n(m)} + A^{(2)}B^{(2)}{}_{km} + e2_{ikm} + e2\_r_{o(ikm)} + A^{(1)}B^{(2)}{}_{l(k)m} + A^{(2)}B^{(1)}{}_{kn(m)} + A^{(1)}B^{(1)}{}_{l(k)n(m)} + \varepsilon_{ijklmno}$$

$$\begin{cases} i = 1,2 \\ j = 1 \\ k = 1,2 \\ l = 1,2 \\ m = 1,2 \\ n = 1,2 \\ o = 1 \end{cases}$$

Figure 6.20. Transformation of the treatment structure [A: 4, B: 4, (R): 2] to [A(2): 2, A(1) -> A(2): 2, B(2): 2, B(1) -> B(2): 2, (R): 2] to analyze the partially restrictive design structure {(R, 1, 1), (A, 2, 2), (B, 2, 2)}.

In Figure 6.20, factor A and factor B effects are both split into two different sources of variability. The AB interaction effect, however, is split into four different independent sources of variation which is a result of the factors A and B splitting. All these sources of variability have exact tests and can be estimated. $A^{(2)}$, $B^{(2)}$, and $A^{(2)}B^{(2)}$ effects are tested using error2 and the rest are tested using error3.

Splitting sources of variation causes the single original test to be replaced with multiple tests. This increase in number of tests will consequently increase the type I error of the statistical tests for an effect. There are several well-known approaches to adjust the significance level, or type I error rate of individual tests so that the type I error rate of the overall test stays at a specified level.

One of the most commonly used approaches for multiple comparisons is the Bonferroni correction (1936). It uses Boole's inequality which guarantees that if each of k tests is performed at type I error rate α/k, the total type I error rate will not exceed α. It is worth mentioning that, because of this guarantee, Bonferroni correction is considered a conservative approach and reduces the individual type I errors more than actually needed. This is specially the case when the tests are positively correlated. As a result, the overall computed power of the tests will be underestimated.

7. Expected Time and Cost of Design Structures

In this chapter, the models and algorithms for finding the expected total level changing cost and time to run all experiments in an experimental design will be presented. The models are for the expected costs and time since all experimental designs will have some randomization of run order. The focus is on factor level changing since the total cost and time to conduct the experiments will be constant for a specific treatment structure. This chapter begins with the derivation of the expected number of level changes for a single factor when there are a fixed number of factors and replicates per factor. Next the models for the expected total time for factor level changes (TLCT), and models for the expected experimental cost (TLCC) will be presented. This will be followed by a presentation of recursive algorithms.

This section is organized as follows:

- Derivation of the expected number of level changes for a particular factor
- Derivation of a general formula for calculating the total level changing time in a fully randomized design
- Derivation of a general formula for calculating the total level changing cost in a fully randomized design
- The effect of randomization restrictions on level changing time and cost
- Development of a recursive algorithm for calculating the total level changing time
- Development of a recursive algorithm for calculating the total level changing cost

7.1 Expected Number of Level Changes

The expected number of level changes for a single factor (one factor within a balanced multi-factor experiment) is derived. It is assumed for a particular experiment that there is no difference between running two or more consecutive experiments at the same level with no factor “resetting”, and running the same experiments but with the factor levels reset. A factor level is reset to the same level by first changing it to a different level, and then resetting it back to the same level. Whether this assumption holds, is one of engineering judgement and must be made for different experiments. When it is a reasonable assumption

"inadvertent split-plot" experiments (Bradley Jones and Christopher J. Nachtsheim, 2009) will not occur. An extreme example of this are computer experiments where there is no need to reset parameter values. Computer experiments are also not affected by randomization of experimental order, or the lack thereof since there is no unknown, uncontrollable experimental outcome influences. If factor levels are always reset after each run, then the total number of level changes for all the factors is equal to the total number of experimental runs.

Theorem: The expected number of level changes for a single factor with L levels and R replicates within each level (independent of other existing factors in the experiment) is,

$$R * (L - 1) + 1 \qquad (7.1)$$

Proof:

With R runs within each level, the total number of experiments is L*R. For example, in a balanced completely randomized single factor experiment with four levels and three replicates per level, there will be a total of 12 experimental runs. Figure 7.1 shows a random assignment of these 12 runs to different levels of this factor.

$$\frac{Run\ number}{Level\ number}: \frac{1}{2}, \frac{2}{4}, \frac{3}{1}, \frac{4}{1}, \frac{5}{3}, \frac{6}{2}, \frac{7}{3}, \frac{8}{4}, \frac{9}{2}, \frac{10}{4}, \frac{11}{1}, \frac{12}{3}$$

Figure 7.1. The assignment of twelve runs to four levels of a factor

Define D_{li} as:

$$D_{li} = \begin{cases} 1 & if\ the\ ith\ run\ belongs\ to\ the\ lth\ level \\ 0 & otherwise \end{cases} \qquad (7.2)$$

In the example above, $D_{13} = 1$ because the 3rd run is at level 1. Similarly, since the 8th run is at level 4 and not level 2, $D_{28} = 0$.

The expected value for $D_{li} * D_{l'i'}$ is,

$$E(D_{li} * D_{l'i'}) = \begin{cases} \dfrac{R-1}{RL-1} & for\ l = l' \\ \dfrac{R(L-1)}{RL-1} & for\ l \neq l' \end{cases} \quad (7.3)$$

Consider D_{13} and D_{14} in the example above. The outcome of the product D_{13}*D_{14} is 1 if and only if D_{13} and D_{14} are equal to 1, or in other words, when the 3rd and 4th runs are at level 1. Using straightforward combinatorics,

$$P(D_{13} * D_{14} = 1) = \frac{\binom{3}{1}}{\binom{12}{1}} * \frac{\binom{2}{1}}{\binom{11}{1}}$$

In general, for a specific level *l*,

$$P(D_{li} * D_{li'} = 1) = \frac{\binom{R}{1}}{\binom{RL}{1}} * \frac{\binom{R-1}{1}}{\binom{RL-1}{1}} = \frac{R-1}{L(RL-1)} \quad (7.4)$$

For any arbitrary level *l* in *L:*

$$P(D_{li} * D_{li'} = 1) = \binom{L}{1} * \frac{\binom{R}{1}}{\binom{RL}{1}} * \frac{\binom{R-1}{1}}{\binom{RL-1}{1}} = \frac{R-1}{RL-1} \quad (7.5)$$

Since $D_{li} * D_{li'}$ is either zero or one, the prior probability is also the expected value,

$$E(D_{li} * D_{li'}) = \frac{R-1}{RL-1}. \quad (7.6)$$

Next consider D_{13} and D_{24}. The outcome of their product is 1 if and only if both D_{13} and D_{24} are equal to 1. The 3rd run is at level 1 and the 4th run at level 2. Using straightforward combinatorics,

$$P(D_{13} * D_{24} = 1) = \frac{\binom{3}{1}}{\binom{12}{1}} * \frac{\binom{3}{1}}{\binom{11}{1}}$$

In general, for a specific level *l* and a specific level *l'*,

$$P(D_{li} * D_{l'i'} = 1) = \frac{\binom{R}{1}}{\binom{RL}{1}} * \frac{\binom{R}{1}}{\binom{RL-1}{1}} = \frac{R}{L(RL-1)} \quad (7.7)$$

For arbitrary levels l and l' ($l \neq l'$) in L:

$$P(D_{li} * D_{l'i'} = 1) = \binom{L}{1} * \binom{L-1}{1} * \frac{\binom{R}{1}}{\binom{RL}{1}} * \frac{\binom{R}{1}}{\binom{RL-1}{1}} = \frac{R(L-1)}{RL-1} \quad (7.8)$$

Therefore,

$$E(D_{li} * D_{l'i'}) = \frac{R(L-1)}{RL-1} \quad (7.9)$$

The derived expected value holds for any run i and i', and in particular when $i' = i+1$.

$$E\left(D_{li} * D_{l'(i+1)}\right) = \frac{R(L-1)}{RL-1} \quad (7.10)$$

Since there are RL runs, there are RL-1 pairs of $\{i, i'\}$ where $i' = i+1$. Therefore, the expected number of level changes is:

$$E\left(\sum_{i=1}^{RL-1} D_{li} * D_{l'(i+1)}\right) = \sum_{i=1}^{RL-1} E\left(D_{li} * D_{l'(i+1)}\right) = \sum_{i=1}^{RL-1} \frac{R(L-1)}{RL-1} = R(L-1) \quad (7.11)$$

Since setting the level for the very first run needs the same amount of setup time or cost as any other level change, the expected number of level changes for a factor is:

$$E(number\ of\ level\ changes) = R(L-1) + 1 \quad (7.12)$$

7.2 Total Level Changing Time

A formula is derived for the expected total level changing time (TLCT) for a fully randomized balanced experiment. It is assumed that all factor levels are changed at the same time. Under this assumption, the time to change all factor levels for a single experiment is equal to the level change time for the factor with the largest level change

time. For example, suppose there are three factors A, B, and C in an experiment with level changing times of lct_a, lct_b, and lct_c, respectively. Assume that $lct_b > lct_a > lct_c$. If for a particular transition (going from one experimental run to another), the levels of all three factors have to be changed, then the corresponding time is equal to lct_b. However, if the level of factor B does not have to be changed at this transition (i.e., the next run will be at the same level of factor B), then the corresponding factor level change time will be equal to lct_a since factor A has the next largest factor level changing time.

Next consider an experiment with *I* factors. Define z_i as:

$$z_i = The\ factor\ with\ ith\ largest\ level\ changing\ time$$

For example, z_1 is the factor with the largest level changing time and the number of levels for this factor and the number of replicates within each level are denoted by L_{z_1} and R_{z_1} respectively. z_2 is the factor with the second largest level changing time with L_{z_2} levels and R_{z_2} replicates within each level; and so on.

From equation (7.8), the probability of a level-change at a particular transition is $\frac{R(L-1)}{RL-1}$. Thus, the probability of a no-level-change at a particular transition is $\frac{R-1}{RL-1}$.

For a particular transition, define the binary variable lc_i as:

$$lc_i = \begin{cases} 1 & if\ factor\ i\ changes\ level \\ 0 & otherwise \end{cases} \qquad (\ 7.13\ )$$

The expected factor level changing at a particular transition is:

$$Pr(lc_{z_1} = 1) * lct_{z_1} + Pr(lc_{z_1} \neq 1) * (Factor\ level\ change\ time\ when\ lc_{z_1} \neq 1) \qquad (\ 7.14\ )$$

Substituting the probability expressions into the prior equation gives:

$$E[Factor\ change\ time] = \qquad (\ 7.15\ )$$

$$\frac{R_{z_1}(L_{z_1}-1)}{R_{z_1}L_{z_1}-1} * lct_{z_1} + \frac{R_{z_1}-1}{R_{z_1}L_{z_1}-1} * \left(Factor\ level\ change\ time\ when\ lc_{z_1} \neq 1\right)$$

When $lc_{z_1} \neq 1$ at a transition, the factor level change time will depend on z_2. Therefore, the expected factor level change time when $lc_{z_1} \neq 1$ is equal to:

$$Pr(lc_{z_2} = 1) * lct_{z_2} + Pr(lc_{z_2} \neq 1) * \left(Factor\ level\ change\ time\ when\ lc_{z_2} \neq 1\right) \quad (7.16)$$

The probability expressions to substitute into the prior equation are conditional on $lc_{z_1} \neq 1$. The difference in the probability expressions is that the number of replicates for z_2 is now reduced to $\frac{R_{z_2}}{L_{z_1}}$. Since z_1 is not changing levels the level changes for other factors including z_2 are within the same level of z_1 and thus only the factor z_2 replicates within the same level of z_1 are relevant. Therefore,

$$E\left[Factor\ level\ change\ time\ when\ lc_{z_1} \neq 1\right] =$$

$$\frac{\frac{R_{z_2}}{L_{z_1}}(L_{z_2}-1)}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1} * lct_{z_2} + \frac{\frac{R_{z_2}}{L_{z_1}}-1}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1} * \left(Factor\ level\ change\ time\ when\ lc_{z_2} \neq 1\right) \quad (7.17)$$

If *I = 3*, the expected factor level changing time for a particular transition will be:

$$\frac{R_{z_1}(L_{z_1}-1)}{R_{z_1}L_{z_1}-1} * lct_{z_1} + \frac{R_{z_1}-1}{R_{z_1}L_{z_1}-1} *$$

$$\left(\frac{\frac{R_{z_2}}{L_{z_1}}(L_{z_2}-1)}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1} * lct_{z_2} + \frac{\frac{R_{z_2}}{L_{z_1}}-1}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1} * \left(\frac{\frac{R_{z_3}}{L_{z_1}L_{z_2}}(L_{z_3}-1)}{\frac{R_{z_3}}{L_{z_1}L_{z_2}}L_{z_3}-1} * lct_{z_3} + \frac{\frac{R_{z_3}}{L_{z_1}L_{z_2}}-1}{\frac{R_{z_3}}{L_{z_1}L_{z_2}}L_{z_3}-1} * 0\right)\right) \quad (7.18)$$

Note that the "$Factor\ time\ when\ lc_{z_3} \neq 1$" is set to zero since none of the three factors are changing levels.

