AD-A035 109 HUGHES AIRCRAFT CO CULVER CITY CALIF DISPLAY SYSTEMS--ETC F/G 5/10
METHODS OF HANDLING SEQUENCE EFFECTS IN HUMAN FACTORS ENGINEERI--ETC(U)
DEC 74 C W SIMON F44620-72-C-0086
UNCLASSIFIED HAC-P74-541A NL

1 OF 3
AD
A035109

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: HAC-P74-541 A
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): Methods of Handling Sequence Effects in Human Factors Engineering Experiments,
5. TYPE OF REPORT & PERIOD COVERED: Technical report, Jul 73 — Jul 74
6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Charles W. Simon
8. CONTRACT OR GRANT NUMBER(s): F44620-72-C-0086

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Hughes Aircraft Company, Culver City, California 90230 (209 p.)
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Project No. 9778

11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, 1400 Wilson Blvd., Arlington, VA 22209
12. REPORT DATE: Dec 74
13. NUMBER OF PAGES: 197

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): -
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: -

16. DISTRIBUTION STATEMENT (of this Report)

Revised edition: Distribution list included at end of report.

61102F

DDC RECEIVED FEB 1 1977

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

-

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

18. SUPPLEMENTARY NOTES

Original report, P74-541, July 1974
Revised edition, P74-541A, August 1976

Superseded by A006240 per ltr 2nc

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sequence effects	Serially balanced sequences
Change-over designs	Experimental design
Trends	Human factors research
Presentation Order	Within-subject designs

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

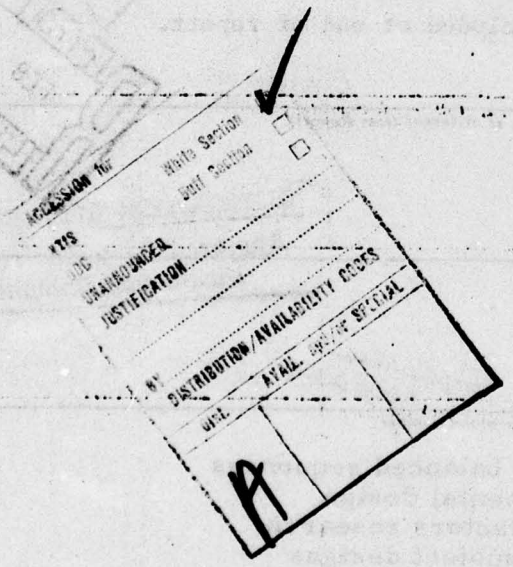
A common procedure in human factors engineering experiments is to test the same subject sequentially on a series of different experimental conditions. These "within subjects" designs produce sequence effects that may be unwanted (in the comparison of equipment designs) or of considerable interest (in the development of training devices). In this report, the more common sequence effects are identified and a variety of (identifies) (describes) (over-)

DD FORM 1 JAN 73 1473A EDITION OF 1 NOV 65 IS OBSOLETE 405 763 UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (continued) (fr p 1473A.)

procedures, experimental designs, and statistical techniques to minimize, adjust for, or isolate these effects, ~~are described.~~ Change-over designs are described that permit reasonably unbiased estimates of treatment effects when it is suspected that residual effects from prior treatments may have carried over to affect the direct performance at the time of measurement. Some designs are suitable when a number of subjects are available for short periods of time; others are suitable when a single subject can be tested on an extended sequence of trials. Designs are described that enable reasonably unbiased estimates of treatment effects to be made even in the presence of underlying shifts in performance when such trend effects cannot be associated with changes in the experimental conditions. Other designs recommend sequences that may be robust against trend effects while limiting the number of times the level of one or more factor must be changed as experimental conditions are changed, thereby increasing the efficiency with the experiment can be run. Some methods of comparing the effectiveness of different treatment presentation orders and for optimizing training schedules are also discussed. (82 references).



METHODS OF HANDLING SEQUENCE EFFECTS IN HUMAN FACTORS ENGINEERING EXPERIMENTS

Charles W. Simon
Hughes Aircraft Company

Technical Report No. P 74-541A

Prepared under contract with the Air Force Office of Scientific Research, (AFSC), United States Air Force.

December 1974

AEROSPACE GROUP

Hughes Aircraft Company, Culver City, California

ABSTRACT

To conserve personnel and to make more precise comparisons among experimental treatments in human factors engineering experiments, subjects are often tested sequentially on a series of experimental conditions. As a result, sequence effects are created which, if improperly or inadequately handled, can distort the results obtained from the study in a way that will cause erroneous conclusions to be drawn from the data.

Sequence effects are conventionally handled in behavioral research by introducing "randomization" or "counterbalancing" into the experimental design. However, unless the design is carefully selected and more important, unless the effects artificially introduced by design sequence itself are removed, procedures may actually increase the bias in the data and the chances of drawing incorrect conclusions. Special procedures and experimental designs are described in this report for handling the different types of sequence effects commonly found in human factors engineering research using "within subjects" designs.

There are designs that permit the measure and isolation of "carry-over" effects. These are the residual effects from a previous treatment that carry over to affect performance on the treatment that follows, or to affect the performance of the treatment two trials away. To measure this effect, some designs require a number of subjects (often a large number) for a few trials, while others require only a single subject tested over a larger number of trials.

There are designs for ordering a series of experimental treatments in a way that linear and quadratic trends running through the experiment can be isolated from experimental effects. In training studies, these trend effects may be of as much interest as the experimental conditions. In many equipment design studies, they are often artificial and unwanted.

There are designs that show how to sequence a series of treatments in a way that the number of times a factor-level must be changed can be minimized. This can do much to reduce the work required to run equipment design experiments. Certain designs with the above properties can also be planned so that the sequences will be robust to trend effects.

A limited discussion of the methods and problems of measuring the effects of presentation order is provided. This is a particularly important problem in devising optimized training schedules and one that has not been adequately developed.

The report attempts to introduce the reader to a wide range of practical techniques for handling sequence effects often found in applied behavioral research. It identifies them and shows how designs for different purposes can be constructed when it's easy to do, or supplies finished designs when construction is difficult. Emphasis is on design and not on analysis.

Future needs are discussed.

FOREWORD

The techniques for handling sequence effects presented in this report are intended to be used with the economical multifactor experimental designs described in an earlier report (Simon, 1973). In order to study a large experimental space with a minimum number of observations, it is important that each measurement be as error free as possible. When a subject is tested on set of different experimental conditions over a series of trials, measurements of the effects of experimental variables can become confounded with spurious effects associated with the sequencing. The presence of such phenomena as learning and transfer can introduce biases into the data. The recognition that such dangers are inherent in these "within subjects" designs has not diminished their popularity among those conducting human factors engineering research. If such effects are not to be avoided, then they must be dealt with.

To find methods of handling these sequence effects, I went through a great many statistical journals published during the past thirty years in search of techniques that have been successfully employed for that purpose in agricultural, chemistry, engineering, and biological research. Those methods that I believed would be particularly applicable in the design of human factors* experiments and preferably those with which most psychologists were likely to be unfamiliar were included in this report. In addition, certain techniques that commonly appeared in popular statistics textbooks were included when they appeared useful yet seldom or never used in human factors experiments. From the multitude of specific experimental designs associated with these techniques, I selected for inclusion those that require relatively few subjects and relatively few total observations while at the same time are capable of estimating the various effects with the highest efficiency possible. I also sought the most

---

* Emphasis on human factors should not be construed to mean that these techniques would not be useful in other fields of psychology. I am sure in many cases they would be.

up-to-date versions of the different classes of designs, although it became apparent that certain designs were still relatively primative when applied to problems involving human performance. This is probably because the physical science research for which they were developed seldom faced the complexities -- which relate to the assumptions required to use a design -- that often occur in psychological research. Some of these less well-developed designs however were also included in the report for the sake of completeness and to encourage their further development should the need be established.

My emphasis in preparing this report, therefore, was to bring together under a single cover as many different but potentially useful techniques for handling sequence effects as I could find. If nothing more, a consolidated and organized report would hopefully alert the reader to the reality of these types of effects and make him aware that some relatively simple ways of reducing and measuring their effects are available. This emphasis on breadth of coverage however, given the normal restrictions on time and funding in preparing this report, reduced the degree to which certain techniques were covered in depth.

In place of a complete and detailed discussion on each and every technique presented here, particularly since all were neither equal in importance nor development, the basic information supplied in this report was supplemented by referencing the reader to pertinent sections of the original papers and occasionally secondary references. When a technique for constructing a design was easy to explain, this was done in the report, often in a form that was clearer than had been the case in the original presentation. However if understanding a construction technique required a knowledge of advanced mathematics, only the completed designs were given in the report. The reader who wishes to learn how to construct these designs is referred to the original paper. Similarly, discussions of the mathematical bases of these designs are found only in the original papers. A major omission however is the details on how to analyze the data once it has been collected. Reference is made instead to those papers that described both the general method of analysis and provided an arithmetric example.

While the interweaving of this report with other sources results in a less convenient document than might be desired, the decision to do so was based on certain practical considerations. The report was already larger than had been originally planned, and including material that essentially would have been copied verbatim could not be justified. The report can stand alone as an introduction to the design and interpretation of this general class of experiment, and the reader who seeks more information has the option of pursuing the recommended references. Because of the amount of culling and filtering and simplification that has gone into the report, using it as a guide can save the serious experimenter, literally hundreds of hours searching for, selecting from,and understanding original papers on handling sequence effects. There is also the hope that by inter-relating this secondary source with the primary ones, the tendency among some psychologists to apply experimental designs in "cookbook" fashion will be reduced.

These techniques have proven useful when employed in the research from other disciplines and they can be useful for research involving human performance. Before the full impact can be assessed however we must gain more experience using them.

I welcome comments and criticisms from any reader. I would appreciate being notified whenever errors are discovered. I am most willing to discuss how these techniques might be applied to specific problems.

Charles W. Simon
1974

ACKNOWLEDGEMENTS

This paper was prepared as part of a research project conducted at the Hughes Aircraft Company, Culver City, California, under Contract No. F44620-72-C-0086 with the Air Force Office of Scientific Research, Air Force Systems Command, United States Air Force. During the period in which this report was being written, Dr. Glen Finch and Dr. Charles Hutchinson both served as technical monitor for AFOSR. Currently, Dr. Ralph R. Canter is the technical monitor for AFOSR. The encouragement given this project by all three of these men is gratefully acknowledged.

The help of several persons deserves special recognition. Rose E. Konrath and Masse Bloomfield of the Hughes Aircraft Company, Culver City Library, materially hastened the pace of the project through their assistance in obtaining copies of critical papers from various libraries throughout the country. Edward J. Dragavon and Prentiss N. Robinson helped me to understand some of the more complex mathematical discussions. Glen L. Gray directed the assembly and the production of the finished report. John G. Bean provided a general support that materially assisted in the fulfillment of this contract.

TABLE OF CONTENTS

Page

FIGURES

TABLES

TERMINOLOGY

Occasionally throughout the text, different phrases have been used to imply the same thing. This is not necessarily a tribute to the richness of the English language; instead it reflects a certain laxity on the part of the author, and a degree of propinquity with the original papers from which the information was taken.

For example, to indicate that subjects and experimental conditions have gotten together, different phrases were used at different times. It was said that "subjects were given a different treatment", "run under different experimental conditions," "tested on different conditions," or "the treatments were applied...". Each reflects the diverse backgrounds of those who devised the experimental design or conducted the experiments. When one "gives" or "applies" treatments, there is generally a medical, biological, or chemical background involved. When one speaks of "running" subjects through an experiment, once upon a time that individual ran rats through a maze. Similarly, to talk of "testing" subjects suggests the residual of an earlier evaluation and measurement background. However, in spite of direct interest in applying these designs to human factor engineering problems, at no time was the phrase used: "the subjects operate under the different conditions" (where operate means to work, not to dissect). However, in spite of the fact that many designs originated for agricultural research, about the only term that survived was "periods". While certain statistics books speak of "plots" (and "split-plots" were mentioned in passing here too), the term "blocks" was always used here. In writing this report, I used the phrase that came naturally at the moment. In certain cases, the same symbol --in different parts of the report -- was used to mean different things, in order to agree with the terminology used in the original paper. It is unlikely that these variations will be too disturbing to an involved reader. However, some of the more frequently occurring ones are listed below:

Equivalents

- treatments, experimental conditions, conditions, factor levels, $\underline{t}$
- periods, trials, runs, $\underline{p}$
- residual, carry-over, $\underline{r}$
- number of subjects, total units, total sequences, $\underline{n}$
- trend, trend effect
- factor, variable
- total number of observations, $\underline{N}$
- tested on, tested under, run on, applied, given (different experimental experimental conditions)
- error, remainder, what's-left-over, residual

Inequivalents

- "order" may refer to the specific positions of a set of treatments in a sequence, or to the highest power in a polynomial equation, or to the number of treatments in a serially balanced sequence.
- "r" for residual is not the same as the $\underline{r}$ in $(t = 2r)$, $(2r + 2)$, $(r + 1)$, and so forth where the latter r's refer to any number and is used in these values because it was so used in the original papers.
- "m" can be the number of treatment replications in a sequence, or the number of sequences used in balancing for trend effects.
- "k" can be the number of units in a block of a change-over design or the number of complete sequences in SBS and trend designs.
- "balance" will differ as to exactly what factors of an experiment has been balanced, depending on the particular design.
- "residual" will generally refer to the effect that has carried over from a previous treatment when a series of treatments are tested on the same subject; occasionally it will be the name of the error term in an analysis of variance table, being what was left over after all other sources of variance had been isolated.

CHAPTER I.
INTRODUCTION

To conserve subjects and to make more precise comparisons among treatments, a common procedure in human factor engineering experiments is to test the same subject sequentially on a series of different experimental conditions. In practice, however, intended advantages of this type of design may be outweighed by effects artificially created by the sequential presentation itself. Responses to a condition on one trial may be influenced by the conditions that have preceded it.

Sequence effects, whether they are trends that run through the data or short term transfer effects from trial to trial, if undetected or allowed to remain confounded with the remainder of the experimental data can seriously distort the effects of primary interest. While trend and transfer effects may not be as unwanted or irrelevant in training research as they usually are in equipment design studies, they still can distort the information obtained unless precautions are taken ahead of time to handle them. Although several rituals of experimental design are commonly employed by behavioral scientists who recognize that sequence effects can occur, they have not always been effective nor appropriate.

How to properly handle sequence effects is one of the least understood facets of human factors experimentation. While the danger of failing to optimize experimental designs in other ways is usually obvious, the consequences of improperly ordering experimental conditions presented sequentially are generally less visible to the uninformed. As a result, this source of error is all the more insidious, having the potential of being confounded with every other source of variance in the experiment and thereby distorting the results in undetectable ways.

In this report, various procedures, experimental designs, and statistical techniques to minimize, adjust for, or isolate sequence effects will be described, along with a discussion of the circumstances in which they are most useful.

TRADITIONAL METHODS OF HANDLING SEQUENCE EFFECTS

A survey of 239 experiments appearing in 118 papers published in the journal, Human Factors, between 1958 and 1972 revealed that more than 77% used a data collection plan in which a series of different experimental conditions were administered sequentially to the same subject (Simon, 1974). Listed below are the various methods reported in these studies for handling sequence effects that might occur in the data:

1. The method of handling sequencing was never mentioned in the description of the experimental procedure.

2. If the sequence was determined by the nature of the task rather than by the experimenter (e.q., in a reconnaissance task, the targets were embedded in the radar imagery and for a moving-scene presentation, the order was fixed by their positions on the film), a reversed order was sometime but not always possible.

3. Some investigators' described how they handled sequence effects by simply stating they "juggled", "varied", or "semirandomized" their presentation.

4. When the set of experimental conditions was repeated in the same study, some investigators "randomized" the sequence of treatment presentations on the first trial and:

   a. Repeated that order on subsequent trials; or

   b. Reversed that order on a subsequent trial; or

   c. Used a different randomization on subsequent trials.

5. When more than one subject was tested under the same set of experimental conditions, some investigators "randomized" the sequence in which the treatments were presented, and

   a. Maintained the same "randomized" order for all subjects, or

   b. Used a different "randomized" order for each subject, or

   c. Used a different "randomized" order for groups of subjects.

6. Some investigators acknowledged that they used "restricted randomization" to determine their order of presentation. This meant that after

some presentation order had been determined essentially by chance, treatments that occurred in a sequence of positions which the investigator believed might create unwanted effects were rearranged.

7. Some investigators preferred more systematic methods of determining order of presentation. The concept of "counterbalancing" is frequently referred to by this group. Counterbalancing in these studies took several forms, such as:

   a. Seeing that each condition occurred an equal number of times in each position of a sequence of trials.

   b. Seeing that each condition occurred an equal number of times in every trial position and for every subject and (with the exception of the treatments at each end of the series) appeared once before and after every other condition.

   Familiar experimental designs used to handle these systematic methods of counterbalancing were in the form of Latin square and Graeco-Latin square designs. There was insufficient information in most cases to determine whether the same or different Latin squares were employed within the same experiment.

8. Systematic and random ordering were both used in some studies.

Experimenters provided essentially no justification for selecting the particular method they used to handle sequence effects. Furthermore, of those experimenters concerned enough to do something about the sequence effects (as shown by the way he planned his experiment), only a very small proportion actually analyzed their data to remove these effects from treatment and error estimates. Those that did looked at the following effects:

1. Differences among mean performances for trials.
2. Differences among the mean performance for different presentation orders totally confounded with differences among subjects.
3. Differences among mean performance for different presentation orders.
4. Trend effects.
5. Interactions between treatments and trials and subjects and trials.

In only a very few cases were these analyses of sequence effects mentioned or interpreted during a discussion of the experimental results.

TYPES OF SEQUENCE EFFECTS

The design in Figure 1 is typical of those commonly used in human factors engineering experiments. Basically it is built around a Latin square arrangement in which a series of treatments are presented in different orders to a set of homogeneous subjects. Treatments are presented in a different order to each pair of subjects. The entire design is repeated on a second day. An experiment that follows this data collection plan will contain a variety of sequence effects, which, if ignored, can bias treatment and/or error estimates. Sequence effects that can be examined in this design are:

Period (trial) effects

Block (day) effects

Trend effects

Residual (carryover) effects

Presentation order effects

These are not necessarily mutually exclusive effects. In some cases they represent different ways of analyzing the same data -- an examination of mean differences or of functions formed by a series of means. Each represents the answer to a different kind of question an experimenter can ask of his data. The detection of an effect may be a matter of great practical interest or may be merely an indication that the artifact must be isolated to purify the data. Whatever the case, sequence effects must be identified, measured, and removed from the remaining data.

Trial (Period) Effects *

When subjects are tested repeatedly in the same experiment, performance may vary from trial to trial as a result of effects not due to the particular experimental

* In the body of this report, trials will be referred to more often as periods, in line with the terminology used in most of the papers being reviewed.

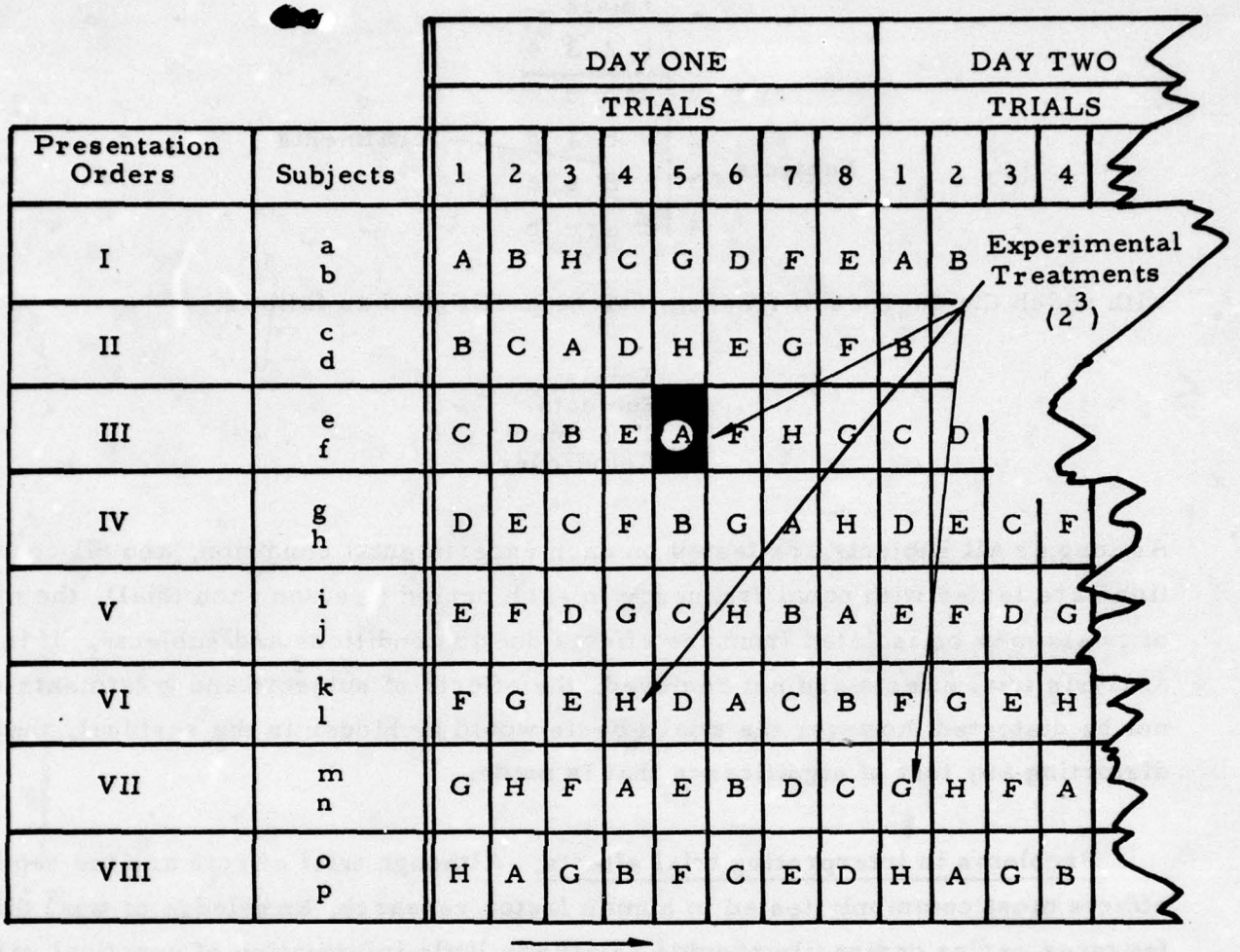


Figure 1. Typical within-subject experimental design

conditions being tested. In some experiments, either the same experimental conditions or different experimental conditions may be tested sequentially from trial to trial. The interpretation of trial effects in these two cases would differ. This report will be concerned with trial effects that result when different experimental conditions are tested sequentially.

An experimenter may have an interest in determining the presence of trial effects. But even if he has not, he should plan his experiment so that trial effects can be partitioned from other effects of interest. For example, a typical design -- a simplified versions of the one in Figure 1 and one commonly employed in human factors research -- is:

		Trials			
		1	2	3	4
Subjects	1	a	b	d	c
	2	b	c	a	d
	3	c	d	b	a
	4	d	a	c	b

Treatments

with which the degrees of freedom can be partitioned as follows:

Trials	3
Subjects	3
Treatments	3
Remainder	6.

As long as all subjects are tested on each experimental condition, and all conditions are tested with equal frequency in each period (i.e. on each trial), the effects of trials may be isolated from the effects due to conditions and subjects. If in the analysis trial effects are not removed, the effects of subjects and treatments would not be distorted; however the trial effects would be hidden in the residual, thereby distorting any test of significance that is made.

Problems in interpreting trial effects. Although trial effects are the sequence effects most commonly tested in human factor research, knowledge of trial differences *per se* ordinarily provide relatively little information of practical value. This is because of the problems that exist in the interpretation of trial mean differences.

For one thing, what may appear to be a trial effect may not actually be one. In designs built around Latin squares, trial effects are confounded with the subject-by-treatment interaction*. Presumably these designs are to be used only when an investigator believes that such an interaction does not exist; however, in practice the matter is often ignored even though with human experimentation the possibility of large subject-by-treatment interactions is high. As a

* Campbell and Stanley (1963) discuss this in their section on "Counter-balanced designs", as do most authors of statistical texts that discuss the Latin square design. Campbell and Stanley's book however is a particularly excellent one in that it relates a great many experimental (and quasi-experimental) designs to a series of factors that can jeopardize the internal and external validity of the research data.

result, even if there is care in minimizing these interactions, interpreting the meaning of statistically significant trial differences must be done cautiously. However, as the number of experimental conditions increase (i.e. the larger the Latin square), the more completely the interaction effect will tend to cancel out in the comparison of the main effects -- treatments, subjects, and trials (Linquist, 1953, pp. 261-264).

Then too, results of tests of the statistical significance of mean differences only show with some probability that a reliable difference exists among trial means. This doesn't say that an observed difference will be repeated (were the study repeated), only that it has a certain probability of being so. But even more important, tests of statistical significance result in ambiguous information about trial differences that must be interpreted subjectively by the experimenter. For example, the mean performance of five trials might be:

Trial	1	2	3	4	5	
Mean	3	5	1	2	4	(Pattern A)

or it might be:

Trial	1	2	3	4	5	
Mean	1	2	3	4	5	(Pattern B)

In both cases, the variability among trials is the same and the results of a test of statistical significance would be the same. However, the interpretations of the two sets of data would be considerably different. In the first example, the presence of a statistically significant irregular pattern of means across trials would be difficult to explain in a human factors experiment assuming that differences of this magnitude are of practical importance. If the effect is significant, one might suspect the presence of some unknown, unwanted, and undesirable concomittant event that is disrupting the measurements of interest. An examination of the raw data may or may not reveal the cause. Since the Latin square design is to be used only when the investigator feels assured that there are no interactions affects among treatments, subjects, and trials, an irregular trial pattern may suggest that this was not the case.

In the second example, the presence of a statistically significant regular pattern probably would be interpreted as evidence that learning had taken place. To discover the existence of such regular patterns is generally the reason why the significance of trial effects is tested. In vigilance studies, the interest would lie mainly in discovering whether the pattern is downward, indicating performance decrement. The point that is being emphasized here is that the interest is in the pattern rather than in whether trial means differ.

Block Effects

When an experimental design includes more trials than can be run during a single session, it is often necessary to divide the experimental conditions into blocks. In Figure 1, separating the total experiment into two days is an example of blocking. In this example, all of the treatments to be studied are found in each block. It is also possible to have "incomplete block designs" (Simon, 1970 a & b; Cochran and Cox, 1957) with only some of the total number of treatments in each block. The advantage of properly blocked designs is that the comparisons of performances among treatments within blocks will not be affected by average performance differences between blocks. If the average performance increased, for any reason, between days in the example, this change would not affect the comparison of treatments within the blocks. Furthermore, if our example were extended to include a series of days (blocks), the average performances of blocks could be examined in the same way average performances of trials are examined – to look for differences, or to see if there is a general performance trend over days (blocks).

Trend Effects

An experimenter can be interested in the presence and nature of trends in his data for at least two reasons. One, in learning and vigilance experiments, he may wish to know whether progressive changes in performance are taking place. Two, in equipment design studies in which a subject is tested on different experimental conditions sequentially, he may wish to assure himself that observed differences are due to the treatments and not to some artificially induced trend running through the experiment. In the second case, there are methods (to be described later in this report) of ordering the treatment presentation sequence so that the existence

of simple trends will not affect the comparisons among treatments. In both cases, however, trend effects (over trials) should be partitioned from the effects of other sources of variance including the error variance.

Trend effects can be determined from the analysis of the trial means. Thus, given Pattern B above, it is possible to partition the four degrees of freedom associated with five trial means as follows:

Trials	4
Linear trend	1
Quadratic	1
Cubic	1
Quartic	1

N-1 trend components with one degree of freedom each can be determined for N trials. Ordinarily, only linear, quadratic, and occasionally cubic components are partitioned; all higher-order effects are combined since they are generally negligible anyway. With this form of analysis, the existence of each trend component can be tested statistically. Although interpreting trend components of the trial means still suffers from the same problems of data purity as the analysis of trials in the Latin square design, at least, it is the preferred one for the kind of interpretations that are usually being made. The presence of a pattern rather than a point (mean) difference itself indicates a regularity to the data that inspires confidence.

The trend effects traditionally analyzed in human factors experiments using Latin square designs have generally been based on the trial means across all subjects and conditions. This is an overall or group trend. In Latin square designs where subjects are tested sequentially on different conditions, if a trend is observed -- for example, an upward trend -- the interpretation that learning is taking place is different from a learning trend found when a subject is tested repeatedly on the same condition. In the former case, it is a "learning-to-learn" phenomenon, reflecting the learning (and transfer) that might occur from learning general procedures or possibly from developing manipulation skills. The presence of a trend here might also reflect a change taking place in the environment.

When experimental designs can be divided into a series of orthogonal blocks, mean performances of blocks can also be examined for trend effects in the same way that trial means can be examined.

When only overall trial means are examined, there is an implicit assumption that individual performances as a function of trials (individual trend effects) will correspond to group trend effects. When Latin square designs are used, this is the same as assuming there are no subject-by-trial interaction effects. Unless special care is taken, however, individual curves will not necessarily correspond to the group curve. Therefore, particularly in training experiments, an examination of individual trends and the differences among them become a proper source of variance worthy of special investigation. With proper experimental designs, both of these sequence effects can be examined. Designs that measure trend effects are discussed in Chapters V and VI.

Carry-Over (Residual) Effects

When subjects are tested on different experimental conditions (i.e. treatments) sequentially over time, the performance on any particular condition may be a function of not only the direct effect of that treatment alone but also of residual effects carried over from the immediately preceding treatments. In training studies, such carry-over effects may be considered a relevant source of variance, sometimes referred to as "transfer effects". In most human factors experiments on equipment design, however, these effects are generally irrelevant and unwanted. Poulton (1969, p. 686) goes so far as to say: "If it is worth running an experimental trial to compare alternative designs of equipment, it is worth using a separate group design. A balanced treatment design can bias experimental results, and give an incorrect order of difficulty... There is often no way of detecting this except by repeating the trial using a separate group design." However, this recommendation does not consider the alternative of designing the experiment so that these carry-over effects can be measured independently of the direct treatment effects and adjusting for them in the analysis. In general, psychological experiments that have employed within-subjects designs have ignored this class of sequential effects.

Ordinarily residual effects are short-lived. For all practical purposes performance on a condition is influenced mainly by the condition that immediately preceded it (First residual). Less frequently, circumstances may exist in which a residual effect may have carried over from a condition two conditions earlier (Second residual).

Residual effects may be either simple or complex. An effect is simple when the magnitude of the residual is independent of the treatments that follow it. The effect is complex when the treatment is not independent of the treatment that follows it, that is, the residual of Treatment A is different depending on whether Treatment A is followed by Treatment B or C. This lack of independence may appear in several forms. For example, the residual effect of the previous treatment may be proportional to the magnitude of the direct effect of the treatment being measured, or it may include the effect of an interaction between the preceeding and immediate treatment.

Designs that measure residual (or carry-over) effects are discussed in Chapters III (when several subjects are required) and IV (when only one subject is used but for a larger number of observations).

Order of Presentation Effects

Up to this point, the discussion has centered around sequence effects that occur from condition to condition in a series of treatments. There can also be interest in the sequence (or order of presentation) itself. For example, in training studies, it may be important to discover in which order a set of lessons should be presented so as to maximize the final performance level.

Order effects exist in many human factors experiments (particularly those using the Latin square design) because the experimenter has put them there. Therefore, even when there is no real interest in order effects per se, it is still desirable to isolate this artificially induced source of variance from experimental error when the data is analyzed. However, when there is an interest in the effects of different orders on performance, the conventional method of analyzing their effects (i.e. removing the order effect in the analysis of variance) leaves much to be desired. This is so because the test of significance of the order effect only indicates whether or not a reliable difference probably exists between the average performance on all conditions that make up the complete sequence for each order. In most situations, there will be little concern as to whether or not the average performance of one order or the other is higher. Of more practical interest,

for example to establish a preferred training schedule, would be in finding the order that results in the fastest rate of learning and terminates at the highest performance level. Thus, an analysis of the performance curves (trends) of different presentation orders will ordinarily yield more useful information than a comparison of mean order effects.

Designs for studying the effects of presentation orders are discussed in Chapter VII.

RANDOM VERSUS SYSTEMATIC TREATMENT PRESENTATION ORDERS

In this report, a number of techniques will be proposed for systematically presenting a series of treatments to subjects in a way that will minimize, measure, or isolate sequence effects that could distort the effects of primary interest. However, Fisher (1949, p. 63) has had this to say about systematic presentation orders: "The results of using arrangements which differ from random arrangements in either direction are thus in one way or the other undesirable." He proposed instead that when faced with a decision of how to design any aspect of an experiment wherein uncontrolled systematic biases may occur -- whether they be known or unknown -- the order should be determine by chance, e.g. by the roll of a die, the flip of a coin, or the drawing of numbers from a hat. Traditionally both of these basic techniques -- systematic versus randomized presentation orders -- have been employed in human factors research. What then are the advantages and disadvantages of each method and to what degree can the differences of opinion regarding the two approaches be recomciled?

Randomized Arrangements

Fisher introduced the concept of randomization into the design of experiments as a means of offsetting irrelevant sources of systematic variance that could not be controlled by the experimenter and which would possibly distort the results of primary interest. He reasoned that with randomization the systematic effects would be broken up and scattered throughout the experiment among the treatments of interest. If the experiment was therefore repeated enough times, the randomized elements would be averaged out while the treatments still systematically varied could be estimated with a great deal of accuracy.

Fisher believed that a still more important reason for randomization was to preserve a valid estimate of the error term. This was critical because without it, no valid test of statistical significance could be made. Fisher (1949, p. 62) argued as follows:

> "The effect of a systematic arrangement on the test of significance may be seen by imagining it carried out on an area under uniform treatment, so that the actual yields are not at all affected by the reallocation of plots. In the analysis of variance, therefore, the total sum of squares is unchanged, as is the portion ascribable to blocks. If, therefore, the agronomist's ingenuity has been successful in diminishing the differences in fertility between treatments, the diminution of the sum of squares in that line of the table will have been exactly counterbalanced by an increase in the sum of squares upon which the estimate of error is based. The effect of the rearrangement will have been to diminish the real errors of the experiment, but at the expense of increasing the estimate of error; so that, although the comparisons have really been improved in precision they will appear to have been less accurate than before, and less reliance will be placed on the result. In the opposite case, likewise, if by bad luck or bad judgment the systematic arrangement adopted has increased rather than lessened the real errors of the experiment, then the estimate of error will be even diminished, and will be, for both reasons, an under-estimate of the errors actually incurred."

Thus Fisher concluded that either way systematic arrangements affected the error estimate, it was "rendered unreliable for the purpose for which it was made."

Psychologists, even before Fisher's time, were sensitive to unwanted systematic biases that could affect experimental results. Psychophysicists, concerned with psychological thresholds of physical stimuli, were critical of Wundt's method of minimal change, for he used serial presentation orders that increased or decreased systematically. They were aware that this method "with complete knowledge" introduced constant errors of habituation and expectation. (Guilford, 1936, p. 116). Randomization therefore has proved useful in those studies, as well as in more recent research on clinical drugs, as a means of concealing from the subject what treatment is to be applied on a particular trial (Cox, 1958, p. 79).

Most statistical tests require that each observation in an experiment be independent of one another, yet in fact, testing the same subject on a series of treatments sequentially can often lead to correlated responses. One of the earliest methods of overcoming these trend effects was to randomize the treatment order a la Fisher and to repeat the study enough times so that these "nuisance" factors (Hays, 1963, p. 449) would be neutralized. Cox (1958, Chapter 5) provides a short but interesting discussion on the use of randomization in experimental designs.

Disadvantages and limitations. When enough replications are taken, randomization will ordinarily yield reasonably accurate estimates of the treatment effects and can supply a valid estimate of error with which a test of statistical significance can be made. But there are problems.

First, the number of replications required to neutralize an uncontrolled systematic effect when the treatments are presented in a random order can be quite large. One problem with this approach therefore is that it can be expensive in time and money.

Second, in some cases, randomization and replication have been used more for the purpose of escaping responsibilities than because they represented the best method of obtaining the desired information. Simon (1973, pp 19-32) discusses circumstances in which replication is used unnecessarily. This approach has been most commonly employed in field experiments where one can expect to have a great many uncontrolled factors affecting the performance of interest. For example, in flight tests to compare the effectiveness of several instrument configurations, variations in the weather, the equipment, the personnel and other flight conditions can introduce biases into the measures. To overcome this, an assumption is made that from flight to flight these unwanted conditions will vary at random and their effect eliminated from the comparison among instruments if enough flights are made. An alternative and preferred approach would be to measure the unwanted conditions if they can't be controlled and to introduce them into the experiment as a variable if they can. The result of this latter approach will be to reduce the magnitude of the effort and to increase the precision of the

information. More important, the information relating instrument differences with other conditions in the real world would be more generalizable.

Third, while ordering a series of experimental conditions randomally can neutralize the effects of an unwanted systematic effect occurring concomittantly, there are times when we may want to measure the effect (e.g. learning in a training study). Attempting to do this with fully randomized designs would be the less efficient procedure to use to isolate these sources of variance.

Fourth, in any single experiment involving a small number of conditions, randomization can often lead to unsuitable arrangements. By chance, it is possible to obtain what appears not to be a random presentation order but a very meaningful one. When this occurs, several approaches can be or have been employed. Representing one point of view, the strict statistician might say that the arrangement that was obtained should be used since it is the result of a random selection process and that in the long run over many experiments it will satisfy the criteria for randomness. However, Youden (1958, p. 16) notes "the difficult position of the statistician who rules against such systematic sequences when advanced on the grounds of convenience and insists on it when it pops out of a hat." Later (p. 19), he adds: "Surely statisticians have a responsibility to evolve a philosophy that does not elevate the statistical state at the expense of the individual experiment". Of course, the behaviorial scientist has an equal responsibility not to blindly employ experimental techniques that will not serve his purpose nor result in valid data. Youden and others have proposed a form of "restricted" randomization that would reduce the possibility of unsuitable arrangements.

Systematic Arrangements

When systematic presentation orders are employed they will ordinarily result in a more precise estimate of the treatment effects. Furthermore, this precision enables effects to be detected more economically than with a randomized design. When experiments are very small, a systematic arrangement is the only practical way of minimizing sequence effects.

Fisher's (1949) position on the danger of using a systematic design has already been quoted. He noted that while real errors may be diminished by such systematic arrangements, the apparent magnitude of the error is increased. This causes less reliance to be placed on the results of primary interest although the experimenter has successfully been able to exclude large unwanted components of error from the comparison. But Fisher (1949, p. 77) goes on to say: "It should be noted that this unfortunate consequence only ensues when a method of diminishing a real error is adopted, unaccompanied by their elimination in the statistical analysis." Fisher thus was not adverse to systematic designs per se, for he advocated the use of blocking and Latin square designs, but to the misuse of such designs. He was concerned with those experimenters -- and human factors specialists have been among the more delinquent in this regard -- who counterbalance or otherwise systematize their presentation orders but fail to remove the order effect from the error term. His point is that to make a valid test of statistical significance, if one introduces systematic presentation orders into a design to obtain a more precise comparison among treatments, these effects must be isolated statistically during the data analysis.

Among applied statisticians there is essentially no real disagreement as to whether or not a randomized or systematic arrangement should be used. Daniel and Wilcoxon (1966, p. 259) discuss the history of randomization since Fisher, pointing out that many factorial experiments using systematized presentation orders have produced useful, even valid, results. They said: "Following Fisher every text on the planning of experiments insists that the allocation of 'treatments' to 'experimental units' be in some degree random. But almost as soon as it is urged, sometimes dogmatically, the demand is modified and for obvious reasons. Full randomization would in many situations guarantee results which while entirely valid, would not detect any effects."

Cox (1958, p. 77) expressed this opinion: "If we have good knowledge of the form of the uncontrolled variation and if a systematic arrangement is much easier to work with, it may be right not to randomize.... What is important, however, is to randomize except where there is a very good reason not to, to understand that the conclusions from nonrandomized experiments depend on the correctness

of what is assumed about the uncontrolled variation and to state this explicitly in reporting the experiment."

Daniel and Wilcoxon (1966, p. 260) placed the emphasis the other way when they wrote: "Randomization is used, then, after we have exhausted our knowledge of the behavior of the system under study and have taken serious steps to control what can be controlled. It is done to make the odds as high as possible that we do not deceive ourselves by identifying unknown and unforseeable shifts in response as the results of the treatments we apply."

Cochran and Cox (1957, p. 7) wrote: "The occasions on which randomization is required vary with the type of experiment and must be left to the judgment of the experimenter." This of course assumes that the experimenter is fully aware of the alternatives and their consequences.

Thus, it would seem that for most human factors engineering research designed to acquire generalizable, quantitative, multifactor data systematic arrangements would be preferred over randomized presentation orders, with randomization being used only as a last resort. This conclusion is reached since for this class of research the following conditions are generally true:

1. Precise estimates of treatment effects will be more important than valid tests of statistical significance (Simon, 1971; 1973).
2. The ability to predict from experimental data will depend on how well all critical sources of variance affecting behavior -- including those artifically introduced in the experiment -- are identified and isolated.
3. Sequence effects in these studies can usually be anticipated and with some exploratory effort, their nature fairly well defined.
4. Economy of data collection is an important consideration.
5. An investigator may have a primary interest in the sequence effects.

Balance and orthogonality. Systematic designs are planned in a way that will permit certain effects to be isolated and measured and not to be confounded

with other less interesting effects. Stated in another way, systematization tends to seek balanced plans that orthogonalize effects. Since the terms "balance" and "orthogonality" do represent the primary characteristics and goals of the majority of the experiments to be discussed in this report and will be referred to frequently, it is appropriate to explain in a somewhat over-simplified fashion just what they mean. A design is balanced when certain specified conditions occur equally often in regard to other specified conditions, for example, each treatment precedes every other treatment equally often. When this is so, the residual effect of every treatment could carry over equally often to the direct effect of every other treatment, but not to the treatment itself. This distinguishes a balanced condition from an orthogonal one, for direct and residual effects would not be orthogonal until residuals from every treatment carried over to direct effects of every treatment. Orthogonality is found only when every condition of one factor or effect appears equally often in conjunction with every condition of another factor or effect. For example, when every subject is tested equally often on every treatment the effect of subjects and treatments are orthogonal. Thus the "equal frequency" requirement can lead to designs that are only partially balanced across a set of conditions, i.e. balanced in one regard but not another. On the other hand, for two effects to be orthogonal (i.e. statistically independent) the balance must be complete between factors.