With *N* total experimental runs, there will be *N-1* transitions. Therefore, the expected level changing time for all the transitions is equal to:

$$\left[\left(\begin{array}{c}\dfrac{R_{z_1}\left(L_{z_1}-1\right)}{R_{z_1}L_{z_1}-1}*lct_{z_1}+\dfrac{R_{z_1}-1}{R_{z_1}L_{z_1}-1}*\\ \dfrac{\frac{R_{z_2}}{L_{z_1}}\left(L_{z_2}-1\right)}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1}*lct_{z_2}+\dfrac{\frac{R_{z_2}}{L_{z_1}}-1}{\frac{R_{z_2}}{L_{z_1}}L_{z_2}-1}*\\ \left(\dfrac{\frac{R_{z_3}}{L_{z_1}L_{z_2}}\left(L_{z_3}-1\right)}{\frac{R_{z_3}}{L_{z_1}L_{z_2}}L_{z_3}-1}*lct_{z_3}+\dfrac{\frac{R_{z_3}}{L_{z_1}L_{z_2}}-1}{\frac{R_{z_3}}{L_{z_1}L_{z_2}}L_{z_3}-1}*0\right)\end{array}\right)\right]*(N-1) \qquad (7.19)$$

Since the total number of experiments $= N = R_{z_1}L_{z_1} = R_{z_2}L_{z_2} = R_{z_3}L_{z_3}$, expression (7.19) for I = 3 simplifies to,

$$R_{z_1}\left(L_{z_1}-1\right)*lct_{z_1}+\frac{R_{z_2}\left(L_{z_2}-1\right)}{L_{z_1}}*lct_{z_2}+\frac{R_{z_3}\left(L_{z_3}-1\right)}{L_{z_1}L_{z_2}}*lct_{z_3} \qquad (7.20)$$

The factor level changing time for the very first run must be added to this expression. For the first experimental run, all factors including z_1 are set resulting in a factor level change time of lct_{z_1}. Therefore,

$$Expected\ total\ level\ changing\ time\ (with\ no\ randomization\ restriction) =$$
$$R_{z_1}\left(L_{z_1}-1\right)*lct_{z_1}+\frac{R_{z_2}\left(L_{z_2}-1\right)}{L_{z_1}}*lct_{z_2}+\frac{R_{z_3}\left(L_{z_3}-1\right)}{L_{z_1}L_{z_2}}*lct_{z_3}+lct_{z_1} \qquad (7.21)$$

Equation (7.21) can be generalized to *I* factors as follows.

$$Expected\ total\ level\ changing\ time\ (TCLT)\ for\ an\ experiment\ with\ I\ factors =$$
$$R_{z_1}\left(L_{z_1}-1\right)*lct_{z_1}+\sum_{i=1}^{I-1}\left(\frac{R_{z_{i+1}}\left(L_{z_{i+1}}-1\right)}{\prod_{j=1}^{i}L_{z_j}}*lci_{z_{i+1}}\right)+lct_{z_1} \qquad (7.22)$$

If it is assumed that factor levels are changed serially and not at the same time, the expected level changing time for factors can be calculated independently form one another. For instance, for the three-factor experiment in the example above, the expected total level changing time can be calculated as follows.

$$(R_1(L_1 - 1) + 1) * lct_1 + (R_2(L_2 - 1) + 1) * lct_2 + (R_3(L_3 - 1) + 1) * lct_3 \quad (7.23)$$

As a result, the expected total level changing time for an experiment with I factors is calculated as

$$\sum_{i=1}^{I} (R_i(L_i - 1) + 1) * lct_i \quad (7.24)$$

7.3 Total Level Changing Cost

Calculating the total factor level changing cost for a balanced randomized experiment is straightforward since the total level changing cost for each factor are independent of the other factors.

$$Expected\ total\ factor\ level\ changing\ cost\ for\ a\ single\ factor\ =$$
$$Expected\ number\ of\ level\ changes\ for\ that\ factor * factor\ level\ change\ cost\ = \quad (7.25)$$
$$(R(L - 1) + 1) * lcc$$

Therefore, the expected total factor level changing cost for an experiment with three factors is:

$$(R_1(L_1 - 1) + 1) * lcc_1 + (R_2(L_2 - 1) + 1) * lcc_2 + (R_3(L_3 - 1) + 1) * lcc_3 \quad (7.26)$$

Equation (7.26) can be easily generalized to the one for I factors.

$$Expected\ total\ level\ changing\ cost\ (TLCC)\ for\ an\ experiment\ with\ I\ factors$$
$$= \sum_{i=1}^{I} (R_i(L_i - 1) + 1) * lcc_i \quad (7.27)$$

Equation (7.27) is also applicable if it is assumed that factor levels are changed serially and not at the same time.

7.4 Recursive Algorithm for Calculating TLCT

The prior results have been obtained under the assumption that the experimental run order is fully randomized. If the experimental design structure is represented as a box, a fully randomized design is represented as a single box since the experimental order makes it possible to move from one treatment combination to any other treatment combination.

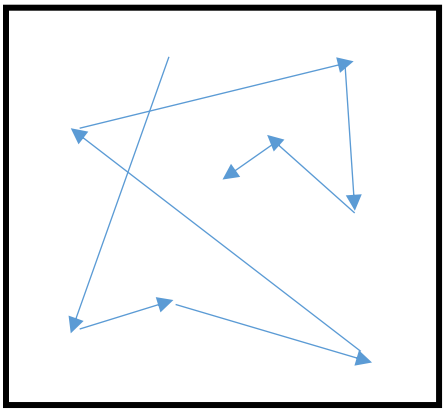


Figure 7.2. A single box representing a fully randomized design structure

In the Figure 7.2, the single box represents the design structure and the arrows show the transitions from one experimental run (treatment combination) to another.

With randomization restrictions as part of the design structure, not all transitions between treatment combinations is possible. This is shown in Figure 7.3 with multiple smaller boxes within the larger box. Within each smaller box, the experimental run order is randomized.

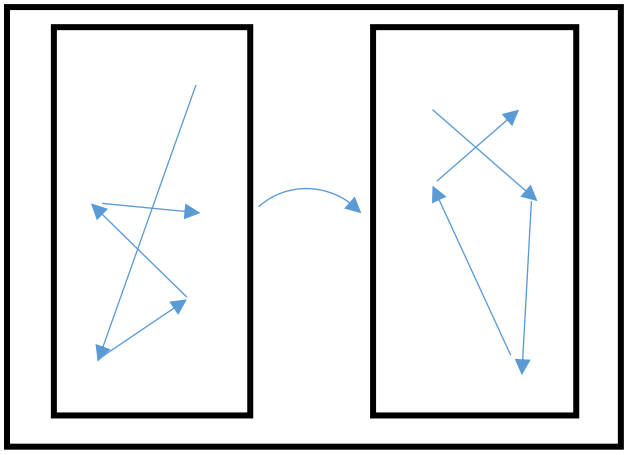


Figure 7.3. Two smaller fully randomized boxes representing a restricted design structure

For the “experiment” represented within each smaller box, the formulas obtained earlier can be applied. The difference is that the number of levels and replications of the factors involved must be revised accordingly.

There can be multiple layers of randomization restrictions that define an experimental design structure. For example, suppose there is another randomization restriction on the levels of another factor. If this randomization restriction is nested within another restriction (e.g., a split-plot experiment) it is represented as smaller fully randomized boxes within the prior randomization restriction. Note that selection of the smaller boxes within each larger box is also randomized.

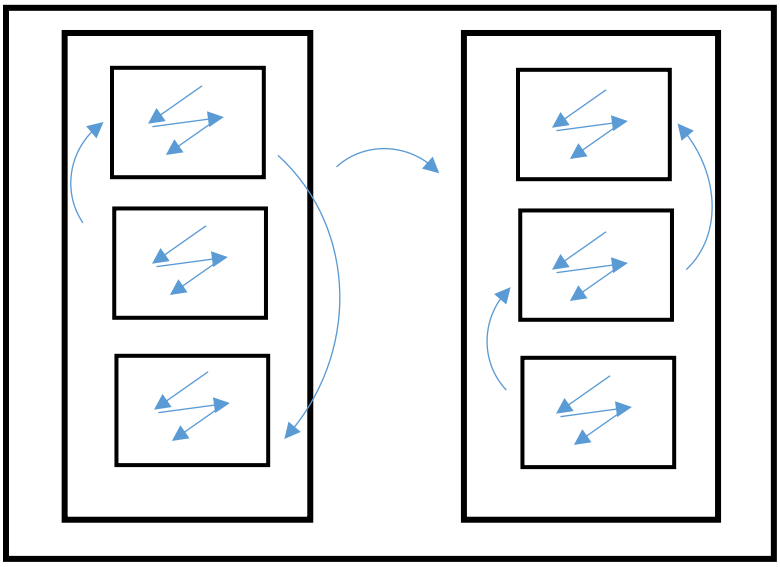


Figure 7.4. A design structure with nested randomization restrictions. The arrows are showing the randomized selection of elements within each box

This can be repeated until no more layer of randomization restrictions is applied. This nested randomization restriction structure suggests a recursive procedure for calculating the TLCT and TLCC.

When the experiments transition from one box to another, however, there is a small probability that the factor level does not change. Consider the example shown in Figure 7.5.

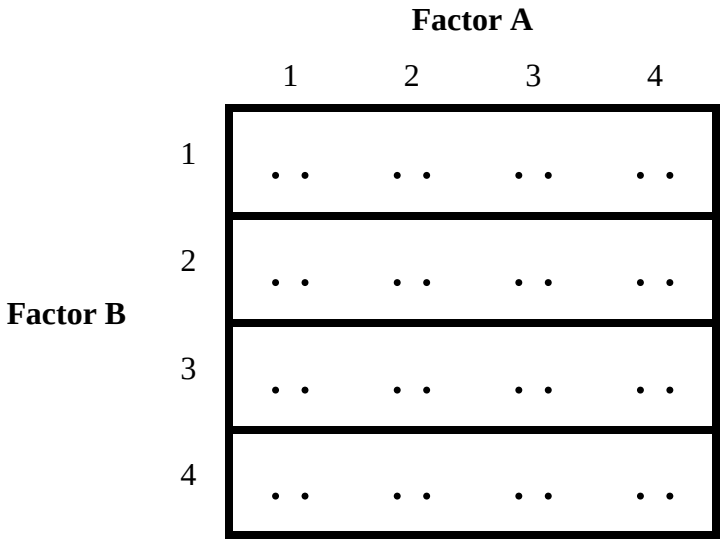


Figure 7.5. The treatment structure is partitioned into four horizontal boxes. There is one fourth chance that factor A does not change level when transitioning from one box to another.

In this design structure, one of the four horizontal boxes is selected randomly. Then, all the eight experiments within that box are run in random order. Then, another horizontal box is selected randomly and this process continues.

Assume that factor A has the larger *lct*. During the transition from one horizontal box to another, there is $\frac{1}{4}$ chance that the level of factor A is not changed since there are four levels of A within each box. And since there are three transitions of this type, the expected number of level changes for factor A will be $\frac{3}{4}$ less than what is calculated using the recursive procedure explained above. To account for this the probability of not changing factor levels when transitioning from one fully randomized box to another, and the expected number of such transitions must be computed at each randomization restriction layer.

Consider the design in Figure 7.6. The experiment is divided into twelve fully randomized "boxes" or smaller treatment structures, so there will be eleven between-box transitions.

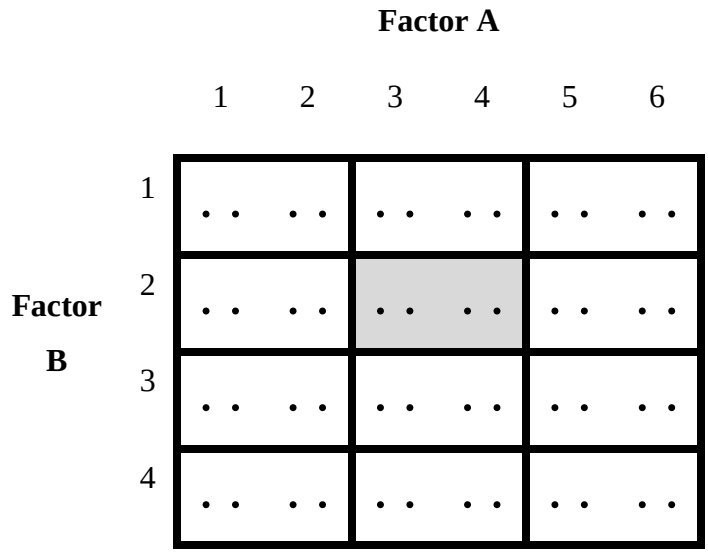


Figure 7.6. Assuming the experiment is at the highlighted box, the probability of factor A not changing level is zero if the next transition is to a box on the left or right column.

If the experimental runs being concluded are in the highlighted box in Figure 7.6, the probability of "factor A not changing level" is equal to zero if the next runs are in a box in another column. Therefore, three out of eleven transitions have a larger than zero probability of no factor A level change.

If the next transition is within the same column this probability is equal to $\frac{1}{number\ of\ levels\ of\ factor\ A\ whitin\ each\ box}$. Since there are two levels of factor A within each box in Figure 8, the probability of "factor A not changing levels" at is $\frac{1}{2}$. Therefore, the overall probability of "factor A not changing levels" for any transition is:

$$\frac{Transtions\ within\ the\ same\ column\ as\ the\ current\ position}{Total\ number\ of\ transitions} * \frac{1}{Levels\ of\ factor\ A\ whitin\ each\ box} = \frac{3}{11} * \frac{1}{2} = \frac{3}{22} \quad (7.28)$$

To find the expected number of times that factor A does not change levels, the probability equation (7.28) has to be multiplied by the "total number of transitions between boxes". Since there are 11 such transitions the expected number of times factor A does not change level is calculated as:

$$\frac{Transtions\ within\ the\ same\ column\ as\ the\ current\ position}{Levels\ of\ factor\ A\ whitin\ each\ box} = 1.5 \qquad (7.29)$$

The "expected total level changing cost" calculation can be adjusted by subtracting the expected cost not incurred since the expected level changing cost for each factor is calculated independently from each other. For each factor at each layer, the "expected number of times of no factor level change" multiplied by the "level changing cost" should be subtracted from the total expected level changing cost.

However, it is more complicated to adjust the "expected total level changing time". In equation (7.21) it is assumed that the very first run within each fully randomized box is always dominated by the factor with the largest level changing time. If the factor with the largest "*lct*" does not change level, then the factor with the second largest "*lct*" might have a level change. If neither changes level, then the factor with the third largest "*lct*" needs to be considered and so on. For an experiment with I factors, the following expected time should be subtracted from the expected level changing time at each randomization restriction layer.

$$\begin{aligned}
&E[z_1\ does\ not\ change\ level\ but\ z_2\ does] * \left(lct_{z_1} - lct_{z_2}\right) \\
&\quad + E[z_1\ and\ z_2\ do\ not\ change\ level\ but\ z_3\ does] * \left(lct_{z_1} - lct_{z_3}\right) \\
&\quad + \cdots \\
&\quad + E[z_1, z_2, \ldots, z_{I-1}\ \ do\ not\ change\ level\ but\ z_I\ does] * \left(lct_{z_1} - lct_{z_I}\right) \\
&\quad + E[None\ of\ the\ factors\ change\ level] * lct_{z_1} \\
&\\
&\quad = E[None\ of\ the\ factors\ change\ level] * lct_{z_1} \\
&+ \sum_{i=2}^{I} E[\ z_1\ to\ z_{i-1}\ do\ not\ change\ level\ but\ z_i\ does] * \left(lct_{z_1} - lct_{z_i}\right)
\end{aligned} \qquad (7.30)$$

The pseudocode shown in Figure 7.7 uses this logic to recursively fine the expected TLCT for an experiment with the treatment structure [TS] and design structure [DS] as inputs.

```
TLCT( [TS] , [DS] )

        if [DS] is empty then                                   //base case
                Randomized_Box_TLCT( [TS] )

        for each factor in [TS]\first_element
                If l_i = 1                                      //fully restricted
                remove this factor from [TS]

        rr = pop from [DS]                                      //[DS] is updated
        multiple = 1

        for each restricted factor in rr
                multiple = multiple*(l_i/restriction size)
                l_i = restriction size                          //[TS] is updated

        for each factor in [TS]
                update r_i                                      //[TS] is updated


        subtraction = use equation (  7.30 )

        if l = 1 for the first_element of [TS] then
                return (multiple-1)* TLCT( [TS]\first_elem , [DS] )
                        +  TLCT( [TS] , [DS] ) - subtraction

        else
                return multiple * TLCT( [TS] , [DS] ) - subtraction
```

```
Randomized_Box_TLCT( [TS] )

        Use equation (  7.22 ) to calculate expected total level changing time
```

Figure 7.7. Pseudocode of the recursive procedure for calculating the TLCT

7.5 The Recursive Algorithm for Calculating TLCC

The procedure for calculating the TLCC is more straightforward since the factor level changing costs are computed for each factor independent of the others. A recursive approach is used here as well. The following is the step by step description of this procedure.

```
TLCC( [TS] , [DS] )

        if [DS] is empty then                                   //base case
                Randomized_Box_TLCC( [TS] )

        constant = 0

        for each factor in [TS]
                If l_i = 1                                      //fully restricted
                constant += lcc_i
                remove this factor from [TS]

        rr = pop from [DS]                                      //[DS] is updated
        multiple = 1

        for each restricted factor in rr
                multiple = multiple*(l_i/restriction size)
                l_i = restriction size                          //[TS] is updated

        for each factor in [TS]
                update r_i                                      //[TS] is updated

        subtraction = 0
        for each factor in [TS]
                subtraction += E[factor i does not change level]*lcc_i

        return constant + multiple * TLCC( [TS] , [DS] ) - subtraction
```

```
Randomized_Box_TLCC( [TS] )

        Use equation (  7.27 ) to calculate expected total level changing cost
```

Figure 7.8. Pseudocode of the recursive procedure for calculating the TLCC

The correctness of this procedure was verified through multiple Monte Carlo simulations of various experimental designs.

8. Evaluating the Power of Design Structures

When considering experimental designs there is a trade-off between the effectiveness of a design with respect to detecting differences from the null hypotheses, and the experimental cost/time. The randomization restriction concept for creating different design structures was introduced (chapter 2), and it was shown that multiple different design structures for a single experimental treatment structure exist (chapter 4). The focus thus far with respect to detecting differences from the null hypotheses has been on identifying for which factors statistical tests in an ANOVA exist, and for which factors the effects are confounded with a randomization restriction.

In this chapter, the evaluation of the design structure power is addressed. In general, more restricted design structures with multiple randomization restriction layers reduce the power of the tests in an ANOVA. To compare the power of different design structures for a given treatment structure, a generalized approach is developed to estimate the power of the tests that includes non-restricted, fully-restricted, or partially-restricted factors.

This chapter starts with a brief overview of the concept of the statistical power. It is then shown how the power of a non-restricted F-test in an experiment is computed. This is followed by the presentation of a general approach for computing the power of restricted F-tests. The chapter concludes with the presentation of a simple method for quantifying the relative *statistical effectiveness* of different design structures, for an experiment with a given treatment structure.

8.1 Statistical Power of Tests

The power of a test of significance to evaluate a statistical hypothesis is the probability of rejecting the corresponding hypothesis when it is false. A null hypothesis is rejected when there is statistically enough evidence to show that there is significant difference between the treatments' means. The failure to reject the null hypothesis when it is false is referred to as "type II error" with standard notation β. Therefore,

$$Power = 1 - \text{"type II error"} = 1 - \beta$$

There are three parameters that influence the statistical power:

1. *Level of significance* or the probability of a Type I error, which is rejecting the null hypothesis when it is true and is denoted α. As the significance level increases the likelihood of rejecting a hypothesis increases and thus power increases. However, α is typically fixed and commonly set to 0.05.

2. *Effect size* or "Δ to detect" which is typically measured as the differences among treatment means in units of standard deviation. Larger effect sizes are easier to detect, and thus result in higher statistical power.

3. *Sample size* or *n*. Increasing the sample size ensures higher degrees of freedom and usually results in lower standard deviations for sample averages. This improves the precision of the test and, therefore, increases the power.

In the next section, it will be shown how power is calculated for the tests of significance in an analysis of variance, where the three parameters described either affect the critical value of the test or the non-centrality parameter of a non-central F distribution.

8.2 Power of F-Tests in ANOVA

The test hypotheses for a factor "A" main effect with "a" levels in an analysis of variance are:

$$\begin{cases} H_0\colon\ \mu_1 = \mu_2 = \cdots = \mu_a = \mu \\ H_1\colon\ Otherwise. \end{cases} \qquad (\ 8.1\ )$$

Where, μ_i is the mean of the *i*th factor level and μ is the overall mean. This test hypothesis can be also represented in terms of treatment effects $A_i = \ \mu_i - \mu$:

$$\begin{cases} H_0\colon\ A_1 = A_2 = \cdots = A_a = 0 \\ H_1\colon\ Otherwise. \end{cases} \qquad (\ 8.2\ )$$

The statistical power for a test of significance in an analysis of variance can be defined as the probability of detecting any possible deviation from the overall mean by one or more factor means. It is the probability of rejecting H_0:

$$Power = \Pr(Reject\ H_0 | H_0\ is\ false\ )$$
$$= \Pr(F_a > F_{crit} | H_0\ is\ false\ ) = 1 - \Pr(F_a < F_{crit} | H_0\ is\ false\ ) = 1 - F_{NC}(F_{crit}) \quad (8.3)$$

Where, F_a is the factor A test statistic and F_{crit} is the critical value, which is the *1-*α^{th} percentile of an *F* distribution with v_1 and v_2 degrees of freedom, and F_{NC} is the cumulative distribution function of a non-central *F* distribution.

According to equation (8.3), to calculate the power the distribution of the test statistic *F* must first be identified. The test statistic is formed as the ratio of mean square of *A* over the mean square error:

$$\frac{MS_A}{MS_E} = \frac{SS_A/(a-1)}{SS_E/(N-a)} \quad (8.4)$$

Where SS_A and SS_E are the sum of squares of treatments and error, with $a-1$ and $N-a$ degrees of freedom, respectively. It was also shown that:

$$E(MS_A) = \sigma^2 + r * \frac{\sum_{i=1}^{a} {A_i}^2}{(a-1)} \quad \text{and} \quad E(MS_E) = \sigma^2 \quad (8.5)$$

Under the null hypothesis all A_i are zero. Thus, both MS_A and MS_E are estimators of σ^2. Thus, the random variables SS_A/σ^2 and SS_E/σ^2 both have a Chi-Square distribution with *a-1* and *N-a* degrees of freedom, respectively. Consequently, the ratio of MS_A/MS_E, shown in equation (8.4), is the ratio of two independent Chi-Square random variables divided by their degrees of freedom. This implies that under the null hypothesis, this ratio has an *F* distribution with *a-1* and *N-a* degrees of freedom.

However, statistical power is the probability of rejecting the null hypothesis when it is false. Under the alternative hypothesis, the treatment effects are <u>*not*</u> all equal to zero and $E(MS_A)$ will be larger than σ^2. This causes the random variable shown in equation (8.4) to have a "<u>non-central</u>" *F*-distribution.

Non-central *F* distributions are specified with three parameters:

- v_1: Numerator degrees of freedom (same as a central *F*)

- $\boldsymbol{v_2}$: Denominator degrees of freedom (same as a central *F*)
- $\boldsymbol{\lambda}$: Non-centrality parameter

The non-centrality parameter (NCP) is a measure of the degree of deviation from the null hypothesis. For a non-central *F* distribution, the NCP is:

$$\lambda = noncentrality\ parameter = r * \frac{\sum_{i=1}^{a}(\mu_i - \mu)^2}{\sigma^2} = r * \frac{\sum_{i=1}^{a} {A_i}^2}{\sigma^2} \quad (8.6)$$

Where *r* is the number of replicates within each treatment. According to the expression in equation (8.6):

- If the treatment effects are zero (H_o is true), the NCP will be equal to zero and the distribution will be a central *F*.
- If the treatment effects are relatively large, the resulting NCP will also be large. Larger values of NCP correspond to higher statistical power since it is easier to detect large deviations from H_o.

8.3 Estimating the Power of F-Tests

In the last section it was shown that,

$$Power = 1 - F_{NC}(F_{crit}) \quad (8.7)$$

Where, F_{NC} is the cumulative distribution of the F_a test statistic under the alternative hypothesis. F_{NC} is a non-central *F* distribution with *a-1* and *N-a* degrees of freedom and the non-centrality parameter of $r * \sum_{i=1}^{a}\left(\frac{A_i}{\sigma}\right)^2$.

In this research, the objective of estimating power of tests of significance is to be able to compare the power of different design structures for a given treatment structure. Power will be computed for specific sources of variation (factors and/or interactions) in an experiment with specific design and treatment structures.

The experimental design and treatment structure along with effect sizes (i.e. A_i/σ) provide the information needed to compute power. The specification of effect sizes is the topic of the following section.

8.3.1 Setting the Effect Size

Statistical power is typically computed for a specific number of replicates and a particular effect size, or Δ. In an analysis of variance, this effect size is defined as:

$$\Delta = \text{maximum pairwise difference between treatment means in units of } \sigma$$

$$= Max\left\{\frac{|\mu_i - \mu_j|}{\sigma}\right\} = Max\left\{\frac{|(\mu_i - \mu) - (\mu_j - \mu)|}{\sigma}\right\} = Max\left\{\frac{|A_i - A_j|}{\sigma}\right\} \quad (8.8)$$

$$for\ all\ i, j\ \ in\ \{1, \dots, a\}$$

For example, for a factor with four levels and treatment effects of 0, -4, -2, and 6 (note that the sum must be equal to zero), the effect size is: $\Delta = {(6-(-4))}/{\sigma} = {10}/{\sigma}$.

It is a common practice to estimate the power of a test at different levels of effect sizes. Although the effect sizes are arbitrary, some authors have suggested effect size levels. Lorenzen and Anderson (1993) for instance, suggested the three effect sizes of 1.5, 3, and 5 to be the low, medium, and high levels of detectability (power) respectively. The values selected for Δ throughout this research will be from this suggested range. Once the value of Δ has been established, the non-centrality parameter can be computed.

8.3.2 Determining the Non-Centrality Parameter for a Selected Effect Size

For a particular Δ, the valid range of values for the NCP is derived as follows:

- The NCP is at its maximum when one half[1] of treatment effects are equal to $\frac{\Delta}{2}$ and the other half are equal to $-\frac{\Delta}{2}$

$$\begin{cases} A_1 = A_2 = \cdots = A_{\frac{a}{2}-1} = \frac{\Delta}{2} \\ A_{\frac{a}{2}} = A_{\frac{a}{2}+1} = \cdots = A_a = -\frac{\Delta}{2} \end{cases} \quad (8.9)$$

[1] When the number of factor levels is odd, one treatment effect is set to zero.

The NCP in this case is:

$$\lambda = r * a * \frac{\Delta^2}{4} \quad (8.10)$$

- The NCP is at its minimum when all but two treatment effects are equal to zero. This is shown as:

$$\begin{cases} A_i = \frac{\Delta}{2} \\ A_j = -\frac{\Delta}{2} \\ all\ other\ A_k = 0 \end{cases} \quad (8.11)$$

The NCP in this case is:

$$\lambda = r * 2 * \frac{\Delta^2}{4} \quad (8.12)$$

For instance, choosing a Δ value of 2, results in the minimum and the maximum NCP of $2*r$ and $a*r$ respectively. Note that in equation (8.12), unlike equation (8.10), the number of levels does not affect λ. Since this research focuses on comparing different experimental designs, λ will be set to its maximum, as shown in equation (8.10) since this includes the number of factor levels.

In the next section, some examples of power calculations for design structures with no, or fully restrictive randomization restrictions are shown. The calculation of power for design structures with partially restrictive randomization restrictions is more complicated and will be explained in additional sections.

8.4 Examples of Fully Randomized and Fully Restrictive Design Structures

In this section, the power calculations for three different design structures for an experiment with the treatment structure of [A: 4, B: 4, (R): 2], is shown. The three design structures are listed:

1. {}: Fully randomized design
2. {(A, 1, 1)}: Fully restricted design
3. {(R, 1, 1), (A, 1, 2)}: "Split-Plot" design structure

A Δ value of 1.5 and an α value of 0.05 are used in all three examples.

8.4.1 Example 1: A Fully Randomized Design

This design structure is illustrated in Figure 8.1.

		Factor A			
		1	**2**	**3**	**4**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.1. Example 1: Fully randomized design structure {}

There are two tests for the main effects, *A* and *B*, and one for the interaction effect *AB*. The corresponding parameters and the power for these three tests are listed in Figure 8.2.

Effect	# of levels	d.f.	Error d.f.	Replicates	*Δ*	NCP	Power
A	4	3	16	8	1.5	18	0.89
B	4	3	16	8	1.5	18	0.89
AB	16	9	16	2	1.5	18	0.63

Figure 8.2. Power for example one.

Since there is no randomization restriction in the design, and since this is the most expensive way to run this experiment, the power values are the highest possible for this treatment structure.

8.4.2 Example 2: A Fully Restricted Design

The second design is shown in Figure 8.3.

		Factor A			
		1	**2**	**3**	**4**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.3. Example 2: Fully restricted design structure {(A, 1, 1)}.

The power values are,

Effect	# of levels	d.f.	Error d.f.	Replicates	*Δ*	NCP	Power
A	4	3	0	8	1.5	18	0.0
B	4	3	16	8	1.5	18	0.89
AB	16	9	16	2	1.5	18	0.63

Figure 8.4. Power for example two.

Since there is no test for *A* since it confounded with the randomization restriction, there are no error degrees of freedom and hence the statistical power is set to zero.

8.4.3 Example 3: A "Split-Plot" Design

This split-plot design is depicted in Figure 8.5.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Factor B	**1**	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 8.5. Example 3: Split-plot design structure {(R, 1, 1), (A, 1, 2)}

The power results are shown below.