WHEN RANDOMIZATION SHOULD BE USED

Although most of the techniques for handling sequence effects described in this report involve systematic designs, there still is a place for randomization. Whenever the experimenter has a choice, particularly in the selection and assignment of subjects, treatments, and even the Latin squares of an experimental design, and there is no known or rational basis for choosing one or the other, then the selection and/or assignment should be done in some random manner.

In many of the papers cited in this report, the writers suggest how to include randomization along with their systematic plans. In this regard, only general suggestions will be made here. An investigator who uses the plans for handling sequence effects presented in this report should not take the matter of randomization lightly, when it can be applied. When no better method is known,

randomization still represents the best method, the best insurance, the best precaution against unknown, uncontrolled biases.

Randomization ought to be employed in conjunction with the systematic designs in this report to:

1. Select the sample of experimental subjects from the larger homogeneous population.
2. Select the levels of experimental factors (except where fixed levels are to be used, the more common case in human factors engineering research).
3. Allocate the subjects to the sequences of the experimental design.
4. Allocate groups of subjects to the different blocks of an experimental design. However, within blocks, subjects should be selected to be as homogeneous as possible (See Finney, 1956, p. 60; Lucas, 1957, p. 229).
5. Select an initial sequence or particular overall pattern of the experimental design. (Designs can generally be permuted into other designs of equal quality. Experimenters should not get into habit of using the same pattern over many experiments.)
6. Allocate the treatments to the treatment symbols in the completely balanced experimental designs.
7. Decide which comparisons will be of equal precision when partially-balanced incomplete block designs are involved, or which treatments will not appear in the same block or sequence. (Note: Practical considerations may make it more desirable not to leave this decision to chance.)

CHAPTER II

PROCEDURES THAT HELP TO AVOID OR MINIMIZE SEQUENCE EFFECTS

Through experience gained from previous research and observations made during the preliminary phase of the particular experimental program, an experimenter should be able to identify what conditions are most likely to create sequence effects and the nature of these effects. If the investigator is not interested in such effects and does not want them to bias the experimental estimates of interest, he can prepare for and conduct the data collection phase of his experiment in a way that will enable them to be avoided or minimized. Six not necessarily mutually exclusive procedures for this purpose are described here.

USING DIFFERENT SUBJECTS ON EACH EXPERIMENTAL CONDITION

Poulton (1973, p. 119), long an advocate of different-subjects-for-different-conditions, stated: "The day should come when no reputable psychologist will use a within-subjects design, except for special purposes, without combining it with a separate-group design." If different groups of subjects are run on each experimental condition, then many of the sequence effects associated with within-subjects designs would be eliminated. Certain short term condition-to-condition transfer effects and long term trend effects such as learning and fatigue would have no opportunity to occur.

But even so, this approach does not completely guarantee that certain sequence effects will not be present. For example, while trend effects associated with subjects may be eliminated by this approach, trend effects associated with the equipment and the environment are not. If there are changes in the equipment throughout the day, as new conditions (with new subjects) are being tested, a trend can run through the data that can bias the effects of interest. Experimental designs to be described later in this report might still be appropriate to use. At the same time, the experimenter must monitor and maintain equipment adjustments and eliminate or measure environmental variations that might affect operator performance.

Certain practical considerations may prevent the use of different subject-groups in all experiments. First of all, it may not be feasible because of the limited number of available subjects. This is particularly the case if truly multi-multivariate studies are to be performed. Second of all, the nature of the experiment may not justify the increased costs arising from the additional measurements generally associated with separate-group designs. Particularly early in an experimental program, within-subject designs (properly planned and analyzed) can provide more good data more economically. In line with Poulton's suggestion, within-subject designs could be used initially to obtain enough information about a great many conditions to make intelligent decisions concerning follow-up studies of more modest size using separate-group designs.

ALLOWING ADEQUATE REST PERIODS BETWEEN TRIALS

When trial-to-trial transfer is likely to occur, and when each trial is particularly fatiguing for the operator, sequence effects can often be reduced by introducing adequate rest periods between trials. The assumption is often made that for certain tasks, these transfer effects or residuals that carry over from the previous treatment will be inversely proportional to the time between each pair of treatments.

One disadvantage of this approach, however, is that it increases the temporal demands on the individual subjects as well as the data collection time of the complete experiment. Furthermore, in certain cases, new problems may arise depending on what the subject does during his rest periods. If the rest periods are too long, subjects may grow bored and a portion of each trial must be spent adjusting during a new warm-up period.

For certain types of carry-over effects, rest periods may be totally ineffective. For example, in certain target acquisition studies, once a subject has learned the location of a target under one level of illumination, any reasonable interval between trials is not likely to cause him to forget that location when asked to find the target again under another illumination level.

INTRODUCING A STANDARD "DUMMY" CONDITION BETWEEN EXPERIMENTAL CONDITIONS

An experimenter might expect to absorb first residual effects by introducing into his experimental design a standard "neutral" (dummy) condition between each of the experimental conditions of interest. In using this procedure of course it is assumed that either there will be no residual effects from the dummy conditions, or if there is, the effects will not differ as a function of the treatment that follows.

One disadvantage of using dummy treatments between experimental conditions is that they increase the size of the experiment. This extends the total work period and increases the chances that an observer may become fatigued before the sequence is completed. If the experimental conditions are levels of a quantitative variable, the dummy treatment will be more similar to some levels than others and may affect performance on subsequent experimental conditions as a function of their proximity on the scale. Other "scale" effects such as those suggested by Poulton (1973) might also occur since the one standard (dummy) treatment occurs t times to each single time the other t treatments occur. When the experimental conditions are qualitatively different, then there is always the special problem of deciding what the standard, dummy treatment should be.

RUNNING EXTRA TRIALS AND DROPPING SOME IN THE ANALYSIS

If each experimental condition is tested sequentially several times before changing to a new experimental condition and only the results of the last trial are included in the analysis, then residual effects of one period duration (or there about) will be absorbed during the first trial (or so) of each series. Theoretically then the final trial of these short series will be uncontaminated from such sequence effects. This technique at least avoids the problems of deciding what to make a dummy treatment or what to do during a rest period.

The major disadvantage of this procedure is that the length of the experiment increases with each replication. This may introduce long term trend effects such as fatigue.

USING HIGHLY PRACTICED SUBJECTS

With highly practiced subjects that are thoroughly familiar with the experimental procedures as well as the specific task involved, sequence effects such as learning trends may be reduced or eliminated. However, practice is more likely to reduce long term trend effects than it is to reduce short term transfer effects.

In certain circumstances, if the practice has been obtained using one of the tasks to be compared in the experiment, then this might negatively distort the performance on one or more of the new tasks. A common example of this is in the comparison of pilot flight performance using old and new instrument designs, when his many thousands of hours of experience has been built up using the old instrument design. The problem is similar to that described when using a dummy experimental condition, except that in this case, the over-training on one condition occurred before instead of during the experiment.

Another problem in using highly practiced subjects is that they are more difficult to obtain. This of course is not an acceptable excuse for failing to do do, particularly if using them will minimize certain sequence effects and if the results of the experiment are to be applied to situations involving highly experienced personnel. Conversely, if the experimental results (for example, of a training study) are to be applied to novices, data collected on experienced subjects may be misleading.

ISSUING EFFECTIVE INSTRUCTIONS TO THE SUBJECT

Clear instructions to the subjects can help reduce residual effects and even some trend effects. Properly executed, instructions can serve as a form of verbal, rather than motor, practice and thereby minimize undesirable changes in the way a subject attacks a problem when he finally realizes the purpose of

the task part way through the sequence. Appropriate instructions can also be used to motivate a maximum performance effort, which may partially reduce both transfer and trend effects, although not necessarily eliminating them.

PRECAUTIONS TO CONSIDER WHEN PROCEDURAL TECHNIQUES ARE USED

All of the above procedures should be considered in any experiment where sequence effects might occur. However, when using the procedures, the following precautions should be observed:

1. Certain procedures do not minimize all types of sequence effects with equal effectiveness.
2. Certain procedures while minimizing one type of sequence effect may actually increase another.
3. The effectiveness of certain procedures can only be assumed and is not capable of being measured.

In practice, procedural methods of minimizing sequence effects are most effective in combination with the design and analytic techniques to be described in the chapters that follow.

CHAPTER III

DESIGNS FOR HANDLING SHORT-TERM RESIDUAL EFFECTS USING A GROUP OF SUBJECTS

Change-over designs, as this class will be called, are used to improve the estimate of the direct effects of a treatment when it is suspected that residual effects from prior treatments may have been carried over to affect the performance at the time of measurement. The designs in this chapter are built on the assumption that residual effects persist for one or two trials.

Provisions for separating direct and residual effects are made in these change-over design by balancing the frequency and order in which treatments are applied in successive periods. This balance can be achieved by testing a single subject on an extended series of treatments arranged in all possible orders. These serially balanced sequence designs will be described in the next chapter. The balance can also be achieved by using more subjects and limiting the amount of testing per individual subject, while balancing the various treatment sequences among the subjects. These change-over designs, when properly designed and analyzed, provide the data required to enable direct and residual effects to be separated.

There are a number of different kinds of change-over designs from which the investigator must select the type that best meets his needs. Change-over designs can be classified according to the nature of the residual effects and their assumed relationship with the direct effects. These include such considerations as:

1. The number of periods over which the residual effects persist.
2. The nature of the relationship between direct and residual effects.
3. The degree to which direct and residual effects can be isolated.

Change-over designs in this report are divided into the following types:

1. Those assuming that direct and residual effects are additive.
   a. When the residual persists for one period only (FIRST RESIDUAL)
      - When direct and residual effects are not completely isolated (SINGLY-BALANCED DESIGNS)
      - When direct and residual effects are completely isolated (EXTRA-PERIOD BALANCED DESIGNS)
   b. When the residual persists for two periods (DOUBLY-BALANCED DESIGNS)
2. Those assuming that the residual is proportional to the direct effect.
3. Those assuming that there is an interaction between direct and residual effects.

Each of these types will be treated in the sections that follow.

ADJUSTING DIRECT EFFECTS FOR FIRST RESIDUAL EFFECTS

In certain experiments, the experimenter is interested primarily in the direct treatment effects, and any residual effects that might exist are nuisances, an artifact of the experimental design that ought to be minimized procedurally (Chapter II) and/or statistically. This could be the case in certain equipment design studies in which it is highly unlikely that the sequencing that occurs in the experimental design would be found in the operational situation. In that case, since precise knowledge on residuals would be unnecessary, change-over designs can be used that give more precise estimates of the direct and less precise estimates of the residual effects. In these designs, direct and residual effects remain partially confounded and direct effects estimates must be statistically adjusted for the residual effects. However as the number of treatments being investigated increases, the relative loss in precision in the residual effects to the direct effects decreases and these designs will be adequate for estimating both direct and residual effects.

When these designs are used, the investigator assumes that:

1. Direct and residual effects combine in an additive fashion.

2. Residual effects carry over for only one trial (first residual).

These designs, referred to as singly-balanced change-over designs, result in the following model:

TRIAL PERIOD:	I	II	III	IV	V
TREATMENT APPLIED:	A	B	C	B	A
MEASURED EFFECT:	A_d	$B_d + A_r$	$C_d + B_r$	$B_d + C_r$	$A_d + B_r$

where d and r stand for direct and residual effects respectively. The effect of a residual from the treatment in the previous period is added to the direct effect of the treatment applied in a current period, irrespective of what that current treatment might be.

Singly-Balanced Change-over Designs

Papers by Cochran, Autrey, and Cannon (1941), Williams (1949), Patterson (1950, 1951, 1952), Finney (1956), Davis and Hall (1969), and Patterson and Lucas (1962) and statistics textbooks by Cochran and Cox (1957), Li (1964), and Federer (1955) are among those that discuss methods of constructing and analyzing singly-balanced change-over designs. In some cases, for the sake of economy, the designs are only partially balanced, a fact that reduces their efficiency. *

Only some of all available designs are presented in this paper, selected on the basis of the criteria set forth in the Foreword. At least one singly-balanced change-over design is provided for studying from four to sixteen treatments inclusively. Additional designs are included when the number of factors to be studied are high in order to permit the user to trade-off the efficiency with

*Efficiency factors are the ratios of the variances per period per unit (subject) of complete Latin Square designs to the variances per period per unit (subject) of the change-over designs. (See Patterson and Lucas, 1962, pp 18-19.)

which the effects can be estimated against the number of subjects and total observations required to do the study. The almost total lack of experience with this class of design in human factors engineering research makes it difficult to guess at this time the relative importance of these alternatives.

Characteristics of the designs to be found in Appendix I-A of this report are listed in Table 1. The columns in that table are identified as follows:

t = number of different treatments (experimental conditions)

p = number of periods (trials)

b = number or blocks (Latin squares and rectangles)

k = number of units (subjects/sequences) within a block

n = bk = total number of units (subjects)

N = pn = total number of observations in complete design

E = efficiency factor for the estimated treatment (E_t) effects (ignoring residuals), adjusted direct (E_d) effects, and adjusted residual (E_r) effects

Source = the reference from which the design was taken, along with the design number found in that reference

Type = an identity related to the construction and analysis characteristics of the design

When efficiency factors are not given in this or any other table in this report, it was because they had not been calculated. However, the reader may be assured that they will be as high or higher as alternate plans of the same type for the same number of treatments.

Examples. Examples of the four types of singly-balanced change-over designs listed in Table 1 are shown in Figure 2.

A study of these four designs reveals the various balances and imbalances that exist. In the Latin square plan, the number of treatments equals the number of periods and therefore each subject is tested on every treatment. Since this is not the case in the remaining designs, they are referred to as "incomplete block" designs.

Table 1. Singly-balanced change-over designs to be found in Appendix I-A

t	p	k	b	n	N	E_t	E_d	E_r	Source*	Type of Change-over**
4	4	4	1	4	16	100	91	62	-	LS
5	5	5	2	10	50	100	95	72	P-L 10	LS
6	6	6	1	6	36	100	97	78	P-L 14	LS
7	4	7	2	14	56	88	80	57	P-L 19	BIB
8	8	8	1	8	64	100			-	LS
9	9	9	2	18	162	100			-	LS
9'	5	9	1	9	45	90	76	56	D-H 34	CCO
10	10	10	1	10	100	100			-	LS
10'	5	10	1	10	50	89	74	57	D-H 35	CCO
11	6	11	2	22	132	92	89	72	P-L 26	BIB
11'	5	11	1	11	55	88	67	52	D-H 36	CCO
12	12	12	1	12	144	100			-	LS
12'	5	12	1	12	60	87	71	55	D-H 37	CCO
13'	5	13	1	13	65	86	71	56	D-H 38	CCO
14	14	14	1	14	196	100			-	LS
14'	5	14	1	14	70	86	71	55	D-H 39	CCO
15'	5	15	1	15	75	85	68	51	D-H 40	CCO
16	16	16	1	16	256	100			-	LS
16'	4	4	8	32	128	71	65	45	P-L 129	PB
16''	5	16	1	16	80	85	68	53	D-H 41	CCO

* P-L Patterson and Lucas (1962)
D-H Davis and Hall (1969)
\- Classical LS-CO design
The numbers are the ones they used in their reports.

** LS Latin square
BIB Balanced incomplete block
CCO Cyclic change-over
PB Partially balanced incomplete block

Latin square (LS) change-over design (t = 4)

Periods	Subjects (Sequences) 1	2	3	4
I	1	2	3	4
II	2_1	3_2	4_3	1_4
III	4_2	1_3	2_4	3_1
IV	3_4	4_1	1_2	2_3

Direct treatment

Residual from previous treatment

Balanced incomplete-block (BIB) change-over design (t=7)

Periods	Subjects (Sequences) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
I	0	1	2	3	4	5	6	0	1	2	3	4	5	6	0	1	2	3	4	5	6
II	5	6	0	1	2	3	4	3	4	5	6	0	1	2	6	0	1	2	3	4	5
III	6	0	1	2	3	4	5	5	6	0	1	2	3	4	3	4	5	6	0	1	2
	Block 1							Block 2							Block 3						

Direct treatments

"Cyclic change-over" (CCO) designs* (t=6)

Periods	Subjects (Sequences) 1	2	3	4	5	6	7	8	9	10	11	12
I	1	2	3	4	5	6	1	2	3	4	5	6
II	4	5	6	1	2	3	6	1	2	3	4	5
III	5	6	1	2	3	4	2	3	4	5	6	1
	Block 1						Block 2					

Direct treatments

Partially balanced (PB) change-over designs (t=6)

Periods	Subjects (Sequences) 1	2	3	4	5	6	7	8	9	10	11	12
I	1	4	2	5	2	5	3	6	3	6	1	4
II	4	5	1	2	5	6	2	3	6	4	3	1
III	2	1	5	4	3	2	6	5	1	3	4	6
IV	5	2	4	1	6	3	5	2	4	1	6	3
	Block 1				Block 2				Block 3			

Association Scheme
1, 4
2, 5
3, 6

Figure 2. Examples of the four types of singly-balanced change-over designs listed in Table 1

*Davis and Hall (1969) number their t treatments from 0 to t-1. To promote consistancy among the designs in this section, CCO plans will be written with t treatments being numbered from 1 to t.

In the Latin square plan and in the balanced incomplete block plans, each treatment is preceded an equal number of time s by each other treatment. Although no treatment precedes itself, these designs are still referred to as being balanced.

With "cyclic change-over" (CCO)* designs and partially balanced incomplete block designs, each treatment is not preceded equally often by each other treatment. This leads to the designation, "partially balanced." For example, in the CCO design (Figure 2), treatment 3 is never preceded by treatment 5, nor treatment 5 by treatment 1. In the Partially Balanced Change-Over design, certain treatments are preceded more frequently by some treatments than others. For example, treatment 1 is preceded twice by treatment 4 (and vice versa) but only once by the other treatments. Similarly, treatment 2 is preceded twice by treatment 5 (and vice versa) but only once by the remaining treatments. A similar relationship would be found between treatments 3 and 6. These particular combinations are noted in the Association Scheme (also supplied along with the appropriate designs in Appendix I and used in the analysis of the data). Obviously first associates (with the most frequent comparisons) provide more precise estimates of effects than second associates (with the least frequent or absent comparisons).

The four types of change-over designs represent plans that differ in regard to the complexity of the analysis and the ease with which results can be interpreted. The size of the plan affects the degree of imbalance in the design. If an investigator prefers to make fewer observations or use fewer subjects over a shorter time period, he must pay for this with increased imbalance among the direct, residual, and other effects and a corresponding increase in ambiguity in the results. The practical consequences of these trade-offs will not be known until more experience has been gained using these plans in human factors engineering experiments. One might expect, however, that as the number of treatments increase, the less serious the imbalance will be when estimating the various effects.

*"Cyclic change-over" is a name given to this class of designs by Davis and Hall (1969). While many change-over designs are cyclic in construction, the designation, CCO, will be used to refer specifically to this group of designs.

The one partially balanced change-over plan and the eight CCO plans are included in Table 1 because of their marked economy as compared to the other designs. The investigator who is tempted to use one however should weigh the extra problems of analysis when these designs are used against the additional data collection effort when other designs are used. The little extra effort during the data collection stage will generally be worth it when faced with the problem of interpreting the results.

Construction

All possible methods of constructing change-over designs will not be described here; only the methods required for the types selected for this report will be discussed and these should be adequate for most practical purposes. The reader who is interested in knowing about other methods will find Patterson's (1952) and Patterson and Lucas' (1962) papers excellent ones with which to begin.

The most difficult task in constructing the plans listed in Table 1 is to determine the initial sequence (referred to as the "generating" sequence) of each block of the design. The actual method used to construct these generating sequences will be described in this report for only the Latin square plans. However to aid the reader who intends to use these designs, the generating sequence of every plan will be given in Appendix I, and in a few cases, the complete design will be given.

Methods of completing a design once the generating sequence is known will also be described here for all of the designs listed in Table 1. The cyclic solution is the most frequently used for the Latin square, balanced incomplete block, and CCO designs. Another solution, also described, is needed to complete the partially balanced incomplete block designs. The reader will have to use these methods in conjunction with the data supplied in Appendix I-A when a complete design is not supplied.

Constructing the generating sequence of a Latin square plan. Only one block and therefore one generating sequence is required for a singly-balanced

change-over design of the Latin square type when t (the number of treatments) is even. Two blocks are required when t is odd.

To construct the generating sequence of a LS plan when t is even, the numbers 1 through t are arranged in the following order to represent the sequence of t periods for the first subject:

Period: I II III . . . . . t

Treatment: 1, 2, t, 3, (t-1), 4, (t-2) ... ... (t+2)/2.

Each of the t numbers representing the different treatments will be used once. This sequence, when t is even, is constructed by writing down the first two numbers in their normal sequence, and to continue writing the remaining numbers in a normal sequence but skipping every other period until the t^{th} period has been filled. At that point, the normal sequence of numbering is continued in the reverse direction, filling in the previously vacant periods. The following may help to visualize the process when t = 8:

1, 2, _, 3, _, 4, _, 5,

8, 7, 6,

or

1, 2, 8, 3, 7, 4, 6, 5,

To construct the generating sequence of a LS plan when t is odd, the treatment numbers from 1 to t are ordered as follows:

1, 2, t, 3, (t-1), 4, (t-2) ... ... (t+3)/2

Thus for t = 7, the sequence in which treatments should be presented to the first subject in the first block is:

Subject 1: 1, 2, 7, 3, 6, 4, 5.