Effect	# of levels	d.f.	Error d.f.	Replicates	Δ	NCP	Power
A	4	3	3	8	1.5	18	0.46
B	4	3	12	8	1.5	18	0.87
AB	16	9	12	2	1.5	18	0.56

Figure 8.6. Power for example three.

From this table, it is evident that the overall experimental power has dropped compared to the fully randomized experiment. A better comparison is the power between the split-plot design structure and {(R, 1, 1), (A, 1, 1)} (shown below). This design structure is slightly less restrictive than the split-plot and adds one error degree of freedom to *A*.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 8.7. Design structure {(R, 1, 1), (A, 1, 1)}

Effect	# of levels	d.f.	Error d.f.	Replicates	Δ	NCP	Power
A	4	3	4	8	1.5	18	0.58
B	4	3	12	8	1.5	18	0.87
AB	16	9	12	2	1.5	18	0.56

Figure 8.8. Power of {(R, 1, 1), (A, 1, 1)}.

Because of this change in the design structure, the statistical power of *A* was increased from 0.46 to 0.58, but the experiment is more expensive to run.

8.5 Estimating Power of Partially Restrictive Designs

In the previous sections, a general approach for computing the power of different tests of significance in an analysis of variance was discussed. The design structures analyzed had

either no or only fully restrictive randomization restrictions. The estimation of power in designs with partially restrictive randomization restrictions requires a few additional steps which will be discussed in this section.

It was explained in chapter 6 that the analysis of partially restrictive design structures is completed by splitting the partially restricted factors into multiple sources of variation. This separates the test of significance for the partially restricted factor into a set of multiple independent tests to isolate the restricted parts of the factor from the unrestricted parts. This allows the experimenter to obtain available information from the partially restricted factor.

Computing power for a partially restricted test will be more challenging because of the multiple independent sub-tests. This is because comparing the level of power of different design structures requires only one statistical power value for each factor in the model. Therefore, the first step is to derive the exact correspondence between each partially restricted test of significance and its sub-tests.

8.5.1 Estimating Power Through the Transformation of Treatment Structures

Consider an experiment with the treatment structure of [A: 4, B: 4, (R): 2] (the same treatment structure used for the three examples in section 8.4), and a design structure of {(A, 2, 1)}. This design is depicted in Figure 8.9.

		Factor A			
		1	**2**	**3**	**4**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.9. The partially restrictive design {(A,2,1)}

To make a partial test for the factor *A*, the treatment structure of this experiment needs to be transformed to a nested structure in which the variation explained by *A*, as well as its degrees of freedom are split into two sources of variation.

As a refresher of the analysis of partially restricted designs, the transformation of the treatment structure, along with other underlying changes in assumptions and treatment effects, are summarized in Figure 8.10.

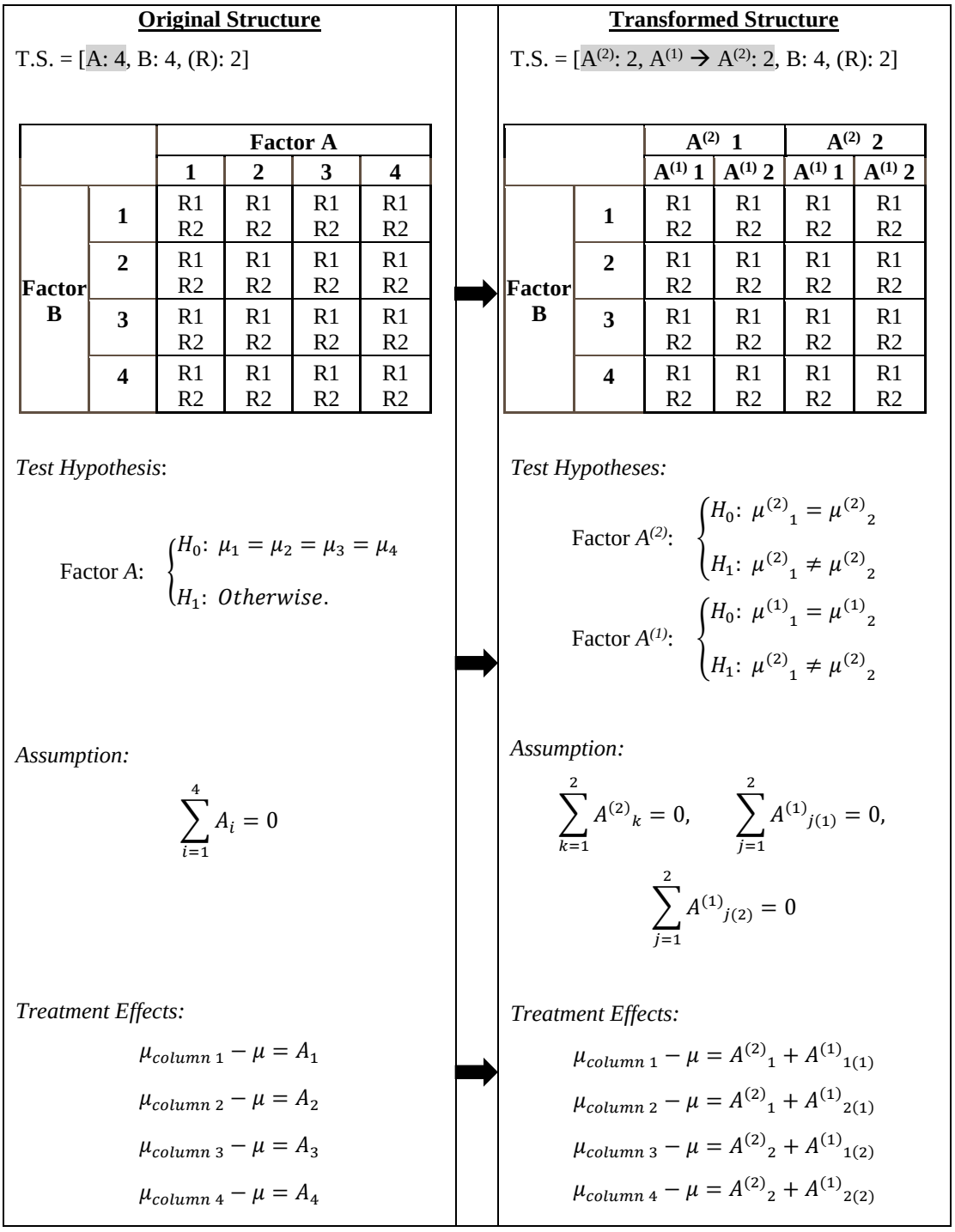


Figure 8.10. Transformation of the treatment structure in the partially restrictive design structure {(A,2,1)}

Figure 8.10 shows the two test hypotheses that replace the single original test hypothesis for factor *A*. The original null hypothesis (with no restrictions) is rejected, if and only if

one of the new null hypotheses on the right is rejected (the significance level α has been reduced to $\frac{\alpha}{2}$ for the new test hypotheses). Therefore,

$$\begin{aligned} Power &= \Pr(rejecting\ H_0\ for\ A|H_1) = \Pr\left(F_a > F_{\alpha,3,16}\right) = 1 - F_{NC}(F_{\alpha,3,16}) \\ &= \Pr\left((rejecting\ H_0\ for\ A^{(2)}|H_1\ )\ OR\ (rejecting\ H_0\ for\ A^{(1)}|H_1\ )\right) \\ &= \Pr\left(\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\right) \cup \left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\right)\right) \end{aligned} \qquad (8.13)$$

F_a, $F_a{}^{(2)}$, and $F_a{}^{(1)}$ are the test statistics used to test for significant factor *A*, $A^{(1)}$, and $A^{(2)}$ effects, respectively. As equation (8.13) shows, the key element to compute the overall power is computing the power of individual sub-tests $power^{(2)}$ and $power^{(1)}$. To estimate the power of the sub-tests, the corresponding non-centrality parameters (indicated with appropriate superscripts) must be calculated. For this treatment structure,

$$\begin{aligned} NCP^{(2)} &= r^{(2)} * \sum_{k=1}^{a^{(2)}} \left(\frac{A^{(2)}{}_k}{\sigma}\right)^2 = 16 * \sum_{k=1}^{2} \left(\frac{A^{(2)}{}_k}{\sigma}\right)^2 \\ NCP^{(1)} &= r^{(1)} * \sum_{k=1}^{a^{(2)}} \sum_{j=1}^{a^{(1)}} \left(\frac{A^{(1)}{}_{j(k)}}{\sigma}\right)^2 = 8 * \sum_{k=1}^{2} \sum_{j=1}^{2} \left(\frac{A^{(1)}{}_{j(k)}}{\sigma}\right)^2 \end{aligned} \qquad (8.14)$$

From equation (8.14), the last step in estimating the sub-tests is to derive the values of $A^{(2)}{}_i$ and $A^{(1)}{}_{j(i)}$ in accordance with values of A_i. This relationship is shown in Figure 8.11. Relationship between original treatment effects and the new effectsFigure 8.11.

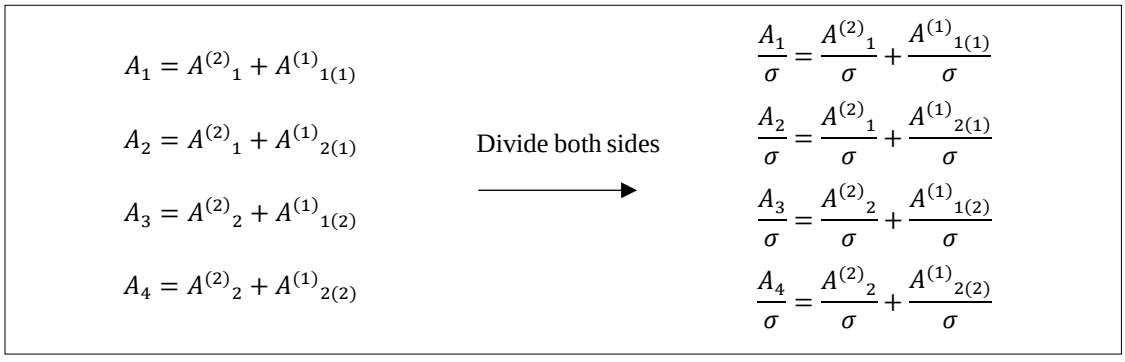


Figure 8.11. Relationship between original treatment effects and the new effects

When the NCP is at its maximum, half of the A_i are set to $\frac{\Delta}{2}$, and the other half to $\frac{-\Delta}{2}$.

For simplicity let Δ to be equal to 2. With $\Delta = 2$ the $\frac{A_i}{\sigma}$ are either 1 or -1. However, different possible permutations of 1 or -1 to the $\frac{A_i}{\sigma}$ results in different values of $\frac{A^{(2)}{}_k}{\sigma}$ and $\frac{A^{(1)}{}_{j(k)}}{\sigma}$, and therefore, results in different values of $NCP^{(2)}$ and $NCP^{(1)}$. All the permutation of these assignments, for the transformation shown in Figure 8.10, are shown in Figure 8.12.

Permutation	Source	Treatment Size	Level 1	Level 2	Level 3	Level 4	Sum Square	NCP
1	A	A_i/σ	1	1	-1	-1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	1		-1		2	32
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	0	0	0	0
2	A	A_i/σ	1	-1	1	-1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	1	-1	4	32
3	A	A_i/σ	1	-1	-1	1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	-1	1	4	32
4	A	A_i/σ	-1	-1	1	1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	-1		1		2	32
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	0	0	0	0
5	A	A_i/σ	-1	1	-1	1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	-1	1	-1	1	4	32
6	A	A_i/σ	-1	1	1	-1	4	32
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	-1	1	1	-1	4	32

Figure 8.12. Different permutation of treatment effects' assignments and the resulting NCP's

Note that the total number of permutations is equal to the number of possibilities of assigning 1 to half of the treatment effects and -1 to the other half. Therefore, for a factor with "a" levels the total number of permutations is $\binom{a}{a/2}$.

The frequency of $\frac{A^{(2)}{}_k}{\sigma}$ and $\frac{A^{(1)}{}_{j(k)}}{\sigma}$ values and the resulting non-centrality values are shown in Figure 8.13.

Frequency	Treatment Size	NCP
2	$A^{(2)}{}_k/\sigma$	32
	$A^{(1)}{}_{j(k)}/\sigma$	0
4	$A^{(2)}{}_k/\sigma$	0
	$A^{(1)}{}_{j(k)}/\sigma$	32

Figure 8.13. Frequency of different permutations as a result of the design structure {(A, 2, 1)}

Using the expression shown in equation (8.13), and by incorporating the role of different permutations in calculating the NCP's, the overall power for the main effect *A* can be expressed as a conditional probability as follows:

$$\begin{aligned} Power &= \Pr\left(\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\right) \cup \left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\right)\right) \\ &= \sum_{\substack{all\ permutations \\ p_i}} \Pr\left(\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\right) \cup \left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\right)\middle| p_i\right) * Pr(p_i) \end{aligned} \qquad (8.15)$$

The hypothesis tests resulting from the sum of squares and degrees of freedom into "*a-1*" orthogonal contrasts are independent from each other (chapter 6). Using this property, equation (8.15) can be simplified:

$$\begin{aligned} Power &= \textstyle\sum_{\substack{all\ permutations \\ p_i}} \Pr\left(\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\right) \cup \left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\right)\middle| p_i\right) * Pr(p_i) \\ &= \sum_{\substack{all\ permutations \\ p_i}} \begin{pmatrix} \Pr\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\middle| p_i\right) + \Pr\left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\middle| p_i\right) \\ -\Pr\left(F_a{}^{(2)} > F_{\frac{\alpha}{2},1,16}\middle| p_i\right) * \Pr\left(F_a{}^{(1)} > F_{\frac{\alpha}{2},2,16}\middle| p_i\right) \end{pmatrix} * Pr(p_i) \\ &= \sum_{\substack{all \\ permutations \\ p_i}} \left(\left(Power^{(2)}|p_i\right) + \left(Power^{(1)}|p_i\right) - \left(Power^{(2)}|p_i\right) * \left(Power^{(1)}|p_i\right)\right) * Pr(p_i) \end{aligned} \qquad (8.16)$$

Since all factor effect permutations are equally likely (Figure 8.13), the overall power for the main effect *A* is calculated as:

$$Power = \left(\frac{2}{6}\right) * \left(\left(Power^{(2)}|p_1\right) + \left(Power^{(1)}|p_1\right) - \left(Power^{(2)}|p_1\right) * \left(Power^{(1)}|p_1\right)\right) + \left(\frac{4}{6}\right) * \left(\left(Power^{(2)}|p_2\right) + \left(Power^{(1)}|p_2\right) - \left(Power^{(2)}|p_2\right) * \left(Power^{(1)}|p_2\right)\right) \quad (8.17)$$

8.5.2 More Examples of Power Calculations

Two additional examples of power calculations for partially restricted design structures are presented in this section to demonstrate the concept of different effect permutations to compute power. Figure 8.14 and Figure 8.15 illustrate two transformations of treatment structure [A: 6, B: 4, (R): 2] to the treatment structures [$A^{(3)}$: 2, $A^{(1)} \rightarrow A^{(3)}$: 3, B: 4, (R): 2] and [$A^{(2)}$: 3, $A^{(1)} \rightarrow A^{(2)}$: 2, B: 4, (R): 2], respectively. These figures only show one half of the total effect permutations. The other half is obtained by multiplying these permutations by -1, and therefore result in the same sum of squares and NCP frequency ratios.

		Factor A					
		1	**2**	**3**	**4**	**5**	**6**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

		$A^{(2)}$ 1			$A^{(2)}$ 2		
		$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 3	$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 3
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Permutation	Source	Treatment Size	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Sum Square	NCP
1	A	A_i/σ	1	1	1	-1	-1	1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	1			-1			2	48
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0	0	0	0	0	0	0	0
2	A	A_i/σ	1	1	-1	1	-1	-1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.66	0.6	-1.33	1.33	-0.66	-0.66	5.33	42.66
3	A	A_i/σ	1	1	-1	-1	1	-1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.66	0.66	-1.33	-0.66	1.33	-0.66	5.33	42.66
4	A	A_i/σ	1	1	-1	-1	-1	1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.66	0.66	-1.33	-0.66	-0.66	1.33	5.33	42.66
5	A	A_i/σ	1	-1	1	1	-1	-1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.67	-1.33	0.67	1.33	-0.67	-0.67	5.33	42.66
6	A	A_i/σ	1	-1	1	-1	1	-1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.67	-1.33	0.67	-0.67	1.33	-0.67	5.33	42.66
7	A	A_i/σ	1	-1	1	-1	-1	1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	0.3333			-0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	0.67	-1.33	0.67	-0.67	-0.67	1.33	5.33	42.66
8	A	A_i/σ	1	-1	-1	1	1	-1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	-0.3333			0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	1.33	-0.67	-0.67	0.67	0.67	-1.33	5.33	42.66
9	A	A_i/σ	1	-1	-1	1	-1	1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	-0.3333			0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	1.33	-0.67	-0.67	0.67	-1.33	0.67	5.33	42.66
10	A	A_i/σ	1	-1	-1	-1	1	1	6	48
	$A^{(3)}$	${A^{(3)}}_k/\sigma$	-0.3333			0.3333			0.22	5.33
	$A^{(1)}$	${A^{(1)}}_{j(k)}/\sigma$	1.33	-0.67	-0.67	-1.33	0.67	0.67	5.33	42.66

$$Power = \left(\frac{1}{10}\right) * \left(\left(Power^{(2)}|p_1\right) + \left(Power^{(1)}|p_1\right) - \left(Power^{(2)}|p_1\right) * \left(Power^{(1)}|p_1\right)\right) + \left(\frac{9}{10}\right) * \left(\left(Power^{(2)}|p_2\right) + \left(Power^{(1)}|p_2\right) - \left(Power^{(2)}|p_2\right) * \left(Power^{(1)}|p_2\right)\right)$$

Figure 8.14. Transformation of treatment structure [A: 6, B: 4, (R): 2] to [$A^{(3)}$: 2, $A^{(1)}$ → $A^{(3)}$: 3, B: 4, (R): 2]

		Factor A					
		1	2	3	4	5	6
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

		$A^{(2)}$ 1		$A^{(2)}$ 2		$A^{(2)}$ 3	
		$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2	$A^{(1)}$ 1	$A^{(1)}$ 2
Factor B	1	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Permutation	Source	Treatment Size	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Sum Square	NCP
1	A	A_i/σ	1	1	1	-1	-1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	1		0		-1		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	1	-1	0	0	2	16
2	A	A_i/σ	1	1	-1	1	-1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	1		0		-1		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	-1	1	0	0	2	16
3	A	A_i/σ	1	1	-1	-1	1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	1		-1		0		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	0	0	1	-1	2	16
4	A	A_i/σ	1	1	-1	-1	-1	1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	1		-1		0		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	0	0	0	0	-1	1	2	16
5	A	A_i/σ	1	-1	1	1	-1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		1		-1		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	0	0	0	0	2	16
6	A	A_i/σ	1	-1	1	-1	1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	1	-1	1	-1	6	48
7	A	A_i/σ	1	-1	1	-1	-1	1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	1	-1	-1	1	6	48
8	A	A_i/σ	1	-1	-1	1	1	-1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	-1	1	1	-1	6	48
9	A	A_i/σ	1	-1	-1	1	-1	1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		0		0		0	0
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	-1	1	-1	1	6	48
10	A	A_i/σ	1	-1	-1	-1	1	1	6	48
	$A^{(2)}$	$A^{(2)}{}_k/\sigma$	0		-1		1		2	24
	$A^{(1)}$	$A^{(1)}{}_{j(k)}/\sigma$	1	-1	0	0	0	0	2	16

$$Power = \left(\frac{4}{10}\right) * \left(\left(Power^{(2)}|p_1\right) + \left(Power^{(1)}|p_1\right) - \left(Power^{(2)}|p_1\right) * \left(Power^{(1)}|p_1\right)\right) + \left(\frac{6}{10}\right) * \left(\left(Power^{(2)}|p_2\right) + \left(Power^{(1)}|p_2\right) - \left(Power^{(2)}|p_2\right) * \left(Power^{(1)}|p_2\right)\right)$$

Figure 8.15. Transformation of treatment structure [A: 6, B: 4, (R): 2] to [$A^{(2)}$: 3, $A^{(1)}$ → $A^{(2)}$: 2, B: 4, (R): 2]

8.5.3 Generalization of Power Calculations for Partially Restrictive Design Structures

The procedure of estimating the overall power of a source of variation (factors and interactions) can be generalized to design structures with multi-layered randomization restrictions. Consider an experiment with the treatment structure [A: 8, (R): 4]. Figure 8.16 shows some of the transformation of this treatment structure to [$A^{(4)}$: 2, $A^{(2)}$ → $A^{(4)}$: 2, $A^{(1)}$ → $A^{(2)}$: 2, (R): 2].

Treatment structure transformation:

Factor A							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
R1	R1	R1	R1	R1	R1	R1	R1
R2	R2	R2	R2	R2	R2	R2	R2
R3	R3	R3	R3	R3	R3	R3	R3
R4	R4	R4	R4	R4	R4	R4	R4

$A^{(4)}$ 1				**$A^{(4)}$ 2**			
$A^{(2)}$ 1		**$A^{(2)}$ 2**		**$A^{(2)}$ 1**		**$A^{(2)}$ 2**	
$A^{(1)}$ 1	**$A^{(1)}$ 2**	**$A^{(1)}$ 1**	**$A^{(1)}$ 2**	**$A^{(1)}$ 1**	**$A^{(1)}$ 2**	**$A^{(1)}$ 1**	**$A^{(1)}$ 2**
R1	R1	R1	R1	R1	R1	R1	R1
R2	R2	R2	R2	R2	R2	R2	R2
R3	R3	R3	R3	R3	R3	R3	R3
R4	R4	R4	R4	R4	R4	R4	R4

Treatment effects transformation:

$$A_i = A_l{}^{(4)} + A_{k(l)}{}^{(2)} + A_{j(k(l))}{}^{(1)}$$

Effect permutations:

$\frac{A_1}{\sigma}$	$\frac{A_2}{\sigma}$	$\frac{A_3}{\sigma}$	$\frac{A_4}{\sigma}$	$\frac{A_5}{\sigma}$	$\frac{A_6}{\sigma}$	$\frac{A_7}{\sigma}$	$\frac{A_8}{\sigma}$
$\frac{A^{(4)}{}_1}{\sigma}$				$\frac{A^{(4)}{}_2}{\sigma}$			
$\frac{A^{(2)}{}_{1(1)}}{\sigma}$		$\frac{A^{(2)}{}_{2(1)}}{\sigma}$		$\frac{A^{(2)}{}_{1(2)}}{\sigma}$		$\frac{A^{(2)}{}_{2(2)}}{\sigma}$	
$\frac{A^{(1)}{}_{1(1(1))}}{\sigma}$	$\frac{A^{(1)}{}_{2(1(1))}}{\sigma}$	$\frac{A^{(1)}{}_{1(2(1))}}{\sigma}$	$\frac{A^{(1)}{}_{2(2(1))}}{\sigma}$	$\frac{A^{(1)}{}_{1(1(2))}}{\sigma}$	$\frac{A^{(1)}{}_{2(1(2))}}{\sigma}$	$\frac{A^{(1)}{}_{1(2(2))}}{\sigma}$	$\frac{A^{(1)}{}_{2(2(2))}}{\sigma}$

Power:

$$\sum_{\substack{all\ permutations \\ p_i}} Pr\left(reject\ A^{(4)} \cup reject\ A^{(2)} \cup reject\ A^{(1)} \middle| p_i\right) * Pr(p_i)$$

Figure 8.16. More general power calculation through the transformation of treatment structures

8.5.4 Interaction Effects

The transformation of a treatment structure always changes the structure of corresponding interaction effects as well. For example, the transformation of the treatment structure shown in Figure 8.10 is,

$$[A: 4, B: 4, (R): 2] \quad ==> \quad [A^{(2)}: 2, A^{(1)} \rightarrow A^{(2)}: 2, B: 4, (R): 2] \qquad (8.18)$$

Factor *A* was replaced by factors $A^{(2)}$ and $A^{(1)}$. This transformation implies that the *AB* interaction is replaced with $A^{(2)}B$ and $A^{(1)}B$. However, the procedure of deriving the overall power for the interaction effect *AB* would be no different than for main effects.

As a continuation to the three examples of fully randomized or fully restricted design structures shown in section 8.4, the next section provides the power for different factors and interactions in four partially restricted design structures.

8.6 Examples of Partially Restrictive Design Structures

The first two examples have the same treatment structure as the three examples in section 8.4 denoted [A: 4, B: 4, (R): 2]. The two design structures used for these examples are:

4. {(A, 2, 1)}
5. {(R, 1, 1), (A, 2, 2), (B, 2, 2)}

To be consistent with section 8.4, the value of Δ is set to 1.5 in these first two examples.

The other two examples are experiments with the treatment structure of [A: 6, B: 4, (R): 2]. The two design structures used in these examples are:

6. {(A, 3, 1)}
7. {(A, 2, 1)}

The power is computed for these two design structures using a Δ of 1.

8.6.1 Example 4: Transformation of [A:4] to [$A^{(2)}$: 2, $A^{(1)} \rightarrow A^{(2)}$: 2]

This design structure is illustrated in Figure 8.17.

		Factor A			
		1	**2**	**3**	**4**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.17. Example 1: Design structure {(A, 2, 1)}

Figure 8.18 shows the power computed for this design. Note that the maximum statistical power for factors *A*, *B*, and the interaction *AB* in a fully randomized design were 0.89, 0.89, and 0.63.

Effect	**# of levels**	**d.f.**	***Δ***	**Power**
A	4	3	1.5	0.60
B	4	3	1.5	0.90
AB	16	9	1.5	0.63

Figure 8.18. Power calculations for example four.

Since the error degrees of freedom, number of replicates, and NCP are not directly defined for the factor *A*, these tables are summarized into the above four columns only.

8.6.2 Example 5: Transformation of [A:4, B:4] to [$A^{(2)}$: 2, $A^{(1)} \rightarrow A^{(2)}$: 2, $B^{(2)}$: 2, $B^{(1)} \rightarrow B^{(2)}$: 2]

The design structure for this partially restrictive design is depicted in Figure 8.19.

		Replicate 1				Replicate 2			
		Factor A				Factor A			
		1	2	3	4	1	2	3	4
Factor B	1	R1	R1	R1	R1	R2	R2	R2	R2
	2	R1	R1	R1	R1	R2	R2	R2	R2
	3	R1	R1	R1	R1	R2	R2	R2	R2
	4	R1	R1	R1	R1	R2	R2	R2	R2

Figure 8.19. Example 5. Partially restrictive design structure {(R, 1, 1), (A, 2, 2), (B, 2, 2)}

The power for the factors and interaction in this experiment are:

Effect	# of levels	d.f.	Δ	Power
A	4	3	1.5	0.77
B	4	3	1.5	0.77
AB	16	9	1.5	0.62

Figure 8.20. Power estimation for example five.

Comparing the power of this design with examples 1 and 3, the effectiveness of this design resides somewhere between the fully randomized and the split-plot design (example 3).

8.6.3 Example 6: Transformation of [A:6] to [$A^{(3)}$: 2, $A^{(1)} \rightarrow A^{(3)}$: 3]

The next two examples compare two slightly different design structures for the treatment structure [A: 6, B: 4, (R): 2]. The maximum possible statistical power for factors *A* and *B*, and the interaction *AB* using a Δ of 1 in a completely randomized design is:

Effect	# of levels	d.f.	Δ	Power
A	6	5	1	0.65
B	4	3	1	0.77
AB	24	15	1	0.35

Figure 8.21. Maximum statistical power for treatment structure [A: 6, B: 4, (R): 2]

The first, and less restrictive, design structure represented as {(A, 3, 1)} is depicted in Figure 8.22.

		Factor A					
		1	**2**	**3**	**4**	**5**	**6**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.22. Example 6. Design structure {(A, 3, 1)}

The power for this design is:

Effect	# of levels	d.f.	Δ	Power
A	6	5	1	0.47
B	4	3	1	0.76
AB	24	15	1	0.34

Figure 8.23. Power for example 6.

8.6.4 Example 7: Transformation of [A:6] to [$A^{(2)}$: 3, $A^{(1)} \rightarrow A^{(2)}$: 2]

The design structure {(A, 2, 1)} is depicted in Figure 8.24.

		Factor A					
		1	**2**	**3**	**4**	**5**	**6**
Factor B	**1**	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	3	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	4	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Figure 8.24. Example 7. Design structure {(A, 2, 1)}

The power is shown in Figure 8.25.