Since t is odd, however, a second Latin square, or block is required for complete balance. The sequences of the second block of t more subjects are obtained by writing the sequences in the first block in reverse. For example, the sequence in which the treatments would be presented to subject number

eight would be the reverse of the sequence presented to subject number one, thus:

Subject 8: 5, 4, 6, 3, 7, 2, 1.

Obtaining the generating sequence of BIB change-over and CCO designs. The generating sequence for all blocks of the BIB and CCO designs listed in Table 1 are given in Appendix I-A. If the reader wishes to develop additional plans, he can refer to the papers by Patterson (1952) and by Davis and Hall (1969) to find out how these sequences were constructed.

Completing the remaining sequences of the LS, BIB, and CCO designs. For the balanced designs described in Appendix I-A, once the generating sequence for each block has been determined, the remaining (t-1) sequences can in most cases be created by using a cyclic solution. The one exception is the design for t = 4, which is formed by a multiplication process.* Cyclic solutions, of course, are far from being the only ones that can be used to complete singly balanced change-over designs, but will be the only type described here. References already cited can be referred to if the reader is interested in learning of other techniques.

Using the cyclic solution to complete these designs, each new sequence is obtained by adding +1 (modulo t) to the treatment numbers of each preceding sequence. Thus, if the first (generating) sequence of the Latin square (t=8) were:

1, 2, 8, 3, 7, 4, 6, 5

the sequence of treatments for the second subject would be:

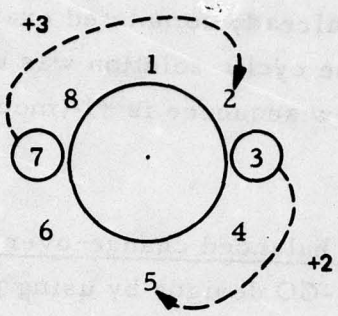
2, 3, 1, 4, 8, 5, 7, 6,

and for the third subject:

3, 4, 2, 5, 1, 6, 8, 7.

*This multiplication method is described in the footnote in the next section, "Constructing partially balanced change-over designs".

Note that when +1 is added to 8 is this example, it did not become 9; instead it becomes 1. "Modulo t" means that t (in this example, 8) is largest number possible, after which the count starts over again. This can be visualized by thinking of a series of numbers (modulo 8, for example) in a circle:



Moving clockwise around the circle and adding 2 to 3 (modulo 8), gives 5. Adding 3 to 7 (module 8), gives 2. This could also be thought of as 3 + 7 (modulo 8) = 10 - 8 = 2.

The successive sequences are obtained in the same way, i.e., adding +1 (modulo 8) to the preceding sequence. This is continued until there is a different sequence of the t treatments for each of t subjects. The complete design would be for t = 8:

	Periods (Trials)							
	I	II	III	IV	V	VI	VII	VIII
1	1	2	8	3	7	4	6	5
2	2	3	1	4	8	5	7	6
3	3	4	2	5	1	6	8	7
Subjects (Sequences) 4	4	5	3	6	2	7	1	8
5	5	6	4	7	3	8	2	1
6	6	7	5	8	4	1	3	2
7	7	8	6	1	5	2	4	3
8	8	1	7	2	6	3	5	4

Treatments

In this text, in the appendices, and in the original papers, the experimental designs are not always formatted the same and they are not always labelled. It is important that the reader ascertain for each design whether columns or rows represent sequences or periods. In some cases, an arrow has been drawn near the design to show the direction of the sequences across periods.

An examination of the already completed examples of BIB and CCO designs (Figure 2) will show how the cyclic solution was used. As in the case of the LS designs, each value in a new sequence is +1 (modulo t) more than the preceding value.

Constructing partially balanced change-over designs. Patterson and Luce (1962, p. 33) construct PB-CO designs by using partially balanced incomplete block designs developed by Bose, Clatworthy, and Shrinkhande (1954) to provide the initial generating sequences for each block, and a method employed in the construction of simpler change-over designs to complete the blocks. In their tables, Patterson and Lucas (1962), p. 49) indicate which Bose design and which change-over design should be used to construct each partially balanced change-over designs. Let us now see how this information was used to construct the example of the partially balanced change-over design for t=6 given in Figure 2.

Patterson and Lucas indicate that Design S1 in Bose et al (1954) is to be used to produce the initial generating sequences for each block and Design 5 in their own report would be a model of how the PB-CO is to be completed.

From Bose's report, we find that Design S1 separates six treatments replicated twice into three incomplete blocks of four treatments each, thus:

	(1)	1	4	2	5
Blocks	(2)	2	5	3	6
	(3)	3	6	1	4

The sequences of treatments in each of these three blocks are to be used as the first generating sequences of three new blocks that will make up the partially balanced change-over design.

Each of these blocks are to be completed using the same technique that Patterson and Lucas employed to construct their model, Design 5.* Design 5 is a t = 4 Latin square design, the first one found in Appendix I of this report and reproduced here:

Patterson & Lucas' Design 5 for t = 4:

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

However with incomplete blocks, the method is complicated because we are forced to work with incomplete blocks, that is, blocks in which all of the treatments are not present. The following substitution technique is easier when the model design (in this case, Design 5) is known.

The method of treatment symbol substitution is used to change the model design (in this case, Design 5) into a design developed from the generating sequence supplied by Bose's PBIB Design S1. Thus we take the first sequence from Bose's design, i.e. 1, 4, 2, 5, and substitute it for the first sequence of the model design, i.e. 1, 2, 3, 4, by putting a 1 in place of 1, 4 in place of 2, 2 in place of 3, and 5 in place of 4. These same substitutions would then be made throughout the model design to create the first block of the partially balanced change-over design shown in Figure 2, thus:

	Sequences in Block 1 (k)				
Period (p)	1	4	2	5	(Block obtained by substituting sequence 1425 into Design 5)
	4	5	1	2	
	2	1	5	4	
	5	2	4	1	

*The LS-CO design for t = 4 shown in Appendix I-A was not created by a cyclic solution. Instead it was completed using a multiplication process that works when (t+1) is a prime number. Once the generating sequence has been determined—in this example, it was simply the treatment values written serially from one to four—each subsequent sequence is formed by multiplying this first sequence by 2, then 3, 4, ... up to t (modulo t+1). Thus the third sequence in the example would be obtained by multiplying: 1x3=3 (modulo 5)=3; 2x3=6(modulo 5)=1; 3x3=9(modulo 5)=4; 4x3=12(modulo 5)=2, or 3, 1, 4, 2. Williams (1949) shows how this LS can be made equivalent to one derived with a cyclic solution (p. 152).

The remaining two blocks of the PB-CO as shown in Figure 2 would be constructed in the same way, but by substituting the other two sequences from Bose S1 design into Design 5. A total of 12 subjects would be needed for all sequences of the three block design.

Analysis

In all of the singly-balanced change-over designs, essentially the same sources of variance are examined. The typical partitioning of the degrees of freedom is shown in Table 2. Modifications of this general form must be made to meet special requirements. For example, when t is even, Latin square plans require only one block for a balanced design. Thus $b = 1$, and all sources pertaining to blocks are eliminated.

The most unique feature in the analysis of this singly-balanced design is the dual partitioning of the treatments effect (as indicated by the letters A and B in Table 2). This is done because the direct and residual effects are correlated with one another (i. e. are non-orthogonal). In order to evaluate the direct effects, they must be adjusted for the entangled residual effects. The same holds true if one wishes to examine the residual effects; they must be adjusted for direct effects.

The rationale for this adjustment may be clearer if the situation is seen graphically as in Figure 3. The two circles represent the sums of squares for the direct and residual effects which are correlated as indicated by the degree of overlap. If residual effects are partialled out first, then the adjusted direct effects are what's left over. If the direct effects are partialled out, then the adjusted residual effects are what's left over. The adjusted values are used in the F-tests of significance.

How this analysis would be performed will not be discussed in any detail in this report. In Table 3 however, references are given in which the method and numerical examples of the analysis of these singly-balanced designs can be found.

Table 2. Sources of variation and partitioned degrees of freedom in the analysis of variance of singly-balanced change-over designs.

Sources of variance	Degrees of freedom
Blocks	$b - 1$
Periods	$p - 1$
Blocks x periods	$(b - 1)(p - 1)$
Units within blocks	$b(k - 1)$
Treatments (direct and residual effects)	$2(t - 1)$
Error	$(n - b)(p - 1) - 2(t - 1)$
Total	$n\,p - 1$

---Alternate methods of partitioning the Treatments degrees of freedom---

A	Direct effects (eliminating residual effects)	$t - 1$
	Residual effects (ignoring direct effects)	$t - 1$
	Total	$2(t - 1)$
or		
B	Direct effects (ignoring residual effects)	$t - 1$
	Residual effects (eliminating direct effects)	$t - 1$
	Total	$2(t - 1)$

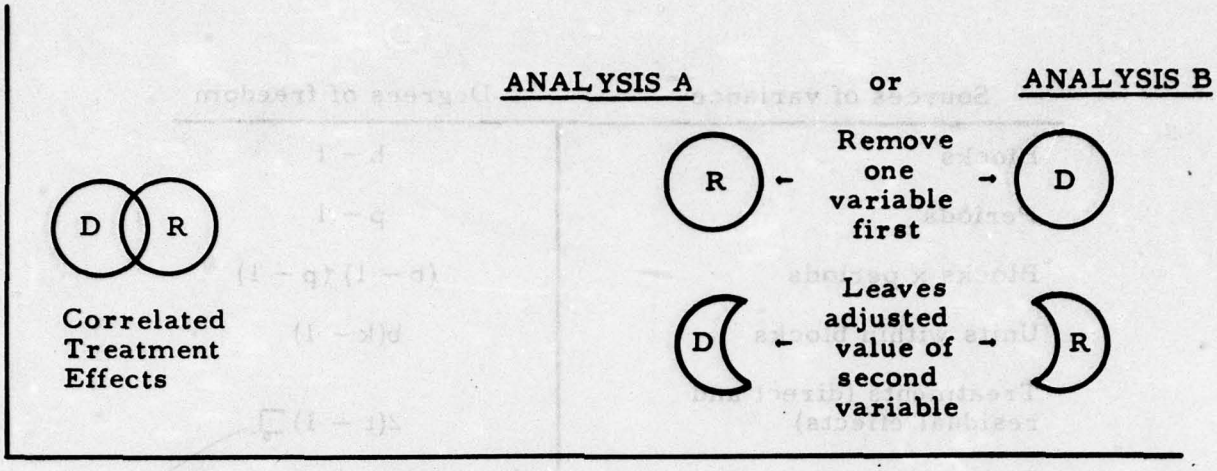


Figure 3. Illustrating the partitioning of correlated effects

Table 3. References on how to analyze singly-balanced change-over designs

Source	Arithmetic Method	Numerical Example	Type of Design
Cochran, Autrey, & Cannon (1941)		pp 946-949	Latin square
Cochran and Cox (1957)		pp 135-139	Latin square
Davis and Hall (1969)	pp 289	pp 290-291	CCO*
Federer (1955)	pp 446-449	pp 449-452	Latin square
Patterson (1950)		pp 375-377	
Patterson (1951)	pp 258-263 pp 268-270		Latin square PB
Patterson & Lucas (1962)	pp 7-10 pp 11-13	pp 26-29 pp 34-38	Latin square PB
Williams (1949)	pp 161-165	pp 165-168	Latin square

*Davis & Hall (1969) wrote: "A program (Fortran for CDC 3200 computer) has been written which analyzes any CCO design and enables suitable designs to be chosen for any given set of parametric values. Copies may be obtained from the authors on request." (p. 289)

ISOLATING DIRECT AND RESIDUAL EFFECTS

In the previous designs, treatments were presented in a balanced order so that each treatment preceded every other treatment equally often but never preceded itself. As a result, residual effects were not completely balanced with direct effects. In the designs to be described in this section, a class referred to as extra-period change-over designs, treatments are repeated in an extra period in such a way that within each design each treatment now follows itself as often as it follows each other treatment.

These designs would be used when the investigator assumes that:

1. Direct and residual effects are additive. That is to say, the residual effect for any treatment will be the same no matter with which direct effect it may be combined.
2. The residual effect will persist for only one period.
3. There is an interest in both direct and residual effects.

The *same* additive model operates with these designs as they did with the singly balanced change-over designs. The primary difference between two types of designs lies in the degree of precision with which the residuals can be estimated when compared to estimates of the direct effects and in the fact that direct and residual effects can be completely isolated from one another. Both are orthogonal to periods, but in this design, residual rather than direct effects are orthogonal to subjects (sequences).

Constructing Extra-Period Change-Over Designs

Lucas (1957) suggested adding an extra-period at the end of the Latin square change-over design in which the treatment of the previous period in each sequence would be repeated. This resulted in each treatment following every other treatment an equal number of times including itself, thus completing the balance between direct and residual effects. Patterson and Lucas (1959) (1962) have extended this approach to incomplete Latin squares and incomplete block change-over designs which are not considered here.

Using a t = 4 basic change-over design as an example, we add the extra-period to obtain the following plan:

		Subjects (Sequences) 1	2	3	4
Periods	I	1	2	3	4
	II	2_1	3_2	4_3	1_4
	III	4_2	1_3	2_4	3_1
	IV	3_4	4_1	1_2	2_3
Extra-period	V	3_3	4_4	1_1	2_2

Direct treatment effect

Residual effect (subscripts)

An inspection of this extra-period design reveals that:

1. One-period residual effects are orthogonal to the direct effects.
2. Residual effects are orthogonal to sequences (subjects).
3. Direct effects are non-orthogonal to subjects, but Lucas (1957, p. 227) says that "the degree of non-orthogonality is not great".
4. The amount of replication of the residual effects relative to the direct effects is a little greater than in the Latin squares.

The same results could have been obtained if instead of the extra period being added at the end of the sequence it were introduced at the beginning with the first treatment being the one that was repeated. Since this extra treatment may not be included in the analysis (although its residual effect combined with the second occurrence of the same treatment is included), there can be an advantage in having it first rather than last, for in first position it can serve as a warm-up trial.

Recommended extra-period designs are described in Table 4. All of these can be constructed by adding an extra period with treatment to corresponding singly-balanced change-over designs listed in Table 1 and given in Appendix I-A.

Table 4. Characteristics of recommended extra-period designs

t	p	k	b	n	N	E_t	E_d	E_r	Parent Plan from Table 1**
4	5	4	1	4	20	96	96	80	LS 4
5	6	5	2	10	60	97	97	83	LS 5
6	7	6	1	6	42				LS 6
7	5	7	2	14	70	84	84	72	BIB 7
8	9	8	1	8	72				LS 8
9	10	9	2	10	180				LS 9
10	11	10	1	10	110				LS 10
11	12	11	2	22	254				LS 11
12	13	12	1	12	156				LS 12
13*									
14	15	14	1	14	210				LS 14
15*									
16	17	16	1	16	272				LS 16
16'	5	4	8	32	160	69	69	59	PB 16'

*No extra-period designs have been supplied for t = 13 and 15. Such designs can be found in Patterson and Lucas (1962). The "best" designs available however are P-L design #53 for t = 13, requiring 39 subjects and 234 observations, and P-L design #150 for t = 15, requiring 60 subjects and 300 observations.

**LS, Latin square; BIB, balanced incomplete block; PB, partially balanced incomplete block

Extra-period cyclic change-over designs. None of the extra-period designs listed in Table 4 were constructed by adding an extra period with treatment to the Davis and Hall (1969) cyclic change-over (CCO) designs. This does not mean that CCO's could not be used if it were necessary to reduce the size of the experimental effort.

Davis and Hall (1969, p. 289) had this to say regarding the adding of an extra-period to their CCO designs:

> "CCO designs may be analyzed after any number of periods, and further periods may be added in such a way as to maximize E_d or E_r, these features adding great flexibility to the experimental situation. As in the designs of Patterson & Lucas, repetition of the final period increases E_r at the expense of E_d. However, orthogonality of direct and residual effects is not achieved by this means as in their designs."

They compare the efficiencies of the treatment, direct, and residual effects when the extra-period is and is not used in the CCO design (t=6) shown in Table 1. The percent efficiencies were:

	E_t	E_d	E_r
Basic CCO design	79	58	34
Extra period CCO design	74	72	61

Berenblut's alternative. Berenblut (1964) proposed a change-over design that would keep the residual and direct effects unconfounded, and do so without confounding direct treatment effects with sequences and subjects as is the case with the Lucas' extra-period designs. This was done by adding more extra-periods. An example of a Berneblut design for t = 3 is as follows:

		Subjects (Sequences)								
		1	2	3	4	5	6	7	8	9
	I	1	2	3	1	2	3	1	2	3
	II	3	1	2	2	3	1	1	2	3
Periods	III	2	3	1	2	3	1	2	3	1
	IV	2	3	1	1	2	3	3	1	2
	V	3	1	2	3	1	2	3	1	2
	VI	1	2	3	3	1	2	2	3	1

Each condition follows every other condition including itself an equal number of times -- in this example, five times. Direct and residual effects are orthogonal.

However, whereas Lucas' designs required $t(t+1)$ observations and t or $2t$ subjects for a complete design, Berenblut's designs required $2t^3$ observations and t^2 subjects. Namboodiri (1972, pp. 58-60) discusses the relative merits of these two designs and indicates that Lucas' designs are better for estimating first residuals. Because of its extra size and limited advantage, no further discussion of Berenblut's design will be made here. General methods of construction and analysis can be found in the original article.

Analysis

A typical partitioning of the degrees of freedom in an extra-period, singly balanced, change-over design is shown in Table 5. Some modifications may be required to meet special requirements, but the essential difference between this analysis and that of the basic singly balanced design (Table 2) without the extra period lies in the handling of the treatment variance. Since direct and residual effects are orthogonal to one another in this extra-period design, no alternate analyses are required.

How this analysis is performed will not be discussed in any detail in this report. In Table 6, however, references are given in which the method and numerical examples of the analysis of these extra-period change-over designs can be found.

Table 5. Sources of variation and partitioned degrees of freedom in the analysis of variance of extra-period designs

Variation	Degrees of freedom
Blocks	$b - 1$
Periods	$p - 1$
Blocks x periods	$(b - 1)(p - 1)$
Units within blocks	$b(k - 1)$
Direct effects	$t - 1$
Residual effects	$t - 1$
Error	$(n - b)(p - 1) - 2(t - 1)$
Total	$np - 1$

Table 6. References on how to analyze extra-period designs

Source	General Method	Numerical Example	Special features
Lucas (1957)	pp. 230-233	pp. 234-237	Extra-period design Blocking Missing data
Patterson & Lucas (1959)	pp. 122-127	pp. 127-128	Extra-period design ANOVAs including and omitting first period
Patterson & Lucas (1962)	pp. 11-12	pp. 30-32	Extra-period design
	pp. 13-14	pp. 34-38	Partially-balanced extra-period design

Basic Change-Over Versus Extra-Period Designs

Extra-period designs improve the quality of information on residual effects. If there is a special interest in residual (transfer) effects, as there might be in training situations, then the use of extra-period designs should be given special consideration. If the cross-over effects however are artifacts of the experimental design and are unlikely to be found in the field -- as in the case when Latin square designs are used in the comparison of several types of equipment -- then the basic change-over design should suffice. However residual effects might be of special interest in an equipment design study when its purpose is to compare performance levels between old and new devices, when the latter is to be substituted for the one currently being used.

Patterson and Lucas (1959) discuss certain biases in the estimates of error that occur in extra period and standard change-over designs, suggesting that with extra-period designs the biases will be somewhat greater. They indicate that a better estimate of the error of residual effects (which is overestimated if there is a linear trend in the sequence of measurements) can be obtained with an analysis of variance that omits the first period in the calculations. However they add: "The calculation of this estimate is sometimes difficult and will not normally be worthwhile unless a critical test of residual effects is required." (p. 125) The reader should pay particular attention to the word, "critical", for in the typical human factors research, the result of an F-test alone should never be critical; it must be tempered considerably by the interpreter's judgement (See Simon, 1971, Objective Two).

Lucas (1957, p. 228) wrote: "The advantages of the extra-period designs over the regular Latin square designs decreases as the number of treatments is increased. In fact, when the number of treatments is six or more, the incomplete Latin squares likely would be chosen, simply to avoid using too many periods". However, it should be remembered that Lucas wrote in the context of dairy farming and the same considerations may not apply when one is using human subjects in equipment design and training research. For cows, the extra period may not fit their lactation schedules, but in many human factors studies, including an extra-period may be the easiest part of the experimental

effort. Certainly it is worth it if it simplifies the analysis and the interpretation of the data.

ADJUSTING DIRECT TREATMENT EFFECTS FOR RESIDUALS THAT LAST FOR TWO PERIODS

In the previous sections, designs were balanced for residual effects that persisted for only one period. There are times however when an experimenter may believe that this limit is too severe and that it may be necessary to balance the experimental conditions for effects that persist for two periods. The "doubly-balanced change-over designs" described here can be used when:

1. Residual effects affecting a particular treatment carry over not only from the preceding treatment (first residual) but also from the treatment two periods before it (second residual).
2. The magnitude of the residual effect from any treatments is not dependent upon the particular treatment in the period in which the residuals are being measured.