Effect	# of levels	d.f.	Δ	Power
A	6	5	1	0.39
B	4	3	1	0.76
AB	24	15	1	0.34

Figure 8.25. Power for example 7.

The trade-off between the efficiency and effectiveness of the design structures is demonstrated in examples 6 and 7. The less restrictive design structure 6 shows higher power for factor *A*. The more restrictive, and also more efficient, design structure 7 shows less statistical power for factor *A*.

8.7 Overall Effectiveness of Design Structures

In this chapter, a procedure was developed to compute the statistical power of all significance tests in an experiment with any type of balanced design structure. In partially restrictive design structures it was shown how power can be calculated by transforming the overall treatment structure.

The objective of this chapter has been to compare the overall effectiveness of different design structures given a specific treatment structure. Judging the effectiveness of a design structure is subjective and depends on the experimenters' priorities. In this section, a simple approach is presented as one method to compute a single effectiveness value for each design structure as a function of the power to detect effects for factors and interactions.

This effectiveness function is a weighted sum of statistical power for factors and interactions.

$$Effectiveness = \sum_{\substack{For\ all\ factors \\ and\ interactions\ S}} \frac{1}{IL_S} * Power_S \qquad (8.19)$$

Where the *IL* is the level of interaction for *S*. For instance, *IL* is one for factor *A*, two for the two-factor interaction *AB*, three for the three-factor interaction *ABC*, and so on. Since design structures tend to have larger number of interactions, this weighted approach guarantees larger values of effectiveness for design structures with high statistical power

for the main factors over design structures with lower statistical power for the main effects but higher statistical power for the interactions.

To better see the use of this effectiveness measure, an example of selecting the most efficient and effective design structures out of all possible design structures for a specific treatment structure is presented in the next chapter.

9. A Case Study

In the previous chapter, an algorithmic approach was developed to estimate the statistical power for detecting main and interaction effects in an experiment with a specific treatment and design structure. In the last section of chapter 8, a simple approach for measuring the overall statistical effectiveness of a design structure was proposed (equation (9.1)), which is a weighted sum of individual statistical powers for the main and interaction effects with weights that consider the interaction level.

$$Effectiveness = \sum_{\substack{For\ all\ sources \\ of\ variation\ S}} \frac{1}{IL_S} * Power_S \qquad (9.1)$$

Equation (9.1) is a measure of the statistical effectiveness of a design structure. Equation (9.2) shows the closed-form formula derived in chapter 7 for calculating the expected time required to run an experiment in a fully randomized order.

$$Expected\ total\ level\ changing\ time\ (TCLT)\ for\ an\ experiment\ with\ I\ factors = R_{z_1}(L_{z_1} - 1) * lct_{z_1} + \sum_{i=1}^{I-1} (\frac{R_{z_{i+1}}(L_{z_{i+1}} - 1)}{\prod_{j=1}^{i} L_{z_j}} * lci_{z_{i+1}}) + lct_{z_1} \qquad (9.2)$$

In the above equation, R_{z_i}, L_{z_i}, and lct_{z_i} are number of replicates, levels, and the level changing time for the factor with i^{th} largest level changing time, respectively. This formula was then incorporated into a recursive algorithm to calculate the overall expected time of a design structure with a multi-layered set of randomization restrictions.

The objective of this chapter is to present a case study utilizing the effectiveness and efficiency measures (equations (9.1) and (9.2)) to identify design structures ranging from the most effective to the most efficient. The process of selecting a design structure illustrates the trade-off between these two often conflicting measures. The most statistically effective design structure can be selected for a given upper bound on the amount of time or resources available.

In the next section, a modified version of the fuel consumption experiment first discussed in the introduction is presented. This experiment will be analyzed as the case study.

9.1 Engine Fuel Consumption Experiment

This experiment was first introduced in chapter one as an investigation of fuel consumption for a specific type of vehicle engine. The two experimental factors and their corresponding levels are:

- Oil Brand (factor A):
  - Valvoline
  - Total
  - Pennzoil
  - Castrol
  - Amsoil
  - Mobil 1
- Fuel Brand (factor B):
  - Chevron
  - Costco
  - Exxon
  - Mobil
  - QuickTrip
  - Shell

With two replicates for each treatment combination the treatment structure of this experiment is represented as:

$$[A\text{: } 6,\ B\text{: } 6,\ (R)\text{: } 2]$$

Changing the oil or fuel from one type to another requires purging of the prior oil or fuel. Purging oil is harder than fuel and takes 15 minutes. Purging fuel takes 5 minutes. The two purges can be done in parallel so both fuel and oil purges can be completed in 15 minutes.

The first step to this case study is to illustrate the analysis of a fully randomized experiment with respect to expected running time and estimated effectiveness.

9.2 Fully Randomized Design Structure

The diagram of a fully randomized design structure for the fuel consumption experiment is depicted in Figure 9.1.

		Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Total	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Pennzoil	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Castrol	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Amsoil	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Mobil 1	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

Statistical model:

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, \dots, 6 \\ j = 1, \dots, 6 \\ k = 1, 2 \end{cases}$$

Figure 9.1. Fully randomize design for the treatment structure [*A*: 6, *B*: 6, (*R*): 2]

The expected time needed to fully run the experiment[2] can be calculated as:

$$Expected\ Time\ = R_{oil}(L_{oil} - 1) * lct_{oil} + \frac{R_{fuel}(L_{fuel} - 1)}{L_{oil}} * lci_{fuel} + lct_{oil} \qquad (9.3)$$

$$= 12(6 - 1) * 15 + \frac{12(6 - 1)}{6} * 5 + 15 = 965\ minutes$$

With a Δ value of 1, the statistical power for the experiment is:

Effect	# of levels	d.f.	Error d.f.	Replicates	*Δ*	NCP	Power
Oil	6	5	36	12	1	18	0.88
Fuel	6	5	36	12	1	18	0.88
Oil.Fuel	36	25	36	2	1	18	0.43

Figure 9.2. Power for the fully randomized design structure

The overall statistical effectiveness for this fully randomized design structure is:

[2] This is only in terms of level changing times. Since the total number of experimental runs is the same among all design structures, the time to running each experiment is not considered.

$$Effectiveness = 1 * 0.88 + 1 * 0.88 + \frac{1}{2} * 0.43 = 1.97 \quad (9.4)$$

The results of equations (9.3) and (9.4) indicate that:

- The estimated effectiveness of 1.97 is the maximum effectiveness that can be achieved for this treatment structure.
- The estimated time of 965 minutes is a lower-bound efficiency for this treatment structure. Design structures costing more than 965 minutes should not be considered further.

In the next section, using the recursive algorithm developed in chapter 4, all possible design structures for this treatment structure will be generated and their corresponding efficiency and effectiveness will be analyzed.

9.3 Enumerating All Possible Design Structures

Using the recursive algorithm developed in chapter 4, there are 255 different design structures that can be generated for the treatment structure of [A: 6, B: 6, (R): 2]. Figure 9.3 shows the scatter plot of expected time versus statistical effectiveness for all these design structures.

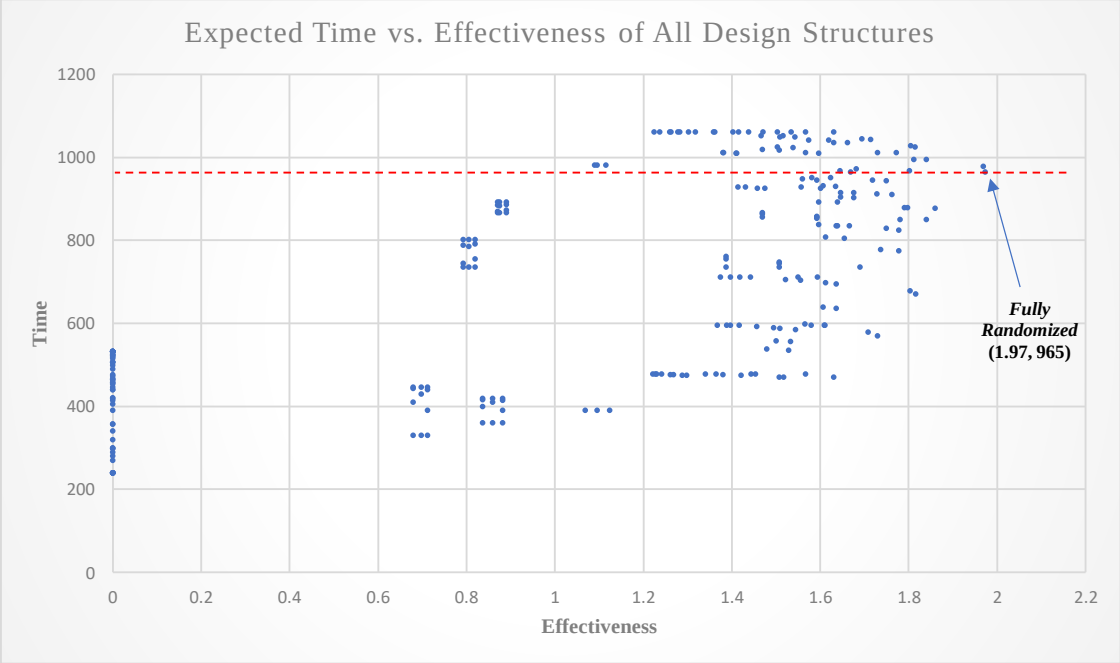


Figure 9.3. Efficiency and Effectiveness for all design structures

The expected time to run a fully randomized fuel consumption experiment is 965 minutes. Although the fully randomized design structure obtains the highest effectiveness and is expected to be the most expensive way to carry out the experiment, some randomization restrictions can create design structures that are even more time consuming.

In the fuel consumption experiment the oil brand is the harder-to-change factor. However, if the applied randomization restriction reduces fuel type changes the corresponding design structure will be more time consuming than the fully randomized experiment. For example, the following three fully and partially restrictive "split-plot" experiments are all more expensive than the fully randomized design. The randomization restrictions applied reduce the fuel type level changes, and as a result they cause the oil type to change levels more than expected in a fully randomized experiment.

			Oil					
			Chevron	Costco	Exxon	Mobil	QT	Shell
Replicate 1	Fuel	Valvoline	R1	R1	R1	R1	R1	R1
		Total	R1	R1	R1	R1	R1	R1
		Pennzoil	R1	R1	R1	R1	R1	R1
		Castrol	R1	R1	R1	R1	R1	R1
		Amsoil	R1	R1	R1	R1	R1	R1
		Mobil 1	R1	R1	R1	R1	R1	R1
Replicate 2	Fuel	Valvoline	R2	R2	R2	R2	R2	R2
		Total	R2	R2	R2	R2	R2	R2
		Pennzoil	R2	R2	R2	R2	R2	R2
		Castrol	R2	R2	R2	R2	R2	R2
		Amsoil	R2	R2	R2	R2	R2	R2
		Mobil 1	R2	R2	R2	R2	R2	R2

Expected Time = 1061.6 minutes

Figure 9.4. Split-plot design structure of {(R, 1, 1), (Fuel, 1, 2)}

			Oil					
			Chevron	Costco	Exxon	Mobil	QT	Shell
Replicate 1	Fuel	Valvoline	R1	R1	R1	R1	R1	R1
		Total	R1	R1	R1	R1	R1	R1
		Pennzoil	R1	R1	R1	R1	R1	R1
		Castrol	R1	R1	R1	R1	R1	R1
		Amsoil	R1	R1	R1	R1	R1	R1
		Mobil 1	R1	R1	R1	R1	R1	R1
Replicate 2	Fuel	Valvoline	R2	R2	R2	R2	R2	R2
		Total	R2	R2	R2	R2	R2	R2
		Pennzoil	R2	R2	R2	R2	R2	R2
		Castrol	R2	R2	R2	R2	R2	R2
		Amsoil	R2	R2	R2	R2	R2	R2
		Mobil 1	R2	R2	R2	R2	R2	R2

Expected Time = 1011.6 minutes

Figure 9.5. Partially restrictive split-plot design structure of {(R, 1, 1), (Fuel, 2, 2)}

			Oil					
			Chevron	Costco	Exxon	Mobil	QT	Shell
Replicate 1	Fuel	Valvoline	R1	R1	R1	R1	R1	R1
		Total	R1	R1	R1	R1	R1	R1
		Pennzoil	R1	R1	R1	R1	R1	R1
		Castrol	R1	R1	R1	R1	R1	R1
		Amsoil	R1	R1	R1	R1	R1	R1
		Mobil 1	R1	R1	R1	R1	R1	R1
Replicate 2	Fuel	Valvoline	R2	R2	R2	R2	R2	R2
		Total	R2	R2	R2	R2	R2	R2
		Pennzoil	R2	R2	R2	R2	R2	R2
		Castrol	R2	R2	R2	R2	R2	R2
		Amsoil	R2	R2	R2	R2	R2	R2
		Mobil 1	R2	R2	R2	R2	R2	R2

Expected Time = 995 minutes

Figure 9.6. Partially restrictive split-plot design structure of {(R, 1, 1), (Fuel, 3, 2)}

As seen in figures Figure 9.4, Figure 9.5, and Figure 9.6, the more restricted the easier-to-change factor, the more time it takes to finish the experiment. Out of 255 design structures 51 are more time consuming, or less efficient, than the fully randomized design structure. They reside above the horizontal dashed line in Figure 9.3.

There are 55 design structures with zero effectiveness. These are the design structures that are fully restricting both factors in the experiment, and therefore, there is no power in the design.

Removing these two sets of design structures gives the plot in Figure 9.7. The complete list of these remaining design structures, sorted by their effectiveness, is presented in Figure 9.8.

Figure 9.7. Efficiency and Effectiveness for remaining design structures

The design structures that reside on the left side of the vertical dashed line in the Figure 9.7 have a statistical power of zero for one of the main factors in the model. Therefore, they are really of no interest.

The design structures of particular interest are the ones lying on the lower edge of this two-dimensional plain (connected with dashed line). Any other design within this area is either less efficient or less effective than these designs. These six design structures are highlighted in Figure 9.8.