The effects of these two conditions can be illustrated as follows:

Trial Period	I	II	III	IV	V
Treatment Applied	A	B	C	B	A
Measured Effect	(A_d)	$(B_d + A_{r1})$	$(C_d + B_{r1} + A_{r2})$	$(B_d + C_{r1} + B_{r2})$	$(A_d + B_{r1} + C_{r2})$

where subscript d stands for a direct effect and subscripts r1 and r2 stand respectively for a residual from the treatment in the preceding period and a residual from the treatment two periods back. The size of the first residual of Treatment B is the same in periods III and V whether it combines additively to the direct effects of C or A respectively.

Characteristics of Doubly-Balanced Change-Over Designs

E. J. Williams (1949; 1950) is generally credited with having published the first description of doubly-balanced change-over designs. In order for direct effects to be balanced with respect to each of the two residual effects and for the first and second residual effects to balanced with respect to each other, the following conditions must hold:

1. Every treatment immediately precedes every other treatment equally often.
2. Every treatment precedes every other treatment by two positions equally often.
3. Every pair of treatments immediately precedes every other treatment equally often.

This third condition will also hold when each ordered trios (e.g. 123, 231, 312, 132, 321, 213) of treatment occur equally often, or when each ordered pair of treatments occurs equally often at the end of the row (sequence) of an otherwise balanced design. When these conditions hold, the designs are also balanced for the residual of the interactions of the two preceding treatments on the third, although this effect will not be examined in the analysis. For this balance, $t(t-1)$ sequences of t treatments are needed. Thus for eight treatments, 56 sequences (or subjects) are required.

Williams (1949, p. 151) notes that these designs are also balanced for the effect of subsequent treatments as well as for previous treatments stating: "Such a possibility is not so academic as at first appears. For example, in carrying out an examination on school children, the result obtained in one test may well be influenced by the anticipation of the subsequent tests, to a greater extent, than the actual result of previous tests."

Construction

There are a number of different methods of constructing doubly-balanced change-over designs, none of which are suitable for all values of t (number of treatments). Some are easy to use, some require advanced mathematical

knowledge, and some can be developed from data given in other sources for other purposes. Only designs for handling from four to ten treatments will be considered here; beyond that number, the number of subjects and observations required for complete balance is excessive.

Methods of constructing doubly-balanced designs can be found in the following references:

1. When t equals an odd prime value, e.g. 5, 7, 11, 13, and so forth, Williams' (1949, pp. 156-159) method can be used. It is easy to understand, but was incompletely described in that paper. Nair (1967, p. 215) points out that Williams' algorithm is good only when the odd prime equals 2r - 1, where r is odd. Another algorithm is required, which Nair supplies, when the odd prime equals 2r - 1, where r is even.
2. When t equals a power of a prime, e.g. 4, 8, 9, 16, and so forth, Williams' (1949, pp. 159-161) method can be used but requires a knowledge of advanced mathematics -- Galois fields -- even to complete the designs after he supplies basic information needed to start them.
3. Complete sets (t-1) of orthogonal Latin squares, when they exist, are doubly-balanced change-over designs. Fisher and Yates (1963) supply them for t = 4, 5, 7, 8, and 9, and these have been used for the designs in Appendix I-B of this report. Stevens (1939) describes how to construct completely orthogonalized Latin squares.
4. Williams (1950, pp. 353 and 355) states that Dudeney's* "Round Table Solutions" can be used to construct doubly-balanced change-over design. The t = 6 design in Appendix I-B is based on a Dudeney solution. Dudeney (1943) supplies round table solutions for all values of t up to 12.

Less efficient estimates of second and higher-order residual effects can be made with partially balanced incomplete block designs. Of these types,

*Dudeney's Round Table Solutions are given in Appendix II-E.

of singly-balanced designs discussed earlier in this report, the analysis of second residual effects was considered in the following cases:

1. Patterson and Lucas (1962) supply designs for t = 4 to 10, and also 13 that can be analyzed for multiple residual effects. Many but not all of these designs were derived from complete sets of orthogonalized Latin squares (designated with a -0 in their Table 5.0, pp. 39-43). In their designs, they did not ordinarily use all rows of the orthogonal Latin squares, thus achieving an economy in data collection at the increased cost of a lowered efficiency in estimating the effects. Efficiency falls off rapidly with these designs as treatment, direct, first residual, and second residual effects are calculated in that order. They also point out (p. 21) that given the property of the orthogonalized Latin squares, after three periods have been tested, an analysis can be made at the end of any period, a feature that "is useful if reports are required on the progress of an experiment or if an experiment is prematurely terminated through misfortune or for some other reason." For balanced designs, t(t-1) subject are required. They also supply some partially balanced designs requiring fewer subjects and fewer observations, but result in a greater inefficiency of estimating the effects.

2. The cyclic change-over (CCO) designs proposed by Davis and Hall (1969, see Appendix I-A) can also be analyzed for higher-than-first order residual effects. The drop in efficiency when estimating higher-order residual effects is considerable.

Designs

In Appendix I-B, doubly-balanced change-over designs for from four to nine treatments are supplied complete. With the exception of the t = 6 design supplied by Williams (1950) the remainder are the complete sets of orthogonal Latin squares given by Fisher and Yates (1963).

Analysis

Williams (1950, pp. 355-360) explains how doubly-balanced change-over designs are analyzed and provides a numerical example of the analysis. In this he does not consider the interactions among pairs of residuals, which in fact becomes part of the error variance. In these designs, the attribute of balance does not necessarily imply orthogonality among the factors. Thus, position (trial) effects will be orthogonal to all other effects. Subject (and sequence) effects will be orthogonal to direct treatment effects, but not to residual effects. Direct effects are not orthogonal to residual effects and first and second residual effects are not orthogonal to each other.

The non-orthogonality of the direct and residual effects in these designs, as in the case of the singly-balanced designs, require that adjustments be made in the analysis. Since estimations of correlated variables depend upon the order in which the estimates are made (i.e. which are adjusted for which first), several alternative analyses are possible. With the singly-balanced designs, with only direct and residual effects, two orders for analysis were possible. With doubly-balanced designs, with direct and first and second residual effects, there are six possible orders for analysis. However, Williams (1950, p. 359) proposes only four of these as being of any practical value.

To estimate the three treatment effects (D, R_1, and R_2) after they have been adjusted for the other two, the treatment sums of squares would be partitioned in the following ways:

1. Direct effects, adjusted for first and second residual effects
   First residual effects, adjusted for second residual effects
   Second residual effects, unadjusted

2. First residual effects, adjusted for direct and second residual effects
   Second residual effects, adjusted for direct effects
   Direct effects, unadjusted

3. Second residual effects, adjusted for direct and first residual effects
   Direct effects, adjusted for first residual effects
   First residual effects, unadjusted

He also proposes that the third partition above be made along with:

> Second residual effects, adjusted for direct and first residual effects
> First residual effects, adjusted for direct effects
> Direct effects, unadjusted

would be needed to test the significance of direct and first residual effects if second residuals are to be ignored. In the actual calculations, of course, the unadjusted term is removed first, followed by the remaining terms in order as more and more adjustments are required.

Patterson and Lucas (1962, p. 15-16) also provide the calculations required to analyze multiple residual effects in their designs including some with and without extra periods. They describe how with designs derived from complete sets of orthogonal Latin squares (p. 21) the standard analysis can be modified as follows:

1. The results can be analyzed at any stage after all subjects have run more than two periods, using the standard method appropriate to a design with the reduced number of periods. This feature is useful if reports are required on the progress of an experiment or if an experiment is prematurely terminated.
2. If there are signs that there are "warm up" effects during the first few periods of the completed design, a neat analysis can still be carried out by ignoring the results of these treatments.
3. Each of any set of contrasts between the periods can be analyzed separately and the results later combined in any desired manner. This can be used as Patterson (1950; 1951) did to allow for trends.

Considerations

Lack of experience with these doubly-balanced change-over designs for behavioral research makes it difficult to anticipate their usefulness.* Certainly they become costly in the requirements for subjects and total observations for even a few treatments. Even then the efficiency with which second residuals can be estimated is low. At present it is difficult to estimate how large second residuals will be, but it is almost a certainty that they will be considerably lower than first residual effects and probably negligible unless first residual effects are large. The designs are included here primary for completeness, to let the reader know that such techniques exist and to provide him with references should be desire to learn more about them. Without good reason for suspecting that residual effects will carry-over for two periods in any practical magnitude, the added size and complexity of the design and the analysis should discourage the use of these designs. If, on the other hand, experience tells us that these effects cannot be ignored, one may want to look into the use of partially-balanced incomplete block designs that will enable first and second residual effects to be estimated at less cost.

ADJUSTING FOR FIRST RESIDUALS THAT ARE PROPORTIONAL TO DIRECT EFFECTS

Patterson and Lucas (1962) provide a means of analyzing certain change-over designs to handle the case when the first residual effect of a treatment is proportional to the magnitude of the treatment effect in the same period as the residual.

They write: "The following methods can be used to investigate a possible linear relationship between direct effects and first residual effects when second, third, ..., residual effects can be assumed to be zero. The relationship will be of the following type:

$$r_i = \theta t_i$$

i = period
t = treatment
r = residual

*Namboodiri (1972, p. 58) suggests that "in clinical trials and psychological experiments, two-period change-over designs are likely to be very useful."

for all treatments, where θ may be known or unknown. This analysis... ... is only worth making if the earlier analysis has revealed substantial differences between direct effects and between residual effects." (p. 16-17)

Because of their unique symbology and the need to understand this method in the context of the rest of their report, it will not be discussed any further here. The interested reader should refer to Patterson and Lucas' (1962) paper. In it, they include a method of testing to see if residuals are proportional to treatments in the same period in the conventional balanced change-over design (p. 17), partially balanced change-over designs (p. 18), and extra-period designs, balanced or partially balanced (p. 18).

Finney (1956) discusses briefly some of the theoretical consideration in this and other change-over designs. When the residual is proportional to the immediate treatment, this is a special case of his Model III in which all residuals are proportional in different degrees to the immediate treatment (Model III. 1) and when only the first residual is proportional (Model III. 2). He illustrates the differences between this model of a change-over design and other types (Table 1, p. 46) and mentions proportional designs briefly on p. 47. He states that the proportion "will usually be positive and less than unity, and in the more general case where a proportion is attached to first, second, third, etc. order residuals, they would almost certainly be a descending series."

It should not be overlooked that while Patterson and Lucas were describing designs in a general sense, they had applied them in the context of argicultural and dairy experimentation. Finney wondered whether they would be equally applicable for biological assays. Some analysis and better still some preliminary effort should be devoted to discovering whether the additive or probabilistic models of change-over designs fit the specific tasks to be investigated. To what extent does this probabalistic model of a change-over design realistically fit the behavior found in human factors problems?

ADJUSTING FOR RESIDUALS WHEN DIRECT AND RESIDUAL EFFECTS INTERACT

In learning and transfer experiments, situations can be found in which the carry-over effect from one treatment is not always constant. Instead, it's magnitude will vary depending on which treatment follows. Statistically, this implies that there is an interaction between the residual of treatment A and the direct effect of treatment B. Unfortunately, only a few papers have dealt with the creation of balanced change-over designs for this class of sequence effect, and then only within a limited scope. The major difficulty lies in the size of the design necessary for balance. The designs are discussed here, albeit briefly, primarily to let the reader know that such designs exist and could be developed should such an effort be warranted.

The designs being discussed here can be used when:

1. The residual effect from a treatment is expected to carry-over only to a treatment in the next period (first residual);
2. The magnitude of the residual effect will depend upon the treatment to which it carries over (D X R interaction);
3. Only the linear direct-by-linear residual interactions need be estimated.

Two sets of designs will be described. One set was developed by Patterson (1973) for handling from five to 16 treatments. The other set includes three specific designs developed by Mason and Hinkelmann (1971) for simple factorial designs of two and three factors. The data required to construct the design or the complete designs are given in Appendix I-C for all conditions listed in Table 7.

Estimating Direct-by-Residual Interactions in One-Way Designs

Bereblut (1968) describes methods of constructing and analyzing change-over designs with four or five treatments, so that linear residual-by-linear direct effects could be estimated. The single degree of freedom for that interaction and those of the residual effect are orthogonal. The design was proposed

Table 7. Designs found in Appendix I-C that adjust for carry-over effects when direct and residual effects interact

Number of Factor	No. of treatments per factor	No. of Periods	No. of Subjects	Total No. Observations
1*	5	10	25	250
	6	12	36	432
	7	14	49	686
	8	16	64	1024
	9	18	81	1458
	10	20	100	2000
	12	24	144	3456
	16	32	256	8192
2**	2	4	8	32
	3	6	18	128
3**	2	6	24	144

*From Patterson (1973). Data needed to construct the design is provided.

**From Mason and Hinkelmann (1971). Complete design is provided.

mainly to test whether or not direct and residual effects were additive rather than to estimate the effect of the interaction.

Patterson (1970) discusses tests of nonadditivity in change-over designs limited to a single quantitative factor at four levels. He also discusses how these designs could be used to handle the factorial case. Patterson shows (pp. 541-543) how to construct these designs and suggests that they are better than Berenblut's for estimating direct-by-residual interaction effects. The designs can estimate the linear direct-by-linear residual interaction in any period after the first one but with decreasing efficiency. Patterson shows how

up to one-third of the information on the interaction effect could be lost by an unsuitable choice of the generating sequence for the design. Sequences for a four treatment design are supplied (p. 544).

In a subsequent paper, Patterson (1973) extends some of Quenouille's (1953) change-over designs for a limited number of conditions to handle any number of treatments and to estimate direct-by-residual interaction effects. On pages 41-42, he supplies the mathematical basis and general rules for constructing change-over designs involving a linear interaction model derived from cyclic permutation matrices. With t treatments, each design requires t^2 subjects who are tested for $2t$ periods each. He explains how the difference values are obtained with which the first row of the design can be constructed, and then how to complete the remainder of the design. This method is described here; although the description is taken directly from Patterson's explanation, it is presented in a form that minimizes mathematical symbology.

Determining the difference values. Patterson denotes three kinds of "difference values" by the symbols μ_m, λ_m, and μ_r. For simplicity here, we will refer to them as X, Y, and Z respectively. Without further discussion, the method of arriving at these values will be described below.

For a design of t treatments, it is first necessary to arrange according to Patterson's rules all of the integers from 0 to (t-1) into a difference matrix that looks like this example for $t = 5$:

	r1	r2	r3
X	0	3	2
Y	1	4	

Each matrix has two rows, X and Y, and r columns, where $r = t/2$ when t is even and $r = (t + 1)/2$ when t is odd. When t is odd, there will be no value in the Y row of the r^{th} column. The single value in X/r3 in this case will be designated Z.

By trial and error (and ordinarily without much difficulty), the matrix is constructed by using each of the integers from 0 to (t-1) once in locations that

make the sum of the X and Y values in each column coprime with t. For one number to be coprime with another, no number that is a factor of one can be a factor of the other (except of course the number 1). For example, the numbers 2 and 5 are both coprime with number 13, but only 2 is coprime with 15 (since 15 = 3 x 5), and neither are coprime with 20 (since 20 = 2 x 2 x 5).

It can be seen how the sample matrix for t = 5 satisfies all of the requirements cited above. Since t = 5 is an odd number, there are (t+1)/2 or (5 + 1)/2 = 3 columns in the matrix, and the last column has no Y value. The sums of the columns for this t = 5 matrix are 1, 7, and 2, all of which are coprime with 5.

Difference matrices for t = 5 through 16 are given in Appendix I-C from which these designs can be constructed.

Writing the initial row of the design. From each complete (i.e., both an X and Y value) column in the difference matrix, 2t values of the first row will be determined. With only an X value, as is the case when t is odd, only t values will be determined. Thus for t = 5 in which there are two complete columns and one half-column, there will be $2t + 2t + t = 5t = 5x5 = 25 = t^2$ values assigned to the first row of the design. Patterson refers to each set from a column as a treatment cycle; these are strung together to create the entire first row. This first row when completed will provide the treatments of the first period for each of the t^2 subjects.

For t = 5, the first column of the difference matrix is 0 and 1. To determine the 2t values of the first treatment cycle, first the number 0 is written down, and then 0 and 1 are added alternatingly to the previous number in the series, thus:

Write down 0

Add 0 and get 0

Add 1 and get 1

Add 0 and get 1

Add 1 and get 2

Add 0 and get 2

and this procedure is continued until 2t = 10 (in this example) numbers have been determined. The entire series for the first treatment cycle in this example would be:

0 0 1 1 2 2 3 3 4 4 .

This is repeated to get the second treatment cycle using the numbers 3 and 4 in the next column. Thus:

Write down 0
Add 3 and get 3
Add 4 and get 7
Add 3 and get 10
Add 4 and get 14
Add 3 and get 17

and this is continued until 2t = 10 (in this example) numbers have been determined. This entire series for the second treatment cycle would be:

0 3 7 10 14 17 21 24 28 31

However, since for five treatments the treatment values only run from 0 through 4, this series must be changed by rewriting the series modulo t (in this case modulo 5). Thus:

0 3 2 0 4 2 1 4 3 1

which lists each treatment twice. The total series now with the two treatment cycles combined is:

0 0 1 1 2 2 3 3 4 4 0 3 2 0 4 2 1 4 3 1 .

There are still t more values to be added to this series to be derived from the one difference value in the third column. It alone is used as the difference between each successive number starting with zero, and the values are written modulo t as before. Thus,

Write down 0

Add 2 to get 2

Add 2 to get 4

Add 2 to get 6 (modulo 5) = 1

Add 2 to get 3.

The last t = 5 values of the total series therefore are 0 2 4 1 3, which are to be added at the end of the two treatment cycles combined above. Altogether this becomes the entire first row for the complete design. The t^2 values represent the treatments in the first period for each of the t^2 subjects (columns). This complete series for t = 5 is:

0 0 1 1 2 2 3 3 4 4 0 3 2 0 4 2 1 4 3 1 0 2 4 1 3.

Completing the design. The remaining rows of the design are constructed from the first generating row by the cyclic permutation (modulo t) that was described earlier in this report. In each design there will be a total of $2t$ rows (representing periods) that are constructed by cyclically shifting each value to the next one within its own treatment cycle. In order to obtain $2t$ rows from the last treatment cycle of designs where t is odd for which only t values were supplied, the cycle is repeated an extra time. This best be illustrated by continuing with the example for t = 5.

The entire first row for t = 5 is:

0 0 1 1 2 2 3 3 4 4 0 3 2 0 4 2 1 4 3 1 0 2 4 1 3.

Each treatment cycle is underlined. Under the first 0, all of the remaining numbers in that treatment cycle are written in a column thus:

0
0
1
1
2
2
3
3
4
4

to supply the sequence of treatment values in all $2t = 10$ periods for subject number 1. Under the second 0, all of the remaining numbers in that treatment cycle are written down, and upon reaching the end at the $(2t - 1)$ period, the final treatment value is obtained by cycling back to the beginning. This procedure is continued with the remaining values in the first treatment cycle. Thus the first ten columns would be:

Subjects/Sequences

Periods							
0	0	1	1	2	. . .		4
0	1	1	2	2			0
1	1	2	2	3			0
1	2	2	3	3			1
2	2	3	3	4	. . .		1
2	3	3	4	4			2
3	3	4	4	0			2
3	4	4	0	0			3
4	4	0	0	1			3
4	0	0	1	1	. . .		4

The next ten, however, permute according to the series of values in their own treatment cycle. Thus the first three sequences within this treatment cycle (or sequences for subjects 11, 12, and 13 within the total design) would be:

0	3	2	0	4	. . .	1
3	2	0				
2	0	4				
0	4	2				
4	2	1				
2	1	4				
1	4	3				
4	3	1				
3	1	0				
1	0	3				

And as an example of how the last (third) treatment cycle would be permuted when t is odd, the first and second sequences of 2t treatments (for subjects 21 and 22 within the total design) would be:

3	0
0	2
2	4
4	1
1	3
3	0
0	2
2	4
4	1
1	3

where the five values are permuted twice. When t is even, then the last treatment cycle will have 2t treatments sequenced in it and the double permutation is unnecessary.

Patterson (1970, p. 545; 1973, p. 44) shows how these designs can be used with factorial problems and both qualitative and quantitative factors. In those circumstances as well as the cases when certain interactive effects are considered more important than others, it is necessary to select a design that least confounds the effects of greatest interest and that provides the greatest possible efficiency of estimation.

Analysis. Patterson (1973, p. 43) writes that the above designs and others that he describe are suitable when Berenblut's (1964) analysis is appropriate. They can be analyzed by fitting additive constants in a polynomial for

direct effects
first residual effects
subjects
periods
linear direct X linear residual effects.

However, he points out both in this and his 1970 paper that certain designs may be more preferable than others. This will depend upon which elements of the design are independent and which are not.

It is outside the scope of this report to discuss the analysis of the papers. As is the policy throughout this report, the reader is referred to the original document. In this particular case, analysis is only treated sparsely, and if it is to be pursued, it is likely that the aid of a competant statistician will be required. When the potential value of this class of change-over design is determined to be valid for a particular problem, the additional effort is justified.

Estimating Direct-by-Residual Interactions in Factorial Designs

Mason and Hinkelmann (1971) suggest how certain designs originally introduced by Berenblut (1967; 1968) could be extended to study the linear-by-linear interaction between direct and first residual effects. Their designs require $2(v^f)$ subjects and vf periods, where v is the number of quantitative levels per factor and f is the number of factors.

In these designs, direct linear main effects of each factor (e.g., A and B) are orthogonal to all other effects. Residual linear main effects are non-orthogonal only to subjects. The two factor direct interaction effects are also non-orthogonal to the two factor residual interactions effects. The direct-by-residual interactions -- e.g. $A_d \times B_r$ and $A_r \times B_d$ -- are non-orthogonal to subjects and to one another.

Designs. Mason and Hinkelmann (1971) supply three complete change-over designs that enable linear residual-by-linear direct interaction effects to be estimated in 2^2, 2^3, and 3^2 factorial designs. They are reproduced in full in Appendix I-C. They state that these designs had been found by trial and error. They added that although it might have been possible to discover a mathematical formulation, they did not believe it worthwhile to try and do so since they believed that with more factors and/or levels "the designs become rather large and hence less value for practical purposes" (p. 435). Of course, these design require considerably fewer subjects than do Patterson's (1973).

Analysis. The sources of variances into which these designs can be partitioned are shown in Table 8. The degrees of freedom associated with each source are shown for the two factor, two levels per factor design only. The complete design, with 32 observations, is shown in Appendix I-C. Since all three designs have the same orthogonality properties, the analyses for any of them would follow this same partitioning pattern.

As was the case for the singly-balanced change-over designs where direct and residual effects are not orthogonal, several alternative analyses are possible due to the overlap of effects. In the analysis in Table 8, sums of squares for two pairs of sources of variance can be calculated in this fashion. Either

Table 8. Partitioning Mason and Hinkelmann's 2^2 change-over design

Source				Degrees of Freedom
Subjects (unadjusted)				7
Periods				3
Direct effects				3
Factor A_d				1
Factor B_d				1
Interaction $A_d \times B_d$	Adjusted	or	Unadjusted	1
Residual effects				3
Factor A_r				1
Factor B_r				1
Interaction $A_r \times B_r$	Unadjusted		Adjusted	1
Interacting direct X residual effects				2
$A_d \times B_r$ (adjusted)				1
$A_r \times B_d$ (unadjusted)				1
or				
$A_r \times B_d$ (adjusted)				1
$A_d \times B_r$ (unadjusted)				1
Error				13
Total				31

you can leave the sum of squares for the interaction among residual effects unadjusted and adjust the sum of squares for the interaction among direct effects, or vice versa. Also either you can adjust the sum of squares for the interaction between a direct effect of factor A and the residual effect of factor B and leave unadjusted the sum of squares of the interaction between the residual effect of factor A and the direct effect of factor B, or vice versa.

The equations needed to perform this analysis are given in Mason and Hinkelmann's (1971, pp. 432-434) paper. No numerical example is given.

CHAPTER IV

DESIGNS FOR HANDLING SHORT-TERM RESIDUAL EFFECTS USING A SINGLE SUBJECT

When experimental subjects are hard to find but can be tested for extended periods of time, the designs in this section can be extremely useful. The frequency with which a particular treatment precedes or follows other treatments, or in certain cases, pairs of treatments, is balanced in such a way that direct effects of treatments can be estimated separately from the residual effects that persist for one or two periods. These designs will be referred to as serially balanced sequence designs (SBSD). They differ from the designs discussed in the previous chapter primarily by the way in which treatment-order balance is achieved. With SBSD, balance is achieved within a sequence run by a single subject rather than among several shorter runs by a group of subjects. The extended length of these sequences can allow the treatments to be presented in blocks, a feature that can be used to reduce time-related trend effects that might run through the data. Useful variations on the basic plans are possible.

TYPES AND PROPERTIES OF SERIALLY BALANCED SEQUENCE DESIGNS

The following properties of serially balanced sequence designs allow them to be classified according to their construction and the nature of the information that can be extracted from them.

Number of treatments and blocking

The "order t" of a sequence refers to the number of treatments that are to be studied in that sequence. The number of treatments, whether odd or even, and if even, whether divisible by four (modulo 4 = 0) or not (modulo 4 ≠ 0), affect the rules by which certain designs are constructed.

Serially balanced sequences have in general the potential of being blocked into segments, each containing t different treatments. Blocking can be used to partially eliminate trend effects. In certain designs, blocks can be permuted to form new sequences without destroying the serial balance of the sequence.

Persistance of Residual Effects

A majority of the designs considered in this paper have treatments arranged in sequences that balance direct effects with first residuals, i.e., those persisting from the treatment applied in the immediately preceding period. These are singly-balanced sequences (r=1). It is also possible to balance direct effects with first and second residuals, the latter being those that have carried over from the treatment applied two periods previously. These are doubly-balanced sequences (r=2). In singly-balanced designs, it is necessary that at least all ordered pairs of different treatments occur equally often in the sequence. In doubly-balanced designs, it is necessary that all ordered triplets of different treatments occur equally often in the sequence.

There exist doubly-balanced sequences that cannot be blocked and those that can be blocked. Only the latter, called "standard" sequences, will be considered in this paper. The advantages and disadvantages of singly- and doubly-balanced sequences are essentially the same as they were in the case of the basic change-over designs. The user must weigh the added size and complexity of the doubly-balance sequence against the potential importance of information concerning the second residual effects. Ordinarily these designs would be used only if second residual effects are expected to be quite large and the investigator wishes to improve the estimates of the direct effects by adjusting for both first and second residuals.

Degree of Balance Between Direct and Residual Effects and Blocks

As in the case of the basic change-over designs, each treatment may be balanced in regard to all other treatments or in regard to all treatments including itself. Type I sequences are those in which each treatment precedes every treatment including itself an equal number of times. As a consequence, direct and residual effects are orthogonal. Direct effects are also orthogonal to blocks, but residuals are not. Some residuals occur twice in some blocks and none in others.

Type II sequences are those in which treatment precedes every other treatment equally often but not itself, resulting in some degree of entanglement

in the estimates of direct and residual effects. With Type II designs, both direct and residual effects are orthogonal to blocks in completely reversible designs when all blocks end with the same treatment. In other Type II designs when blocks do not end with the same treatment, residuals are not orthogonal to direct effects or blocks.

Number of Sequences

The "index k" indicates the number of sequences in the final design. Designs discussed in this report include only one or two sequences in the final design. These limits were set by the developers of the designs "since a value of k greater than 3 or 4 will usually imply excessive replication" (Finney & Outhwaite, 1956, p. 495) and "are of little use as experimental designs, as they would require the application to one subject of an impracticably large number of successive treatments..." (Sampford, 1957, p. 288). However, for some equipment design and training research, these generalizations may not apply and an investigator may wish to consider sequences with an index k of more than 2.

Extent and Nature of Sequence Symmetry

Some sequences have the property of being "palindromic" and others, "completely reversible". Certain Type I, k = 2, sequences are palindromic, which means that with the exception of the conditioning treatment at the beginning of the sequence, the sequence reads the same forward and backward. For example, the sentence, "MADAM I'M ADAM" is palindromic. An extra feature of palindromic sequences is the independence between estimates of direct treatment effects and linear trend effects. This will be discussed more fully in the chapter on trends.

Certain Type II, k = 1, sequences are completely reversible including the conditioning treatment. This does not mean that they read the same backwards and forwards, but that when reversed they maintain their balance and other properties. In these designs, each block of treatments ends (or begins) with the same treatment which is also the same as the conditioning treatment, an extra treatment added to the beginning of the sequence. When the entire

sequence is written in reverse order and redivided into a new preliminary treatment and sets of blocks of t treatments each, the resulting new sequence is also serially balanced. Furthermore, when each block ends with the same treatment, block effects are independent of direct and residual effects.

Variations on Basic Sequence Plans

Two variations may be practical to consider. One is to expand the length of the basic sequence (i.e. $k > 2$) by adding new sequences. The other is to split the basic sequence into shorter subsets on which different subjects would be tested in order to fulfill the desired balance.

In Table 9 are listed the serially balanced sequences provided in Appendix II in completed form or in a form which a user can complete easily with the data supplied.

EXAMPLES OF SERIALLY BALANCED SEQUENCES

Examples of each major type of serially balanced sequence listed in Table 9 are shown in Figure 4. The letters in the brackets to the right of each sequence refers to the column in Table 9 exemplified by that sequence. The classification notations are consistent with those used in other tables.

At the beginning of each sequence is a single letter in parentheses; this is referred to as the "conditioning" treatment. This preliminary treatment supplies a residual for the first treatment of the first block of treatments. The parenthesis is used here to remind the user that if the sequence is repeated, the last treatment of the original sequence would serve as the conditioner for the new sequence, and this preliminary treatment -- being the same as the last -- would be deleted.

A block of treatments is the t treatments in a sequence located between a pair of semicolons. In some designs the treatments are not written out in one long sequence; instead they are piled up in levels which are still read from left to right and top to bottom as a sequence. In these designs, each level is a block and the conditioning treatment is the one located at the top of the pile.

Table 9. Characteristics of serially balanced sequences to be found in Appendix II or easily constructed from the instructions provided in the text.

Number of Treatments	Classifying Characteristics					
	SINGLY-BALANCED (r=1)					DOUBLY-BALANCED (r=2)
	TYPE I		TYPE II			
	k = 1	k = 2 ⟷	k = 1 ⇄		k = 2 ⇄	k = 1
	A	B	C			D
5		X	P		*	X
6	X	*		P	*	X
7	X	X	P		*	X
8	X	X	*	P	*	X
9	X	*	P		*	
10	X	*	X	P	*	
11	X	*	*		*	
12		*	*	*	*	
13		*	*		*	
14	X	*		*	*	
15		*	*		*	
16		*		*	*	
17		*	*		*	
18		*		*	*	
19		*	*		*	
20		*		*	*	
REFERENCES:	Sampford (1957)					Nair (1967)

Notations:
A, B, C, D Sections in Appendix II where design will be found.
⟷ Palindromic designs. ⇄ Completely reversible designs.
X Complete design provided. P Partial data supplied to finish design.
* Design easy to construct from instructions supplied in text

SEQUENCES BALANCED FOR FIRST RESIDUAL (SINGLY-BALANCED):

Type I, k = 1, t = 6, p = 37, b = 6* [A]**

(...A;) A F B E C D; D F C E A B; B A C F D E;
E B F A D C; C B D A E F; F E D B C A;

Type I, k = 2, t = 5, p = 51, b = 10 (Palindromic) [B]

(...A;) A B E C D; D E C A B; B C A D E; E A D B C;
C D B E A; A E B D C; C B D A E; E D A C B;
B A C E D; D C E B A;

Type II, k = 1, t = 5, p = 21, b = 4 [C]

(...E;) A B D C E; B C A D E; C D B A E; D A C B E;

Type II, k = 2, t = 5, p = 41, b = 8 (Reversible) [C]

(...E;) A B D C E; C A D B E; B C A D E; D C B A E;
A B C D E; D A C B E; B D A C E; C D B A E;

SEQUENCES BALANCED FOR FIRST AND SECOND RESIDUALS (DOUBLY-BALANCED):

Type II, k = 1, t = 5, p = 37, b = 6 [D]

(...EA;) B D E C A; D C E B A; E D C B A; C E D B A;
D E B C A; E B D C A; B E C D A; C B E D A;
E C B D A; B C D E A; C D B E A; D B C E A;

Figure 4. Examples of serially balanced sequences described in table 9

*No t = 5, Type I, k = 1 sequence exists.
**The letter in the brackets refers to the column in Table 9 and their location in Appendix II.

Except for the Type I, k = 1 sequence for which none exist for five treatments, the examples in Table 9 enable the reader to visually compare relative lengths of the different types of sequences.

CONSTRUCTION OF SEQUENCES WHEN RESIDUALS PERSIST FOR ONE PERIOD

Some methods of constructing serially balanced sequences will be provided. The reader should refer to the original papers if they desire more detail or the mathematical bases of these designs. Finney and Outhwaite (1955; 1956) are credited as being the original proponents of serially balanced sequences. However, the paper by Sampford (1957) provides a descripton of the construction techniques more completely and was used as the basis for the material in this section.

While each type of design shown in Table 9 will be discussed, construction details will be provided only if they are easy to explain and to use. When construction techniques are provided, enough information will be given in Appendix II to enable the user to complete the design himself. When no construction technique is provided, the completed design will be given in Appendix II. Certain serially balanced sequences differ somewhat in their mode of construction depending on whether the treatments are odd or even.

Type I, Single Sequence (k=1), r=1

Sampford (1957) states that there are no general methods of construction known for these serially balanced sequences. He supplies, however, the complete sequences indicated in Table 9, Column A. These sequences for t = 6 through 11 are reproduced in Appendix II-A. Sanford also describes (p. 297) a method of constructing Type I, k = 1 when t = 2r and r is odd, provided a sequence is already known for t = r. He constructs one for t = 14, which is also included complete in Appendix II-A. Since there are Type I, k = 1, sequences for t = 9 and 11, the reader could if desired construct sequences for t = 18 and 22 according to the method described in Sampford's article, pages 297-298. All of these Type 1, k=1, sequences for a single subject have $p=t^2+1$ periods and all only consider the effects of the first residual.

Type I, Double Sequence (k=2), r=1

How these serially balanced sequences are constructed will depend upon whether the number of treatments is odd or even. In both cases, the result is a palindromic design consisting of $2t^2 + 1$ periods with only the effects of the first residual being considered.

Odd t. When there are an odd number of treatments, the following steps should be performed to create the Type I, k = 2, r = 1 sequences listed in Column B, Table 9:

1. Write out a sequence using the same technique employed in creating the initial sequence of a Latin square change-over design when an odd number of treatments is to be studied. Thus, if t = 7, then:

   a. Put down t spaces — - - - - - - -

   b. Fill in the numbers from 1 to t in numerical order, starting with 1 and 2 in the first two consecutive spaces, skipping the next space and locating the remaining numbers in alternate spaces until no longer possible; then folding back to fill in the spaces.

   1 2 _ 3 _ 4 _
   7 6 5

   or

   1 2 7 3 6 4 5

2. Complete the square (adding t-1 more rows) using the cyclical solution. This is the same as adding +1 (mod t) to each value in the preceding row. The rows will eventually be ordered consecutively to form a single, complete sequence.

   1 2 7 3 6 4 5
   2 3 1 4 7 5 6
   3 4 2 5 1 6 7
   4 5 3 6 2 7 1
   5 6 4 7 3 1 2
   6 7 5 1 4 2 3
   7 1 6 2 5 3 4

3. Begin the sequence with any one of the t values -- an additional trial -- and then follow that by the row beginning with that treatment.

   3; 3 4 2 5 1 6 7;

4. Continue the sequence by selecting the next row which starts with the treatment ending the previous row, and continue in this way until all rows are used. The sequence will end with the same treatment that it started.

3; 3 4 2 5 1 6 7; 7 1 6
2 5 3 4; 4 5 3 6 2 7 1;
1 2 7 3 6 4 5; 5 6 4 7
3 1 2; 2 3 1 4 7 5 6;
6 7 5 1 4 2 3;

5. Complete the sequence by writing in reverse the sequence that has already been constructed, eliminating the extra beginning treatment.

The sequence progresses downward from row to row, each row being one block.

3; 3 4 2 5 1 6 7;
7 1 6 2 5 3 4;
4 5 3 6 2 7 1;
1 2 7 3 6 4 5;
5 6 4 7 3 1 2;
2 3 1 4 7 5 6;
6 7 5 1 4 2 3;
3 2 4 1 5 7 6;
6 5 7 4 1 3 2;
2 1 3 7 4 6 5;
5 4 6 3 7 2 1;
1 7 2 6 3 5 4;
4 3 5 2 6 1 7;
7 6 1 5 2 4 3

The above example provided a Type I, k=2, r=1 sequence for t = 7. In Appendix II-B, the only sequence with an odd number of treatments given completely is for t = 5. However the other odd-treatment sequences of this type are simple to construct.

Even t. When there are an even number of treatments, the following steps should be performed to create a Type I, k = 2, r = 1 sequence, listed in Column B, Table 9.

Example: t=4

1. Prepare a serially balanced Latin square change-over design with rows as sequences (subjects) columns as periods. Use a square constructed with a cyclical solution.

1 2 4 3
2 3 1 4
3 4 2 1
4 1 3 2

Square 1

2. Prepare a second serially balanced Latin square design constructed from the first in the following manner:

a. Write down the every different pair of treatments that begin and end each row (irregardless of order). There will be t/2 pairs.

(1 3)
(2 4)

b. Arbitrarily select one treatment from each pair.

3 and 4

c. Permute cyclically these selected treatments in each row of the first square to construct the second square. In this example, this would mean putting 3 in the second square where 4 was in the first square and 4 where 3 was.

1 2 3 4
2 4 1 3 Square
4 3 2 1 2
3 1 4 2

In the case of a larger square, e.g. with 8 treatments, four pairs would be obtained. One treatment is selected from each pair and all are put into numerical order. For example, if the values 1, 2, 4, and 7, had been chosen, then to permute them cyclically, 2 would always be substituted for 1, 4 for 2, 7 for 4, and 1 for 7 in each row.

3. Arbitrarily start with any row from either square, and select the next row from the other square that begins with the treatment at the end of the previous row.

2 3 1 4 (2nd row, Square 1)
4 3 2 1 (3rd row, Square 2)

4. Continue this way, alternating between squares until t rows are ordered

Rows selected from

2 3 1 4 ← Square 1
4 3 2 1 ← Square 2
1 2 4 3 ← Square 1
3 1 4 2 ← Square 2

5. Continue the sequence by selecting from the remaining rows the one that begins with the same treatment as the first row of the sequence. Complete the sequence according to steps 3 & 4.

Note that this complete design reads the same backwards and forwards.

```
2 3 1 4
4 3 2 1
1 2 4 3
3 1 4 2
2 4 1 3
3 4 2 1
1 2 3 4
4 1 3 2
```

6. Add an extra condition at the beginning of the series the treatment that is the same as the treatment at the end of the sequence.

The sequence progresses downward from row to row, each row being one block.

```
      2
2 3 1 4
4 3 2 1
1 2 4 3
3 1 4 2
2 4 1 3
3 4 2 1
1 2 3 4
4 1 3 2
```

The above example provided a Type I, k = 2, r = 1 sequence for t = 4. In Appendix II-B, the only sequence with an even number of treatments given completely is for t = 8. However, other even-treatment sequences of this type are easy to construct.

Type II, Single Sequence (k=1), r = 1

How this class of serially balanced sequences is constructed depends upon whether the number of treatments are even or odd and whether or not the even numbers can be evenly divisible by 4. All serially balanced sequences of this class with an odd number of treatments can be made completely reversible and palindromic, including the extra conditioning treatment. A special construction technique is required to develop a limited number of completely reversible serially balanced sequences when there is an even number of treatments.

Let us consider each of these various sub-types one at a time:

- odd number of treatments, completely reversible
- even number of treatments, completely reversible
- even number of treatments, not completely reversible

AD-A035 109 HUGHES AIRCRAFT CO CULVER CITY CALIF DISPLAY SYSTEMS--ETC F/G 5/10
METHODS OF HANDLING SEQUENCE EFFECTS IN HUMAN FACTORS ENGINEERI--ETC(U)
DEC 74 C W SIMON F44620-72-C-0086
UNCLASSIFIED HAC-P74-541A NL

2 OF 3
AD
A035109

Odd t, reversible. To construct a serially balanced sequence (Type II, k = 1, r = 1) when there are an odd number of treatments as listed in Column C, Table 9, the following steps should be performed:

	Example: t = 7, k = 1
1. Write down the first row of a serially balanced Latin square change-over design for (t-1) conditions.	1 2 6 3 5 4
2. Add (t - 2) new rows using a cyclical solution, i.e. by adding +1 (mod t-1) to each number in the preceding row. Each row is now a complete block of treatments to be read from left to right.	1 2 6 3 5 4 2 3 1 4 6 5 3 4 2 5 1 6 4 5 3 6 2 1 5 6 4 1 3 2 6 1 5 2 4 3
3. Place one additional t^{th} treatment at the beginning of the sequence as indicated by the + in the example.	a) 7 + b) 1 2 6 3 5 4 7 c) 2 3 1 4 6 5 7 d) 3 4 2 5 1 6 7 e) 4 5 3 6 2 1 7
Add the t^{th} treatment at the end of each row.	f) 5 6 4 1 3 2 7 g) 6 1 5 2 4 3 7

Each row except the first one, a, is a complete block containing all of the treatments. The serially balanced sequence containing all combinations of t (t-1) ordered pairs (Type II designs) can be read by starting at the top and reading cyclically through the square. Although there is a second possible location for the t^{th} condition (as shown in the designs provided in Appendix II-C), locating it at the end of each block creates a sequence with a particularly desirable property when the blocks are properly reorganized, i.e. it can be made into a completely reversible sequence.

When the number of treatments are odd in Type II, k = 1, sequences, it is possible to reorganize the blocks (when the t^{th} treatment is the last in each block) so that the sequence remains serially balanced and will read the same forward and backwards. For example, the t = 7, k = 1 square just constructed would provide the sequence:

7; 1, 2, 6, 3, 5, 4, 7; 2, 3, . . . 2, 4, 3, 7

which would <u>not</u> read the same forward and backwards. A rearrangement of the blocks is required.

The rearrangement is made as follows:

1. Write down the first row or block (a) with the single treatment.
2. Follow this with the first half of the remaining rows (blocks) arranged so that the first treatment values of each row are ordered from lowest to highest (b, c, d).
3. Complete the sequence by ordering the remaining rows (blocks) according to the first treatment value in each row in reverse order, from highest to lowest (g, f, e).