#	Effectiveness	Time	Design Structure
1	1.974012	965	{}
2	1.859773	876.6667	{(R, 1, 1),(A, 3, 2)}
3	1.839769	875	{(R, 1, 1),(A, 2, 2),(B, 2, 2)}
4	1.815649	670	{(R, 1, 1),(A, 1, 2),(B, 2, 2)}
5	1.80302	678.3333	{(R, 1, 1),(A, 1, 1),(B, 2, 1)}
6	1.797439	878.3333	{(R, 1, 1),(A, 2, 1),(B, 2, 1)}
7	1.791047	878.3333	{(R, 1, 1),(A, 3, 1)}
8	1.780443	850	{(A, 3, 1)}
9	1.778759	775	{(R, 1, 1),(A, 2, 1)}
10	1.778616	825	{(R, 1, 1),(A, 2, 1),(B, 3, 1)}
11	1.763198	910	{(R, 1, 1),(A, 3, 1),(B, 3, 1)}
12	1.749186	828.3333	{(R, 1, 1),(A, 2, 2),(B, 3, 2)}
13	1.749128	943.3333	{(R, 1, 1),(A, 3, 1),(B, 2, 1)}
14	1.73724	778.3333	{(R, 1, 1),(A, 2, 2)}
15	1.729884	570	{(R, 1, 1),(A, 1, 2),(B, 3, 2)}
16	1.729213	911.6666	{(R, 1, 1),(A, 3, 2),(B, 3, 2)}
17	1.719142	945	{(R, 1, 1),(A, 3, 2),(B, 2, 2)}
18	1.708608	578.3333	{(R, 1, 1),(A, 1, 1),(B, 3, 1)}
19	1.690112	735	{(A, 2, 1)}
20	1.675555	903.3334	{(R, 1, 1),(A, 3, 1),(B, 3, 2)}
21	1.675555	915	{(R, 1, 1),(B, 3, 1),(A, 3, 2)}
22	1.669703	965	{(R, 1, 1),(A, 2, 1),(B, 2, 1),(B, 1, 2)}
23	1.665742	835	{(R, 1, 1),(B, 3, 1),(A, 2, 2)}
24	1.655007	805	{(R, 1, 1),(A, 2, 1),(B, 3, 2)}
25	1.64681	905	{(R, 1, 1),(A, 3, 2),(B, 3, 3)}
26	1.64681	915	{(R, 1, 1),(B, 3, 2),(A, 3, 3)}
27	1.639056	835	{(R, 1, 1),(A, 2, 1),(B, 2, 2)}
28	1.639056	891.6667	{(R, 1, 1),(B, 2, 1),(A, 2, 2)}
29	1.63724	636.6667	{(R, 1, 1),(A, 3, 1),(A, 1, 2),(B, 2, 2)}
30	1.636862	835	{(R, 1, 1),(B, 3, 2),(A, 2, 3)}
31	1.636483	695	{(R, 1, 1),(A, 2, 1),(B, 2, 1),(A, 1, 2)}
32	1.635043	930	{(R, 1, 1),(A, 3, 1),(B, 2, 2)}
33	1.63111	470	{(R, 1, 1),(A, 1, 2)}
34	1.623546	951.6667	{(R, 1, 1),(B, 2, 1),(A, 3, 2)}
35	1.613149	808.3333	{(R, 1, 1),(A, 2, 2),(B, 3, 3)}
36	1.611917	698.3333	{(R, 1, 1),(A, 2, 2),(B, 2, 2),(A, 1, 3)}
37	1.610668	595	{(R, 1, 1),(A, 2, 1),(A, 1, 2),(B, 2, 2)}
38	1.609639	595	{(R, 1, 1),(B, 3, 1),(A, 1, 2)}
39	1.606398	638.3333	{(R, 1, 1),(A, 3, 2),(A, 1, 3),(B, 2, 3)}
40	1.606163	931.6666	{(R, 1, 1),(A, 3, 2),(B, 2, 3)}
41	1.601861	925	{(R, 1, 1),(A, 2, 1),(B, 1, 2)}
42	1.59679	838.3333	{(R, 1, 1),(A, 2, 2),(B, 2, 3)}
43	1.59679	891.6666	{(R, 1, 1),(B, 2, 2),(A, 2, 3)}
44	1.594741	711.6667	{(R, 1, 1),(B, 2, 1),(A, 1, 2)}
45	1.592839	856.6667	{(A, 3, 1),(B, 3, 1)}
46	1.592839	853.3333	{(A, 3, 1),(B, 3, 2)}
47	1.592839	858.3333	{(B, 3, 1),(A, 3, 2)}
48	1.592479	945	{(R, 1, 1),(A, 2, 1),(B, 3, 1),(B, 1, 2)}
49	1.581688	951.6666	{(R, 1, 1),(B, 2, 2),(A, 3, 3)}
50	1.580208	595	{(R, 1, 1),(B, 3, 2),(A, 1, 3)}
51	1.566787	478.33334	{(R, 1, 1),(A, 1, 1)}
52	1.565428	598.3333	{(R, 1, 1),(A, 2, 2),(A, 1, 3),(B, 2, 3)}
53	1.560426	948.3333	{(R, 1, 1),(A, 2, 2),(B, 3, 2),(B, 1, 3)}
54	1.557677	928.3333	{(R, 1, 1),(A, 2, 2),(B, 1, 3)}
55	1.555227	703.3333	{(R, 1, 1),(A, 3, 1),(B, 2, 1),(A, 1, 2)}
56	1.550557	711.6666	{(R, 1, 1),(B, 2, 2),(A, 1, 3)}
57	1.544876	585	{(R, 1, 1),(A, 2, 1),(B, 3, 1),(A, 1, 2)}
58	1.53383	556.6667	{(R, 1, 1),(A, 3, 1),(A, 1, 2),(B, 3, 2)}
59	1.52869	535	{(R, 1, 1),(A, 2, 1),(A, 1, 2),(B, 3, 2)}
60	1.522521	705	{(R, 1, 1),(A, 3, 2),(B, 2, 2),(A, 1, 3)}
61	1.518156	470	{(R, 1, 1),(A, 1, 1),(B, 2, 2)}
62	1.509611	588.3333	{(R, 1, 1),(A, 2, 2),(B, 3, 2),(A, 1, 3)}
63	1.508375	470	{(R, 1, 1),(A, 1, 1),(B, 3, 2)}
64	1.508056	745	{(A, 2, 1),(B, 3, 1)}
65	1.508056	735	{(A, 2, 1),(B, 3, 2)}
66	1.508056	748.3333	{(B, 3, 1),(A, 2, 2)}
67	1.500488	558.3333	{(R, 1, 1),(A, 3, 2),(A, 1, 3),(B, 3, 3)}
68	1.495658	590	{(R, 1, 1),(A, 3, 1),(B, 3, 1),(A, 1, 2)}
69	1.479875	538.3333	{(R, 1, 1),(A, 2, 2),(A, 1, 3),(B, 3, 3)}
70	1.475737	925	{(R, 1, 1),(A, 2, 1),(B, 2, 2),(B, 1, 3)}
71	1.470071	863.3333	{(A, 3, 1),(B, 2, 1)}
72	1.470071	856.6667	{(A, 3, 1),(B, 2, 2)}
73	1.470071	866.6667	{(B, 2, 1),(A, 3, 2)}
74	1.458442	925	{(R, 1, 1),(A, 2, 1),(B, 3, 2),(B, 1, 3)}
75	1.456794	591.6666	{(R, 1, 1),(A, 3, 2),(B, 3, 2),(A, 1, 3)}
76	1.45383292	478.3333	{(R, 1, 1),(A, 1, 2),(B, 2, 3)}
77	1.444052215	478.3333	{(R, 1, 1),(A, 1, 2),(B, 3, 3)}
78	1.442096717	711.6667	{(R, 1, 1),(B, 2, 1),(A, 2, 2),(A, 1, 3)}
79	1.431553426	928.3333	{(R, 1, 1),(A, 2, 2),(B, 2, 3),(B, 1, 4)}
80	1.422056199	475	{(R, 1, 1),(A, 2, 1),(A, 1, 2)}
81	1.418873322	711.6667	{(R, 1, 1),(B, 2, 1),(A, 3, 2),(A, 1, 3)}
82	1.417739645	595	{(R, 1, 1),(B, 3, 1),(A, 2, 2),(A, 1, 3)}
83	1.414258809	928.3333	{(R, 1, 1),(A, 2, 2),(B, 3, 3),(B, 1, 4)}
84	1.397913043	711.6666	{(R, 1, 1),(B, 2, 2),(A, 2, 3),(A, 1, 4)}
85	1.396857568	595	{(R, 1, 1),(B, 3, 1),(A, 3, 2),(A, 1, 3)}
86	1.388308641	595	{(R, 1, 1),(B, 3, 2),(A, 2, 3),(A, 1, 4)}
87	1.387960176	755	{(A, 2, 1),(B, 2, 1)}
88	1.387960176	735	{(A, 2, 1),(B, 2, 2)}
89	1.387960176	761.6667	{(B, 2, 1),(A, 2, 2)}
90	1.381045727	476.6667	{(R, 1, 1),(A, 3, 1),(A, 1, 2)}
91	1.374689648	711.6666	{(R, 1, 1),(B, 2, 2),(A, 3, 3),(A, 1, 4)}
92	1.367426564	595	{(R, 1, 1),(B, 3, 2),(A, 3, 3),(A, 1, 4)}
93	1.364920063	478.3333	{(R, 1, 1),(A, 2, 2),(A, 1, 3)}
94	1.340849296	478.3333	{(R, 1, 1),(A, 3, 2),(A, 1, 3)}
95	1.298430545	475	{(R, 1, 1),(A, 2, 1),(A, 1, 2),(B, 2, 3)}
96	1.28777118	475	{(R, 1, 1),(A, 2, 1),(A, 1, 2),(B, 3, 3)}
97	1.268335613	476.6667	{(R, 1, 1),(A, 3, 1),(A, 1, 2),(B, 2, 3)}
98	1.261398157	476.6667	{(R, 1, 1),(A, 3, 1),(A, 1, 2),(B, 3, 3)}
99	1.241294409	478.3333	{(R, 1, 1),(A, 2, 2),(A, 1, 3),(B, 2, 4)}
100	1.230635044	478.3333	{(R, 1, 1),(A, 2, 2),(A, 1, 3),(B, 3, 4)}
101	1.228139182	478.3333	{(R, 1, 1),(A, 3, 2),(A, 1, 3),(B, 2, 4)}
102	1.221201726	478.3333	{(R, 1, 1),(A, 3, 2),(A, 1, 3),(B, 3, 4)}
103	1.12364859	390	{(A, 2, 1),(A, 1, 2)}
104	1.096100724	390	{(A, 1, 1)}
105	1.069389901	390	{(A, 3, 1),(A, 1, 2)}
106	0.8901536	873.3333	{(A, 3, 1),(B, 2, 1),(B, 1, 2)}
107	0.8901536	866.6667	{(A, 3, 1),(B, 2, 2),(B, 1, 3)}
108	0.8901536	886.6667	{(B, 2, 1),(A, 3, 2),(B, 1, 2)}
109	0.8901536	891.6667	{(B, 2, 1),(B, 1, 2),(A, 3, 3)}
110	0.882580257	390	{(A, 2, 1),(A, 1, 2),(B, 3, 2)}
111	0.882580257	360	{(A, 2, 1),(A, 1, 2),(B, 3, 3)}
112	0.882580257	415	{(A, 2, 1),(B, 3, 1),(A, 1, 2)}
113	0.882580257	418.3333	{(B, 3, 1),(A, 2, 2),(A, 1, 3)}
114	0.874617719	883.3333	{(A, 3, 1),(B, 1, 1)}
115	0.874617719	866.6667	{(A, 3, 1),(B, 1, 2)}
116	0.874617719	891.6667	{(B, 1, 1),(A, 3, 2)}
117	0.870891515	870	{(A, 3, 1),(B, 3, 1),(B, 1, 2)}
118	0.870891515	866.6666	{(A, 3, 1),(B, 3, 2),(B, 1, 3)}
119	0.870891515	885	{(B, 3, 1),(A, 3, 2),(B, 1, 2)}
120	0.870891515	891.6666	{(B, 3, 1),(B, 1, 2),(A, 3, 3)}
121	0.859511256	410	{(A, 1, 1),(B, 3, 1)}
122	0.859511256	360	{(A, 1, 1),(B, 3, 2)}
123	0.859511256	418.3333	{(B, 3, 1),(A, 1, 2)}
124	0.837207592	400	{(A, 3, 1),(A, 1, 2),(B, 3, 2)}
125	0.837207592	360	{(A, 3, 1),(A, 1, 2),(B, 3, 3)}
126	0.837207592	416.6667	{(A, 3, 1),(B, 3, 1),(A, 1, 2)}
127	0.837207592	418.3333	{(B, 3, 1),(A, 3, 2),(A, 1, 3)}
128	0.81955914	755	{(A, 2, 1),(B, 2, 1),(B, 1, 2)}
129	0.81955914	735	{(A, 2, 1),(B, 2, 2),(B, 1, 3)}
130	0.81955914	791.6667	{(B, 2, 1),(A, 2, 2),(B, 1, 2)}
131	0.81955914	801.6667	{(B, 2, 1),(B, 1, 2),(A, 2, 3)}
132	0.8056805	785	{(A, 2, 1),(B, 1, 1)}
133	0.8056805	735	{(A, 2, 1),(B, 1, 2)}
134	0.8056805	801.6667	{(B, 1, 1),(A, 2, 2)}
135	0.79370405	745	{(A, 2, 1),(B, 3, 1),(B, 1, 2)}
136	0.79370405	735	{(A, 2, 1),(B, 3, 2),(B, 1, 3)}
137	0.79370405	788.3333	{(B, 3, 1),(A, 2, 2),(B, 1, 2)}
138	0.79370405	801.6666	{(B, 3, 1),(B, 1, 2),(A, 2, 3)}
139	0.712398395	390	{(A, 2, 1),(A, 1, 2),(B, 2, 2)}
140	0.712398395	330	{(A, 2, 1),(A, 1, 2),(B, 2, 3)}
141	0.712398395	440	{(A, 2, 1),(B, 2, 1),(A, 1, 2)}
142	0.712398395	446.6667	{(B, 2, 1),(A, 2, 2),(A, 1, 3)}
143	0.698777707	430	{(A, 1, 1),(B, 2, 1)}
144	0.698777707	330	{(A, 1, 1),(B, 2, 2)}
145	0.698777707	446.6667	{(B, 2, 1),(A, 1, 2)}
146	0.679649132	410	{(A, 3, 1),(A, 1, 2),(B, 2, 2)}
147	0.679649132	330	{(A, 3, 1),(A, 1, 2),(B, 2, 3)}
148	0.679649132	443.3333	{(A, 3, 1),(B, 2, 1),(A, 1, 2)}
149	0.679649132	446.6667	{(B, 2, 1),(A, 3, 2),(A, 1, 3)}

Figure 9.8. List of all competing design structures and their effectiveness and expected time

9.4 A Final List of Design Structures for the Fuel Consumption Treatment Structure

The effectiveness and efficiency criteria showed that out of 255 possible different design structures there are six that may best suit the needs of the experimenter. As shown in Figure 9.9 these design structures range from the most statistically effective to the most practically efficient.

Design Structure	Effectiveness	Efficiency
{}	1.974	965.00
{(Replicates, size: 1, order: 1),(Oil, size: 3, order: 2)}	1.859	876.67
{(Replicates, size: 1, order: 1),(Oil, size: 2, order: 2),(Fuel, size: 2, order: 2)}	1.839	875.00
{(Replicates, size: 1, order: 1),(Oil, size: 1, order: 2),(Fuel, size: 2, order: 2)}	1.815	670.00
{(Replicates, size: 1, order: 1),(Oil, size: 1, order: 2),(Fuel, size: 3, order: 2)}	1.729	570.00
{(Replicates, size: 1, order: 1),(Oil, size: 1, order: 2)}	1.631	470.00

Figure 9.9. The most effective to most efficient design structures

Figure 9.10 to Figure 9.15 show representations of these six different design structures.

		Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Total	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Pennzoil	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Castrol	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Amsoil	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2
	Mobil 1	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2	R1 R2

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, \dots, 6 \\ j = 1, \dots, 6 \\ k = 1, 2 \end{cases}$$

Effect	Statistical Power
Oil	0.88
Fuel	0.88
Oil.Fuel	0.43

Figure 9.10. Design Structure {}

		Replicate 1						Replicate 2					
		Oil						Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell	Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Total	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Pennzoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Castrol	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Amsoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Mobil 1	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2

$$Y_{ijklmo} = \mu + e1_i + e1\_r_{j(i)} + A^{(3)}{}_k + e2_{ik} + e2\_r_{o(ik)} \quad \begin{cases} i = 1,2 \\ j = 1 \\ k = 1,2 \end{cases} \quad \begin{cases} l = 1,2,3 \\ m = 1,..,6 \\ o = 1 \end{cases}$$
$$+A^{(1)}{}_{l(k)} + B_m + A^{(3)}B_{km} + A^{(1)}B_{l(k)m} + \varepsilon_{ijklmo}$$

Effect	Statistical Power
Oil	0.78
Fuel	0.87
Oil.Fuel	0.40

Figure 9.11. {(Replicates, size: 1, order: 1), (Oil, size: 3, order: 2)}

		Replicate 1						Replicate 2					
		Oil						Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell	Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Total	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Pennzoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Castrol	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Amsoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Mobil 1	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2

$$Y_{ijk} = e1_i + e1_{r_{j(i)}} + A^{(3)}{}_k + A^{(1)}{}_{l(k)} + B^{(3)}{}_m + B^{(1)}{}_{n(m)} + A^{(3)}B^{(3)}{}_{km} + e2_{ikm} + e2\_r_{o(ikm)} + A^{(1)}B^{(3)}{}_{l(k)m} + A^{(3)}B^{(1)}{}_{kn(m)} + A^{(1)}B^{(1)}{}_{l(k)n(m)} + \varepsilon_{ijklmno}$$

$$\begin{cases} i = 1,2 \\ j = 1 \\ k = 1,2 \\ l = 1,2 \\ m = 1,2 \\ n = 1,2 \\ o = 1 \end{cases}$$

Effect	Statistical Power
Oil	0.81
Fuel	0.81
Oil.Fuel	0.42

Figure 9.12. {(Replicates, size: 1, order: 1),(Oil, size: 2, order: 2),(Fuel, size: 2, order: 2)}

		Replicate 1						Replicate 2					
		Oil						Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell	Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Total	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Pennzoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Castrol	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Amsoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Mobil 1	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2

$$Y_{ijkmno} = \mu + e1_i + e1\_r_{j(i)} + A_k + B^{(2)}{}_m + e2_{ikm} \quad \begin{cases} i = 1,2 \\ j = 1 \\ k = 1,..,6 \end{cases} \quad \begin{cases} n = 1,2 \\ m = 1,2,3 \\ o = 1 \end{cases}$$
$$+e2\_r_{o(ikm)} + B^{(1)}{}_{n(m)} + AB^{(2)}{}_{km} + AB^{(1)}{}_{kn(m)} + \varepsilon_{ijkmno}$$

Effect	Statistical Power
Oil	0.81
Fuel	0.82
Oil.Fuel	0.36

Figure 9.13. {(Replicates, size: 1, order: 1), (Oil, size: 1, order: 2), (Fuel, size: 2, order: 2)}

		Replicate 1						Replicate 2					
		Oil						Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell	Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Total	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Pennzoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Castrol	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Amsoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Mobil 1	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2

$$Y_{ijkmno} = \mu + e1_i + e1\_r_{j(i)} + A_k + B^{(3)}{}_m + e2_{ikm} + e2\_r_{o(ikm)} + B^{(1)}{}_{n(m)} + AB^{(2)}{}_{km} + AB^{(1)}{}_{kn(m)} + \varepsilon_{ijkmno} \quad \begin{cases} i = 1,2 \\ j = 1 \\ k = 1,..,6 \end{cases} \quad \begin{cases} n = 1,2,3 \\ m = 1,2 \\ o = 1 \end{cases}$$

Effect	Statistical Power
Oil	0.75
Fuel	0.80
Oil.Fuel	0.34

Figure 9.14. {(Replicates, size: 1, order: 1), (Oil, size: 1, order: 2), (Fuel, size: 3, order: 2)}

		Replicate 1						Replicate 2					
		Oil						Oil					
		Chevron	Costco	Exxon	Mobil	QT	Shell	Chevron	Costco	Exxon	Mobil	QT	Shell
Fuel	Valvoline	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Total	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Pennzoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Castrol	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Amsoil	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2
	Mobil 1	R1	R1	R1	R1	R1	R1	R2	R2	R2	R2	R2	R2

$$Y_{ijklm} = \mu + e1_i + e1\_r_{j(i)} + A_k + e2_{ik} + e2\_r_{l(ik)} + B_m + AB_{km} + \varepsilon_{ijklm} \qquad \begin{cases} i = 1,2 \\ j = 1 \\ k = 1,..,6 \end{cases} \qquad \begin{cases} l = 1,2 \\ m = 1,..,6 \end{cases}$$

Effect	Statistical Power
Oil	0.56
Fuel	0.87
Oil.Fuel	0.40

Figure 9.15. {(Replicates, size: 1, order: 1), (Oil, size: 1, order: 2)}

With this range of design structures in hand, the experimenter can easily pick the one with the highest detectability power given a particular constraint on the amount of time available.

10. Conclusion

In this research two measures of efficiency and effectiveness were utilized to select a collection of the most suitable design structures for an experiment given a specific configuration of factors, their number of levels, and number of replicates. This process involved several closely connected developments:

- Enumeration of all possible design structures, involving multiple layers of both fully and partially restrictive randomization restrictions, for a given treatment structure.
- Uniquely representing each design structure by its corresponding linear statistical model through the addition of randomization restriction term.
- Transformation of treatment structures by splitting partially restricted factors into multiple sources and creating independent tests of significance.
- Development of an analysis of variance, as well as listing the corresponding expected mean squares, according to the developed statistical model and using the illustrative approach of Hasse diagrams.
- Deriving the expected time or cost of running an experiment, given its treatment and design structure, through a recursive algorithm.
- Estimating the statistical power of every non-restricted, fully-restricted, and partially-restricted factors and interactions in the design algorithmically and through transformation of treatment structures.
- Measuring the overall statistical effectiveness of each design structure in terms of the individual statistical power for its corresponding tests of significance.
- Illustrating the use of effectiveness and efficiency metrics in a case study to select a collection design structures most suited for an experiment.

Although all these processes and algorithms have been implemented and tested through simulation models, the general formulation of this system of selecting experimental designs for larger and more complicated treatment structures may not be feasible through enumeration and comparison of all possible design structures. Experiments with larger number of factors, larger number of levels, more replicates within each treatment combination, or with more complicated structures such as nested factors, require the formulation of an optimization problem through multi-objective mathematical modeling.

Furthermore, relaxing some of the assumptions made in this study opens up a lot of opportunities for future researches. For instance, the creation and analysis of partially restrictive design structures for experiments involving both fixed and random factors, design structures with asymmetric randomization restriction, and imbalanced design structures are all challenges that can be attacked in a future work.

Bibliography

Tack, L., Vandebroek, M. (2004). *Budget constrained run orders in optimum design*. Journal of Statistical Planning and Inference, 124(1), 231-249

Ganjua, J., James, M. L. (1998). *Detecting randomization restrictions caused by factors.* Journal of Statistical Planning and Inference, 81(1), 129-140

Hardin, R. H., Sloane N. J. A. (1993). *A new approach to the construction of optimal designs.* Journal of Statistical Planning and Inference, 37(3), 339-369

Simpson, J. R., Kowalski, S. M., Landman, D. (2004). *Experimentation with randomization restrictions: Targeting practical implementation*. Quality and Reliability Engineering International, 20(5), 481-495

Jones, B., Nachtsheim, C. J. (2009). *Split-Plot designs: What, why, and how.* Journal of Quality Technology, 41(4), 340-361

Jones, B. (1985). *Using bricks to build block designs.* Journal of the Royal Statistical Society, 47(2), 349-356

Fisher, R. A. (1992). The Arrangement of Field Experiments. Springer Series in Statistics Breakthroughs in Statistics, 82-91

Anscombe, F. J. (1948). The Validity of Comparative Experiments. Journal of the Royal Statistical Society. Series A (General), 111(3), 181

Cox, D. R. (2009). Randomization in the Design of Experiments. International Statistical Review, 77(3), 415-429

Paniagua-Quiñones, C., & Box, G. E. (2007). Use of Strip-Strip-Block Design for Multi-Stage Processes to Reduce Costs of Experimentation. Quality Engineering, 20(1), 46-52

Anderson, V. L. (1970). *Restriction errors for linear models (An aid to develop models for designed expeiments).* Biometrics, 26(2), 255-268

Freeman, G. H. (1976). *On the selection of designs for comparative experiments.* Biometrics, 32(1), 195-199

Anderson, V. L., McLean, R. A. (1974). *Restriction errors: Another dimension in teaching experimental statistics.* The American Statistician, 28(4), 145-152

Goos, P. (2002). *The Optimal Design of Blocked and Split-plot Experiments*. New York: Springer

Milliken, G. A., Johnson, D. E. (2004) Analysis of Messy Data: Volume 1 - Designed Experiments. 2nd ed. New York: Chapman & Hall

Lohr, S. L. (1995). *Hasse diagrams in statistical consulting and teaching.* The American Statistician, 49(4), 376-381

Jones, B., Eccleston, J. A. (1980). *Exchange and interchange procedures to search for optimal designs.* Journal of the Royal Statistical Society, 42(2), 238-243

Jones, B. (1979). *Algorithms to search for optimal row-and-column designs.* Journal of the Royal Statistical Society, 41(2), 210-216

Jones, B., Goos, P. (2009). *D-optimal designs of split-split-plot experiments.* Biometrics, 96(1), 67-82

Anscombe, F. J. (1948). *The Validity of Comparative Experiments*. Journal of the Royal Statistical Society. Series A (General), 111(3), 181-211

Bowman, K. O., Kastenbaum, M. A. (1974). *Potential pitfalls of portable power.* Technometrics, 16(3), 349-352

Capehart, S. R., Keha, A., Kulahci, M., Montgomery, D. C. (2001). *Designing fractional factorial split-plot experiments using integer programming.* International Journal of Experimental Design and Process Optimization, 2(1), 34-57

Lorenzen, T. J. (1987). *A comparison of approximate F' test under pooling rules.* Research Publication GMR-5928: Mathematical Department, General Motors Research Laboratories, Warrens, MI 48090-9055

Lorenzen, T. J., Truss, L. T., Spangler, W. S., Corpus, W. T., Parker, A. B. (1992). *DEXPERT: An expert system for design of experiment.* Statistics and Computing, 2(2), 47-54

Donckels, B. M. R., De Pauw, J. W. D., De Baets, B., Maertens, J., Vanrolleghem, P. A. (2008) *An anticipatory approach to optimal experimental design for model discrimination.* Chemometrics and Intelligent Laboratory Systems, 95(1), 53-63

Box, G. E., Hunter, W. G., & Hunter, J. S. (1978). Statistics for experimenters: An introduction to design, data analysis, and model building. New York: Wiley

Wheeler, R. E. (1974). *Portable Power*. Technometrics, 16(2), 193-201

Wheeler, R. E. (1975). *The validity of portable power.* Technometrics,17(2), 177-179

Goos, P., Jones, B. (2011). *Optimal Design of Experiments: A Case Study Approach.* Hoboken, NJ: Wiley

Scott, M. E., Delaney, H. D. (2003). *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. Belmont, CA: Wadsworth Pub

Lorenzen, T. J., Anderson, V. L. (1993). *Design of Experiments: A No-name Approach*. New York: M. Dekker

Montgomery, D. C. (2008). *Design and Analysis of Experiments*. 6th ed. New York: Wiley

Kuehl, R. O. (2000). *Design of Experiments: Statistical Principles of Research Design and Analysis*. Belmont, CA: Brooks Cole

Odeh, R. E., Fox M. (1975). *Sample Size Choice: Charts for Experiments with Linear Models*. New York: Marcel Dekker

Aberson, C. L. (2010). *Applied Power Analysis for the Behavioral Sciences*. New York: Routledge

Jackson, S. A., Brashers D. E. (1994). *Random Factors in ANOVA*. Thousand Oaks, CA: Sage Publications

Atkinson, A. C., Donev. A. N., Tobias. R. D. (2007) *Optimum Experimental Designs, with SAS*. Oxford: Oxford UP

Federer, W. T., & King, F. (2007). Variations on split plot and split block experiment designs. Hoboken, NJ: Wiley-Interscience

Barrentine, L. B. (1999). An introduction to design of experiments: A simplified approach. Milwaukee, WI: ASQ Quality Press

Ryan, T. P. (2013). Sample size determination and power. Hoboken, NJ: Wiley

Liu, X., De Raudenbush, S., (2004) *A Note on the Noncentrality Parameter and Effect Size Estimates for the F Test in ANOVA.* Journal of Educational and Behavioral Statistics, 29(2), 251-255