For the t=7 case, the new sequence, with the rows (blocks) now rearranged as

a, b, c, d, g, f, e

would be:

						7
1	2	6	3	5	4	7
2	3	1	4	6	5	7
3	4	2	5	1	6	7
6	1	5	2	4	3	7
5	6	4	1	3	2	7
4	5	3	6	2	1	7

It can be seen that the treatment values starting at the top and working cyclically down:

7; 1, 2, 6, 3, 5, 4, 7; 2, 3, 1 . . .

or from the bottom and working cyclically up:

7, 1, 2, 6, 3, 5, 4; 7, 2, 3, 1 . . .

are in the same sequential order.

This completely reversible, serially balanced sequence, in addition to permitting certain residual, change-over effects to be isolated from treatment

effects, also keeps the treatment effects isolated from any linear trend effects that might also be affecting performance (See Chapter V). With the t^{th} treatment at the end of the block, residual effects are orthogonal to blocks which simplifies the analysis.

Even t, reversible. With the Type II, k = 1 sequences balanced for a single residual, completely reversible serially balanced sequences can be constructed when the number of treatments are even if a Type I, k = 1 sequence exists with an odd number (t-1) treatments. In Column A of Table 9, the only Type I designs available in this report of this type with an odd number of treatments are for 7, 9 and 11 treatments; this means that the Type II, k = 1, even number of treatment sequences can be made completely reversible for 8, 10, and 12 treatments (Column C, Table 9).

The procedure is a simple one: completely reversible Type II, k = 1, even treatments (sometimes designated as t = 2r) sequences can be constructed from a Type I, k = 1, design with (t-1) treatments by adding the t^{th} treatment at the end of each block.

Even t, not reversible. Two other procedures are available for constructing Type II, k = 1, serially balanced sequences when there are an even number of treatments, although without the completely reversible property (Table 9, Column D). One of these is used when t = 4r + 4, where r is any integer; this is essentially the same as saying when t is evenly divisible for four (e.g. 8, 12, 16). The other is used when t = 4r + 2 or when t is even but cannot be divided evenly by four (e.g. 6, 10, 14). To select the appropriate design procedure, therefore, it is necessary to determine whether t = 4r + 2 or t = 4r + 4.

Although the following sequences are constructed in a rectangular (or square) form, they still represent a single sequence for a single subject that is to be read cyclically from the top row on down with each row being read from left to right.

When t is even but not evenly divisible by 4 then:

Example: t = 10

1. If $t = 4r + 2$, then calculate r, (r+1), and (2r+2).

   $r = 2; (r + 1) = 3; (2r + 2) = 6$

2. To build the first row (block) of the square:

   a. Write down the numbers 1 and 2 and follow this with a space.

      1 2 _

   b. Follow this with a 3 and another space, 4 and another space, and so on until you reach treatment number (2r + 2) which is equal to one-half of (t + 2).

      1 2 _ 3 _ 4 _ 5 _ 6

   c. Go back to the space that immediately follows (r + 1) and put a "T" in that space

      1 2 _ 3 T 4 _ 5 _ 6

   d. Starting at the right of the row, continue to the left to fill in the remaining spaces with the remaining treatment values in numberical order.

      1 2 9 3 T 4 8 5 7 6

3. Complete the remaining (t-2) rows or blocks of the sequence using a cyclical solution (building each row by adding + 1 (mod t-1) to each treatment value in the row above it. Repeat the letter T in the same column for all rows.

   1 2 9 3 T 4 8 5 7 6
   2 3 1 4 T 5 9 6 8 7
   3 4 2 5 T 6 1 7 9 8
   4 5 3 6 T 7 2 8 1 9
   5 6 4 7 T 8 3 9 2 1
   6 7 5 8 T 9 4 1 3 2
   7 8 6 9 T 1 5 2 4 3
   8 9 7 1 T 2 6 3 5 4
   9 1 8 2 T 3 7 4 6 5

4. Put a conditioning treatment at the beginning of the sequence that is the same as the last in the sequence.

   5
   1 2 9 3 10 4 8 5 7 6
   2 3 1 4 10 5 9 6 8 7
   3 4 2 5 10 6 1 7 9 8
   etc

   Substitute t for T

When t is even and can be evenly divided by 4, then:	Example: t = 8
1. If t = 4r + 4, then calculate r, (r+2), and (2r + 3)	r = 1; (r+2) = 3; (2r + 3) = 5
2. To build the first row of the square:	
a. Write down the numbers 1 and 2 and follow by a space	1 2 _
b. Follow this by 3 and another space, 4 and another space, and so on until you reach treatment number (2r +3).	1 2 _ 3 _ 4 _ 5
c. Go back to the space that immediately follows (r+2) and put a "T" in that space.	1 2 _ 3 T 4 _ 5
d. Starting at the right of the row fill the remaining spaces from right to left with the remaining treatment numbers in numerical order.	1 2 7 3 T 4 6 5
3. Put a conditioning treatment at the beginning of the sequence that is the same as the last in the sequence.	4 1 2 7 3 8 4 6 5 2 3 1 4 8 5 7 6 3 4 2 5 8 6 1 7 4 5 3 6 8 7 2 1 5 6 4 7 8 1 3 2 6 7 5 1 8 2 4 3 7 1 6 2 8 3 5 4
4. Complete the remaining (t-2) blocks of the sequence using a cyclical solution for the numbered portions and continuing to fill in a column of T's (as in Step 3 in the previous explanation). Then substitute t = T.	

Splitting a Sequence Among More than One Subject

If one sequence requires more of a single subject's time than he can afford and if there are not enough subjects to employ the basic change-over designs described in the previous sections, a sequence can be split between two or more subjects by breaking it between blocks. The test for the second subject

would begin with an extra conditioning treatment identical to the last one for the first subject. Finney and Outhwaite (1956, p. 505) in proposing this modification to serially balanced sequence designs write:

> "The conditions of balance on residuals would be maintained, and, after omitting results from the first response on each subject, the standard pattern of analysis could be adopted. It is not even necessary that the parts should be of equal length, and the extreme case of one block only (plus initial dose) per subject is permissible."

What they have identified here is the application of blocking as in an incomplete block design, where blocks equal subjects (See Simon, 1970; Cochran and Cox, 1957). As in the case of all blocked designs, treatment effects are independent of differences among blocks (or subjects) only as long as there are no interactions between treatments and blocks (subjects). The more homogeneous the subjects, the less likely that undesirable interactions will occur.

Forming Multiple Sequences

Sequences of k multiples of the basic design can be formed by writing consecutively two or more sequences that begin (and therefore end) with the same treatment, but omitting all extra conditioning treatments except at the beginning of the first sequence.

Quite obviously then, completely reversible Type II, k=2 sequences balanced for a single residual can be written for any number of treatments (Column C, Table 9) as long as a Type II, k=1, r =1 serially balanced sequence exists. The construction simply requires that the first sequence be repeated in reverse. These designs are also palindromic.

CONSTRUCTION OF SEQUENCES WHEN RESIDUALS PERSIST FOR TWO PERIODS

Nair (1967) provides sequences for a single subject that are balanced for pairs of residual effects (SBS2). These effects have carried over to the treatment effect being measured from the treatment in the period immediately preceding it (1st residual) and from the treatment two periods before (2nd residual).

He defines a SBS2 of order t (number of treatments) and index m (number of times the sequence is replicated)* as: "...a sequence involving t distinct letters such that (i) any three adjacent positions are occupied by letters which are all distinct, and (ii) each of the t(t-1)(t-2) ordered triplets of letters occurs serially exactly m times." (p. 205)

As a consequence of this definition, these SBS2 designs will have the following characteristics:

1. Length of sequence: $mt(t-1)(t-2) + 2$
2. Initial and final pairs of treatments are the same
3. The initial pair of treatments occur $m(t-2) + 1$ times and each of the other pairs occurs $m(t-2)$ times.

The standard type of SBS2 is one in which the sequence, after the initial pair of treatments, can be broken into sets of t treatments such that each set (or block) contains all of the treatments once.

Construction

Nair explains in considerable detail the theory and methods of constructing these SBS2 designs. Two of his simplest methods will be described here.

For designs where t equals any odd prime plus one (e.g. 4, 6, 8, 12), SBS2 designs can be constructed using Williams' (1950) designs as a base. However since complete sets of orthogonal Latin squares are equivalent to Williams' designs and are given in Appendix I-B, they will be used in this example. The steps given by Nair (1967, pp. 215-216) for constructing this design are as follows:

*In the discussion of the other serially balanced sequence designs, k rather than m was used to designate the number of sequences involved.

Example: t = 6

1. Write down the complete doubly-balanced change-over design for (t-1) equals an odd prime value as described in Chapter III.

   (Note: The columns in Appendix I-C are written as rows in this example.)

   Set of 5x5 orthogonal Latin squares

12345	13524	14253	15432
23451	24135	25314	21543
34512	35241	31425	32154
45123	41352	42531	43215
51234	52413	53142	54321

2. Identify each of the t(t-1) rows by R_{ij}, where i identifies the block (or Latin square) and j identifies the row in the block. There are (t-1) blocks and t rows.

   e.g.
   R_{11} = 12345 R_{12} = 23451
   R_{21} = 13524 R_{24} = 41352
   R_{32} = 25314 R_{45} = 54321

3. Put the t^{th} treatment value at the end of each row, and write out the rows in sequence: R_{11}, R_{12}, ... R_{45}.

   12345 6; 23451 6; 34512 6;
   45123 6; 51234 6; 13524 6;
   24135 6; 35241 6; ...54321 6

4. Add at the beginning of the sequence the two treatment values at the end of the sequence.

   ....16)(123456)(234516)...(543216)

The row of periods in the middle of a sequence means that some of the sequence has been omitted to save space. The row of periods before the first two numbers reminds the user that were sequences to be strung together, the first two would be omitted after the first sequence since they are present at the end of each sequence. Numbers inside parentheses represent a complete block.

Given Dudeney's (1943) "round table solutions", Nair shows how SBS2 designs can be constructed for <u>any number of treatments</u>. This is based on Dudeney's unsupported claim that "round table solutions" exist for all t's; however, he only supplies solutions for t = 4 through 12, which are given in Appendix II-E. Nair's (1967, pp. 216-217) method of constructing SBS2 from Dudeney's solutions is as follows:

	Example: t = 4
1. From Appendix II-F, obtain a Dudeney round table solution for t treatments.	1 2 3 4 1 3 4 2 1 4 2 3
2. Rearrange by cycling each of the $(t-1)(t-2)/2$ rows so that all would end in the same number. It can be selected at random.	1 2 3 4 2 1 3 4 2 3 1 4
3. Write down each of the rows in #2 in reverse order, excluding the last treatment which is retained in the last position.	3 2 1 4 3 1 2 4 1 3 2 4

4. Divide both of the above sets of rows (#2 and #3) into (t-1) groups of (t-2) rows each, with all rows in a group ending in the same pair of treatments.

Group		First treatment number Lowest	to	Highest
I	(14)	2 3 1 4		3 2 1 4
II	(24)	1 3 2 4		3 1 2 4
III	(34)	1 2 3 4		2 1 3 4

5. Identify each row as R where i refers to the group and j refers to the position within a group when ordered on the value of the first treatment number.

R(1, 1) = 2314 R(1, 2) = 3214
... R(3, 2) = 2134

6. Order the rows into the following sequence:

R(2, 1);R(1, 2);R(1, 3);...;R(1, t-2);R(1, 1);R(2, 2);...;R(2, t-2);

...;R(t-1, 1);R(t-1, 2);...;R(t-1, t-2)

and add at the beginning of the sequence the same pair of treatments as the last pair, i.e. R(t-1, t-2).

The completed serially balance sequence for t = 4 is therefore:

3 4; 1 3 2 4; 3 2 1 4; 2 3 1 4; 3 1 2 4; 1 2 3 4; 2 1 3 4.

The semi-colons identify complete blocks.

Designs

Nair supplied a number of complete serially balanced sequences for pairs of residuals that have been reproduced in full in Appendix II-D. These are for the following:

Number of Treatments:	4	5	6	7	8
Number of Observations:	26	62	122	212	338

Analyzing Serially Balanced Sequences

No detailed discussion of how to analyze these serially balanced sequence designs will be given in this paper. Listed in Table 10 however are references in which the general method and numerical examples of these analyses can be found. It should be noted that the mathematical sophistication needed to understand these analyses is considerably more than in the case of the analyses of the basic change-over designs.

Table 10. References on how to analyze serially balanced sequences

TYPE OF DESIGN	SOURCE	ANALYSIS
Singly Balanced:		
Type I Sequences	Sampford (1957) pp. 300-302	General Method
Palindromic	Sampford (1957) pp. 301-302	General Method
Type II Sequence		
Completely reversible	Sampford (1957) pp. 302-303	General Method
Other Type II	Sampford (1957) pp. 303	Comments
Doubly Balanced:		
	Nair (1967) pp. 218-221	General Method
	Nair (1967) pp. 221-223	Numerical Example

Nair (1967) has the following to say concerning the analysis of serially balanced sequence designs:

> "Despite the symmetry inherent in the sequences, no very simple analysis of variance can be formed, because all the residual effects are not orthogonal with blocks. In general, each sequence must be handled individually by setting up and solving the normal equations. The most satisfactory procedure is to isolate contrasts of importance by direct calculation as some of the useful contrasts may not be mutually orthogonal." (p. 218)

He does supply the general method (pp. 218-221) and a numerical example (pp. 221-223) of formally analyzing these sequences and estimating residual effects if the information concerning the latter is wanted. These are only applicable however for sequence generated from Williams' designs and from the round table solutions.

The analysis assumes that the effects of blocks, direct effects, first residual effects, second residual effects, and error are additive. Although the sequences are balanced against interactions between direct, and first and second residual, these effects are not included in the analysis in line with the model employed by Sampford (1957) and Williams (1950). It must be remembered that no estimate need be made of the residual effects if they are not of direct interest. In that case, they are determined only to correct the bias to the treatment effects.

CHAPTER V
DESIGNS FOR ESTIMATING TREATMENT EFFECTS IN THE PRESENCE OF TRENDS

When subjects are tested sequentially on a great many experimental conditions, it is common to find a trend or gradual shift in performance over the entire run that cannot be associated with changes in the experimental factors. If these are not isolated from the other sources of variance, estimates of the treatment effects and error may be distorted. If the interest is primarily in estimating equipment parameters, the investigator may wish to eliminate these trend effects. If the trend effects themselves are of interest, as they might be in training research, the investigator may want to obtain separate estimates of the trend effects and the equipment effects. In both situations, it is a poor experimental design that allows these two sources of variance to remain confounded.

EXAMPLES OF SITUATIONS IN WHICH TREND EFFECTS MIGHT OCCUR

In human factors research, trend effects can appear in situations such as these:

1. Inexperienced subjects are used to determine which of four display-control relationships is the most effective in the performance of a complex tracking task. As each subject is tested serially on the four configurations, his performance improves. This improvement does not appear to be due to the particular display-control configurations and it is suspected that it occurs because he is learning the general procedures needed to perform the task well.

2. As the day progresses, the temperature in the experimental laboratory rises in the morning and then lowers as evening approaches. Electronic components in the experimental simulator are affected by these temperature changes enough to alter critical parameters affecting operator performance. Changes made by the investigator in the experimental variables are made on the assumption that equipment parameters are otherwise stable. Adjusting the equipment for the unwanted trend effects is difficult to do.

In both examples, the question faced by the investigator is how to separate the effects of trends from those due to experimental variables in order to study the latter.

Techniques for Handling Trend Effects

Behavioral scientists have employed "randomization" or "counterbalancing" to determine in what order a series of experimental variables would be presented to a subject to minimize trend effects that might occur. The pros and cons on these methods were discussed in Chapter I. In spite of this concern during the design and data collection phase of an experimenter, relatively few of these same investigators bothered to analyze their data in a way that would isolate trend and experimental effects from one another. It would almost make one suspect that such a technique as counterbalancing was introduced into the experimental design more ritualistically than done as a part of a well-thought out plan.

In this chapter a number of different methods of handling trend effects will be considered. These are shown in Table 11. They differ in four critical ways:

1. Whether the presentation order must be preplanned or not before the experiment begins
2. The degree to which the effects of trends are isolated from the experimental effects
3. Whether single or multiple experimental factors can be handled
4. The complexity of the regression component of the trend under consideration

Variations on the basic descriptions in the table are sometimes possible for certain methods.

Table 11. Methods of handling trends in experimental data from sequenced treatments

Type of presentation order	Degree of orthogonalization between treatment & trend effects	Techniques involved to separate effects	Characteristics of experimental factors	Trend component under consideration ((Trend source))
Randomized order determined as experiment is run	Partial	Analysis of covariance	Single*	Linear ((Measurable concomittant variable))
Randomized order determined prior to start of experiment	Complete	Angular randomization Orthogonal polynomials	Multiple main effects only	Linear and nonlinear effects ((Trials))**
Replicating random or fixed orders in blocks	Complete	Blocking Orthogonal polynomials	Factorial	Linear and nonlinear up to B-1 ((Blocks))
Fixed order determined in advance	Complete	Counterbalanced Orthogonal polynomials	Single*	Linear and/or quadratic ((Trials))**
Fixed order determined in advance	Partial (robust)	Orthogonal polynomials	Factorial (2^{p-q})	Linear or quadratic ((Trials))**

*Although this is commonly used for a single variable, under certain conditions, multiple variables can be handled.

**The assumption is made that trials are correlated with some trend-producing factor and that trials, with equal intervals, is a measure of time.

ADJUSTING TREATMENTS FOR TREND EFFECTS WHEN NO PRESENTATION ORDER WAS PLANNED

There are circumstances when it is not possible for an experimenter to plan in advance the order in which the subject is tested on each experimental treatment. In field studies, this is particularly the case. Similarly, in field studies as well as some laboratory experiments, concommittant variables that affect performance independently of the experimental effects cannot be controlled. As a result, trends may run through an experiment and be confounded with the treatment effects.

If the experimenter cannot design the experiment to reduce the effects of trends (to be discussed later in this section) or control concommittant effects that produce the trends, then he must find ways of statistically adjusting the treatment effects for these intrusions after the experiment has been completed. At least two methods of analysis exist that he can consider for this purpose:

1. If he can identify the source of the trend and is able to measure it at the same time he is measuring operator performance, adjustments can be made to the treatment data for the trend data.
2. If he cannot identify the source of the trend, but recognizes that it is correlated with the time-flow of the experiment, he can remove the effect of trials, leaving the treatment effects ready for their own analysis. Trials can be examined for the presence of trend effects.

In this unstructured and uncontrolled situation it is unlikely that treatment and trend effects are orthogonal, i.e. completely independent of one another, and for the technique described in this section, it is not assumed that they are. But adjustments are possible, and if the treatments are presented in a random order, then the results after adjustments can be reasonably valid. As in all cases when a randomized presentation order is employed, a great deal more data must be collected to be assured that the order is truly randomized.

Various situations external to the experiment itself may become the source of a trend effect running through the experiment that serves only to mask the effects of primary interest. Typical conditions that can create unwanted trend effects in human performance field studies or laboratory experiments are:

- shifting weather conditions
- changes in illumination
- changes in temperature
- variations in information-flow rate
- drifting electronic equipment
- changes in the demands of a secondary task
- time on the job

Measurement with subsequent analysis is always an appropriate method of handling any critical variable that cannot otherwise be controlled. With sufficient preplanning, all of the examples cited above can be measured at the same time performance measures are being taken. Given the pair of measurements -- one of operator performance and one of the concommitant variable (referred to as the "covariant") -- an analysis of covariance can be used to statistically adjust treatment effects for unwanted trend effects created by these typical sources.

Analysis of Covariance

There are many good references on how to do an anlaysis of covariance. The entire September 1957 issue of Biometrics is devoted to articles by Cochran (1957) and others on this technique. Both the general method of calculation and a numerical example are given in many textbooks on experimental design such as the ones written by Myers (1972), Winer (1971), Cochran and Cox (1957), and Edwards (1950), to name only a few. A discussion of the technique in detail in this report is not justified.

An extremely simple example is given below however, to allow the reader to see how and why analysis of covariance can be used to adjust treatment effects for possible bias due to trend effects. The problem is imaginary and the numbers used were adapted from data presented by Myers (1972, pp. 338-341).

The experiment might have been something like this. There was an interest in knowing which of three devices, A, B, and C, is designed so that an individual will perform the best. Since there was no reason to suspect any condition-by-subject interaction, only one subject was tested. However, this subject was tested six times on each of the three devices. The order in which the subject was tested on the devices was randomized. That randomized sequence over 18 trials is shown in Table 12 along with the operator's performance score (Y) on each trail. It had been noticed that the equipment tended to drift, that is, a certain critical voltage that could indirectly influence operator performance tended to increase over the time the experiment was run. Therefore, a measure of this voltage was taken at the beginning of each trail and these (X) are also included in Table 12.

Table 12. Experimental data required for an analysis of covariance

TRIALS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Devices	B	C	C	A	C	A	B	A	B	B	A	B	A	C	A	C	C	B
Voltages (X)	6	6	7	7	9	10	10	11	11	11	12	12	12	13	14	15	15	18
Performance (Y) Scores	20	23	28	20	30	22	29	26	31	32	26	31	28	35	34	44	41	41

Given the data in Table 12, an analysis of covariance can be performed that will adjust the performance data (in row Y) for a linear trend effect that might be evident from the measurements in row X. To perform an analysis of covariance, analyses of variance are performed on three sets of data: the X data, the Y data, and data derived from the cross-produce between X and Y. From this, a table of sums of squares and sums of products is formed, thus:

Source of Variation	Σx^2	Σxy	Σy^2	Degrees of Freedom
Total	184.94	335.94	838.94	17
Between Devices	.77	-2.23	172.11	2
Within Devices (error)	184.17	338.17	666.83	15

From these values, an adjusted total variability is calculated, from which an adjusted error (within devices) variability is subtracted. The variance that remains is the adjusted variability associated with treatment (between devices) effects. An F-test of the ratio of the Adjusted Between Devices variance divided by the Adjusted Within Devices variance can be made.

The adjusted values are:

	Sum of squares	df	Mean Squares	F	p
Adj. Total	228.71	16			
Adj. within Devices	47.89	14	3.42		
Adj. Between Devices	180.82	2	90.41	26.43	<.001

The probability (p) of <.001 that a ratio as large as 90.41 with 2 and 14 degrees of freedom could occur purely by chance can be interpreted to mean that the performance differences among devices A, B, and C are quite reliable. Had no adjustment for trend been made, however, an analysis of variance on the Y values alone would have resulted in an F-value of 1.92 for the test of the significance of the difference among experimental conditions A, B, and C. This would have been interpreted to mean that no reliable differences in performance occurred among the three devices which, from the covariance analysis, is known to be an invalid interpretation of the data. The true differences were masked by the trend.

Treating trials as the covariant. When a linear trend is suspected to exist in data collected sequentially but the source is not identified, adjustments can be made to the performance scores with analysis of covariance by using the trial number representing "time" as a covariant. Implicit in this approach is the assumption that the trials were spaced at approximately equal time intervals and that the true covariant is highly correlated with equal increments of time.

Analysis of covaraince versus multiple regression analysis. The equivalence should be noted between adjustments made with the analysis of covariance and those in Chapter III when the direct effects were adjusted for residual carry-over

effects. In both cases, the unwanted effects are partialled out in order to test the significance of the remaining effect.

For quantitatively scaled variables, analyses of variance and covariance models provide at best limited information (Simon, 1971) (Bakan, 1966) and should be used mainly as a back-up test in the evaluation of experimental data. If the experimental variables are quantitative -- that is, with levels defined on a continuous numerical scale -- then much more information can be obtained by performing some form of regression analysis on the data (Draper and Smith, 1968). This too is mathematically related to the techniques already mentioned, but the information from the regression model will generally be considerally more useful. With the widespread use of computers for statistical analysis, multiple regression analysis is no longer a difficult task. Instead of one concomittant variable, any reasonable number can be examined. Instead of one experimental variable, any reasonable number can be included. Whereas, ordinarily the analysis of covariance assumes a linear relationship between the experimental and concomittant variables, in the regression analysis it is easy to include non-linear terms as well, thereby increasing the information from the analysis. An alternative of course would be to transform the data to a scale that minimizes the non-linear relationships. In the case of non-orthogonal variables, some variation on a straight-forward multiple regression analysis is required. This may be a step-wise regression (Draper and Smith, 1968) or better still a ridge-regression analysis (Hoerl and Kennard, 1970).

RESTRICTED PRESENTATION ORDERS FOR ISOLATING TREATMENT AND TREND EFFECTS

The two methods of isolating trends in this section both randomize to some extent the order in which a series of treatments are presented to subjects. As with the previous method, adjustment for possible trend effects are made after the data has been collected. These approaches to trend isolation differ in the following ways from the covariance technique described previously:

1. The source of the trend need not be identified nor measured during the data collection phase.

2. Some preexperimental planning regarding the presentation order is required, restricting to some extent the randomization process.

3. Within the limitations of the analysis, the designs keep estimates of trend and treatment effects independent.

4. The analyses not only determines the presence of trend effects, but also identifies and measures it regression components (i. e. whether it's a linear, quadratic, or cubic trend) as well.

In the first design, randomization of treatment order is restricted to the extent that all treatments must be tested before any treatment can be retested. Each complete replication therefore represents a block containing every treatment presented once in a randomized order. This is the classical "randomized block" design. One possible variation would be to have the treatments arranged within blocks in a systematically varied presentation order. Another variation would be to use "incomplete blocking", in which fewer than the total number of treatments are placed in each block but balanced among blocks in such a way to make the treatment effects orthogonal to block effects. This latter condition represents the primary value of any blocked design, namely, any average change in performance from block to block will not affect the comparisons among treatments within blocks. This design is suitable when studying qualitative or quantitative variables and can be used with single dimension or combined with multifactor, factorial designs.

In the second design, the random presentation order must be completely determined before the experiment is started. Through a process called "angular randomization", the presentation order is related to and help determine the levels of a single factor that will be used in the experiment. This design presupposes a quantitative factor and the ability to set as an experimental condition any factor value along a continuous scale. Some involved calculations are required to determine the values required to keep linear and non-linear trend effects independent of treatment effects.

With both designs, orthogonal polynomials are used to identify and measure the regression components of any trend that might be present, either from block to block in the first case, or from trial to trial in the second.

Isolating trend effects with blocked designs

If the exact order in which experimental conditions are to be tested cannot be controlled, some improved efficiency in design can be achieved if the treatments at least can be grouped together in blocks, allowing the randomization to exist only within blocks (Davies, 1967; Cochran and Cox, 1957; Simon, 1970, and many others). These blocks may be complete, having all treatments appearing once within each block, or incomplete, having only some of the treatments within a block but assigned in such a way that comparisons among treatments will not be confounded with differences within blocks. In these designs, treatment effects are kept independent of block effects and the differences in average performance from block to block can be used to determine the presence and nature of a trend effect running through the complete sequence.

A simple example will be used to illustrate what happens when a randomized, complete block design is used versus when an unrestricted randomized sequence is employed. Let us assume there are five treatments (A, B, C, D, and E) on which a subject is to be tested four times each. A completely randomized order of presentation for this case might be:

A B D D C A B D E C E A C E A B B D C E

If over these 20 trials, the subject continued to learn how to perform the task, treatment effects would be confounded with this trend effect. The 19 degrees of freedom could be partitioned as follows:

Sources	Dif.
Between treatments	4
Within treatments (error)	15

The error variance would be inflated and the treatments effects would be biased due to the learning trend effects that could not be isolated. Applying a covariance analysis in this case with trials as the covariant would be inadequate since there is no assurance that learning trend would be linear.

If however every treatment were run each time before another replication occurred, for example,

B D C A E / B E A C D / D A B E C / A C E B D

the comparison of performances among treatments within blocks would be more precise since they can now be measured independently of average performance differences among blocks. With this design, differences among blocks can be used to estimate the effect and the nature of any trend running through the data. This is visually illustrated in Figure 5.

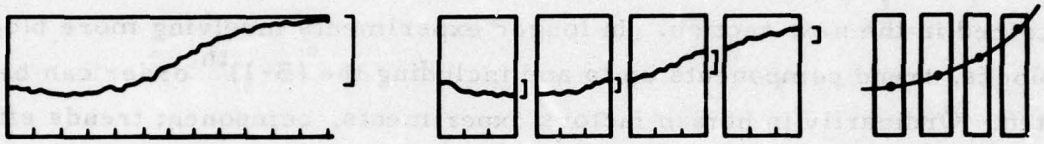


Figure 5. The advantage of blocking in the presence of trend.

In the figure it can be seen that although a major trend is running through all of the data, when the data is blocked, the amount of variations within each block associated with the trend is reduced. Treatments would only be compared with treatments within the same block and their effects would be average over blocks. The assumption in a blocked design is that there are no interactions between treatments and blocks.

The analysis of this randomized block design for five treatments replicated four times with all five treatments in each of four blocks would be:

Between treatments	4
Between blocks	3
Remainder (error	12

While the error term in this example has lost some degrees of freedom, the variance also has been considerably reduced since block differences have been

removed. If upon testing these block differences, no significant effect was discovered then the degrees of freedom associated with blocks could be recovered by recombining block and remainder sources of variance.

To estimate the nature of any trend effect in the data the three degrees of freedom and variance associated with blocks can be partitioned as follows:

Blocks	3
Linear component	1
Quadratic component	1
Cubic component	1

Just how these component effects are isolated using orthogonal polynomials will be described in the next section. In longer experiments involving more blocks, for B blocks, trend components up to and including the $(B-1)^{th}$ order can be estimated. Ordinarily in human factors experiments, component trends effects greater than cubic are unlikely to be important and could probably be combined with the error term.

Within the blocks in this example, the order is randomized. Blocking in this fashion therefore only eliminates a part of the trend, for it is still present (in reduced strength) within the blocks. Obviously, therefore, the steeper the trend and the more conditions within a block, the less effective randomized blocking will be in eliminating trend from treatment effects within blocks.

When all treatments are tested with equal frequency within blocks and are conditions of a factorial design, all main and interaction effects among factors can all be estimated independently of block effects. If incomplete blocks are used (in which fewer than the entire set of treatments are within a block), the within block variations may be reduced; however, this advantage may be paid for by the inability to estimate certain interactions within blocks.

Even when systematic designs are employed, if the treatments are replicated a number of times and it is possible to associate each replication of all treatments with a different block, the trend effect reflected by the differences among blocks can be isolated in the same way it was in a randomized block design.

If from the analysis of variance, block effects are found to be significant, an examination of the mean performance scores per block (as in Figure 5) would suggest the presence of a trend (learning?). The act of removing the block effects in the analysis leaves the treatment effects uncontaminated from differences among blocks (but not within blocks). However, when there is an interest in the trend effect itself, as in a training experiment, the characteristics of the trend should be determined by some more precise method than subjective inspection. This can be accomplished with the aid of orthogonal polynomials.

Orthogonal Polynomials

If polynomials of the form $Y = a_1 + b_1X + c_1X^2$ and of the form $Y = a_1 + b_1X + c_1X^2 + d_1X^3$ are both fitted to the same data, the values of the coefficients, a_1, b_1, and c_1, for the two equations will not be the same. This is to say that after fitting an equation of the second order, the procedure of fitting an equation of the third order is not merely a matter of adding a term in X^3 and coefficient to the original equation. It is not possible to fit in successive stages, testing the goodness of fit as we proceed, unless simultaneous equations are solved at each stage. This means that a new regression equation must be derived for each increase in the order of the equation. From this, the proportion of variance accounted for by the regression of each equation would be calculated and the differences between steps would represent the proportion accounted for as each additional higher-order non-linear component is added. This can be translated into sums of squares that can be used to test the signficance of the component. While the procedure need not be costly with computers and appropriate programs so readily available today, the same kind of trend analysis can be done using ready made coefficients of orthogonal polynomials and a manually operated calculator. Sums of squares of each trend component can be calculated directly and tested in a conventional analysis of variance. Additional steps are required to write an equation if it were desired.

Two assumptions must be met when this method of curve fitting is employed: 1) that there are equal intervals of time btween trials; 2) that each treatment has been measured the same number of times. It should also be noted that the sum of squares of each component merely represents the additional amount due to

the particular component in excess of any contribution of lower order components. Let us now look at some actual examples and work through a trend analysis to show how they can be used.

First, what are coefficients of orthogonal polynomials? Given a series of mean response values, coefficients of orthogonal polynomials are used to weight them in a way that will indicate the presence of only one possible order of a trend relationship. Coefficients of orthogonal polynomials for 8 and for 16 measurements are reproduced in Table 13. These coefficients are independent of one another (i.e. the sum of their cross-products equal zero). They are standardized so as to be symmetrical about zero and to have the sum of the coefficients equal zero. For any number of response values, N, there are N different sets of coefficients with N coefficients each. With these the sum of squares for the mean and up to the N-1th order regression component of n responses can be estimated.

Table 13. Examples of tables of coefficients of orthogonal polynomials from DeLury (1950)

	ξ'_0	ξ'_1	ξ'_2	ξ'_3	ξ'_4	ξ'_5	ξ'_6	ξ'_7
n = 8	+1	-7	+7	+7	+7	-7	+1	-1
	+1	-5	+1	+5	-13	+23	-5	+7
	+1	-3	-3	+7	-3	-17	+9	-21
	+1	-1	-5	+3	+9	-15	-5	+35
	+1	+1	-5	-3	+9	+15	-5	-35
	+1	+3	-3	-7	-3	+17	+9	+21
	+1	+5	+1	-5	-13	-23	-5	-7
	+1	+7	+7	+7	+7	+7	+1	+1
S.S.	8	168	168	264	616	2184	264	3432
I	+7	0	-8.16667	0	-14.83611	0	-16.34306	0
J	8	0	0.66667	0	7.33611	0	29.26716	0

	ξ'_0	ξ'_1	ξ'_2	ξ'_3	ξ'_4	ξ'_5	ξ'_6	ξ'_7
n = 16	+1	+1	-21	-63	+189	+45	-75	-175
	+1	+3	-19	-179	+129	+115	-25	-375
	+1	+5	-15	-265	+23	+131	+45	-235
	+1	+7	-9	-301	-101	+77	+87	+157
	+1	+9	-1	-267	-201	-33	+59	+423
	+1	+11	+9	-143	-221	-143	-39	+143
	+1	+13	+21	+91	-91	-143	-117	-533
	+1	+15	+35	+455	+273	+143	+65	+195
S.S.	16	1360	5712	1007760	470288	201552	77520	1545232
I	+15	0	-37.5	0	-358.125	0	-137.27679	0
J	16	0	1.33333	0	57.33889	0	70.87282	0

*Coefficients up to ξ'_{16} are also available (De Lury, 1950)

The two examples shown in Table 13 are for $\underline{n}$ = 8 and $\underline{n}$ = 16, which may be particularly useful sets to have if trend analyses are applied to screening experiments (see Simon, 1973). Each ξ column supplies the coefficients needed to estimate each unique trend component, with ξ_i designating which i^{th} component it is. The first ξ_1 is for the linear (or first order component), the second for the quadratic, and so forth. The S. S. row in the table represents the sum of the values squared for each column. In some published tables, as illustrated here with the n = 16 table, only one-half of the coefficients are given. The other half is obtained by folding over the given half, maintaining the same sign for even-ordered coefficients (in ξ_2, ξ_4, etc.) and using the opposite signs for odd-ordered coefficients (in ξ_1, ξ_3, etc.).

Second, where do we find the coefficients of orthogonal polynomials? They are in prepared tables in the format shown in Table 13. While certain textbooks (e.g., Hayes, 1963; Davies, 1967; Beyer, 1966;) have tables that supply some coefficients, more complete sets can be found in Fisher and Yates (1970). They supply tables of coefficients useful for isolating up to 5th order regression component effects for 75 or fewer response values. Anderson and Houseman (1942) extend the Fisher and Yates' tables to handle up to and including 104 observations, and DeLury (1950) provides the coefficients needed for all regression components up to the $(N - 1)^{th}$ order for up to and including $\underline{n}$ = 26 response values. Beyer (1966) reproduces Fisher & Yates' table up to and including the polynomials for $\underline{n}$ = 50.

Third, how do we use these polynomials to analyze trend component effects? Let us imagine that we have the means from six trials, e.g. M_1, M_2, ... M_6, (or blocks, for that matter) that represent an ordered sequence. From a table, we would get the linear coefficients (ξ_1) of a polynomial for n = 6, which are -5, -3, -1, +1, +3, +5. Each coefficient would be multiplied by the corresponding trial mean in the sequence, thus:

$$(M_1)(-5) + (M_2)(-3) + (M_3)(-1) + (M_4)(+1) + (M_5)(+3) + (M_6)(+5)$$

and the sum of these would be the sum of squares representing the linear effect. Similar steps would be performed with sets of coefficients for the non-linear components. In Table 14, three numerical examples are worked out to illustrate how these trend component sums of squares would be calculated.

In Table 14, the mean performance of six trials are given at the upper left (a) for three imaginary studies, designated A, B, and C. These means are ordered in a sequence from top to bottom in accordance with the trial numbers.* The results of these three imaginary studies are also shown graphically in Figure 6.

From a subjective examination of these figures, it would appear that there is no performance trend in Study A, a pronounced linear relationship in Study B, and possibly a linear relationship in Study C, although with so few data points and such a spread in the performance scores, it is difficult to say. A component analysis of the regression using orthogonal polynomials could quantify to what extent the function contains a linear, quadratic, cubic, and up to a fifth order component. The operations employed in this analyses can be followed using the information supplied in Table 14.

In the upper right hand corner of this table (b), parallel to the corresponding means of each trial (for all three studies), are the coefficients of orthogonal polynomials for n = 6, up to and including fifth order components. These were obtained from Fisher and Yates' (1963) tables. The sum of the coefficients squared are given in the row (c) below each vertical set of coefficients.

To obtain the Sum of the Inner Products (d), each mean is multiplied by the coefficient in its corresponding trial position. Thus, to determine the sum of the inner product for the linear component of the means of study A, each mean is multiplied by the corresponding coefficient in column LIN. and these cross-products are added, thus:

$$4(-5) + 2(-3) + 6(-1) + 1(+1) + 3(+3) + 5(+5) = 3$$

*It does not matter whether we are referring to the means of trials or blocks, or for that matter, experimental levels of a single factor. Exactly the same procedure is used to isolate and estimate the effects of the regression components of a trend running through the sequential data.

Table 14. Numerical calculations involving trend analysis with orthogonal polynomials.

(a) TRIAL	MEAN PERFORMANCE BY TRIAL FOR STUDIES A	B	C	(b) (n = 6) ORTHOGONAL POLYNOMIAL COEFFICIENTS LIN.	QUAD.	CUB.	QUART.	QUINT.
1	4	1	4	-5	+5	-5	+1	-1
2	2	3	1	-3	-1	+7	-3	+5
3	6	2	2	-1	-4	+4	+2	-10
4	1	4	6	+1	-4	-4	+2	+10
5	3	6	3	+3	-1	-7	-3	-5
6	5	5	5	+5	+5	+5	+1	+1
(c) SUM OF THE COEFFICIENTS SQUARED =				70	84	180	28	252
(d) SUM OF THE INNER PRODUCTS FOR THE THREE STUDIES			A	3	12	18	8	54
			B	31	-3	-9	-9	9
			C	15	9	-25	13	31
(e) SUM OF SQUARES (1 d.f. each)								
Total across each study equals 17.5			A	.1	1.7	1.8	2.3	11.6
			B	13.7	.1	.4	2.9	.3
			C	3.2	1.0	3.4	6.0	3.8
				LIN.	QUAD.	CUB.	QUART.	QUINT.

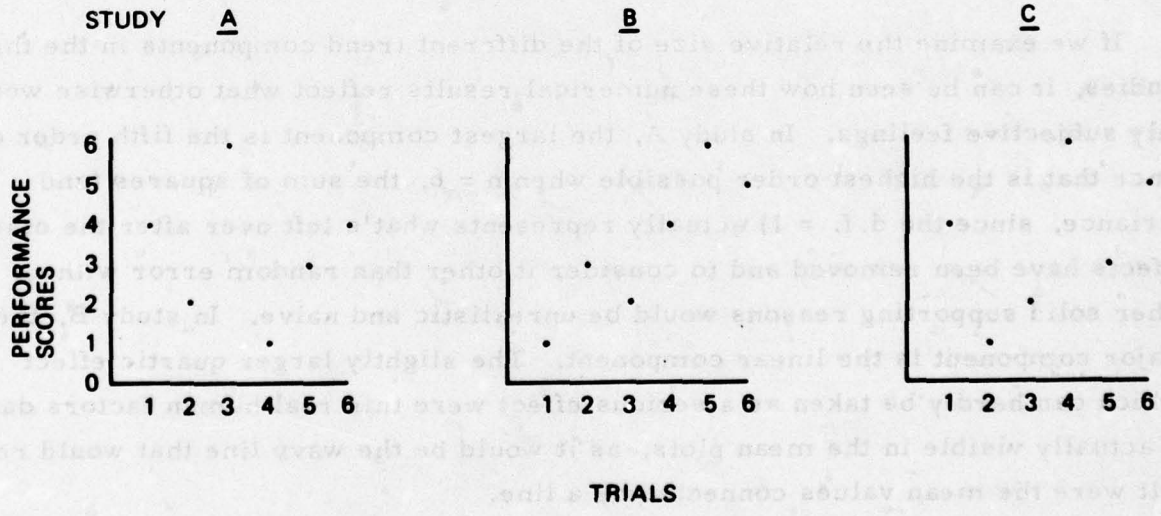


Figure 6. Performance plots of the results from three imaginary studies in Table 14

To obtain the SUM OF SQUARES at the bottom of the table (e) with one degree of freedom for each component, the SUM OF THE INNER PRODUCTS is squared and then divided by the corresponding SUM OF THE COEFFICIENTS SQUARED. In the same study (Study A, linear component), the sum of squares would equal:

$$\frac{\text{SUM OF INNER PRODUCTS, QUANTITY SQUARED}}{\text{SUM OF COEFFICIENTS SQUARED}} = \frac{3^2}{70} = .13 \text{ (or .1)}$$

It can be seen that the total sum of the sum of squares across linear through quintic components of each study (A, B, and C) all equal 17.5 (within the accuracy of rounding). Since in each study, each regression component sum of squares is independent of the other, they can be used to:

1. estimate the proportion of total performance variability attributable to each trend component; and
2. make F-tests of the significance of the trend components.

With only the six values (and a total of five degrees of freedom), effects greater than the cubic component can probably be attributed to random error. Particularly in human factors experiments, higher-than cubic-order effects are probably negligible (see Simon, 1973; 1974). Thus, the sum of squares for the fourth and fifth order trend components could be pooled and considered to be (with 2 degrees of freedom) an estimate of the error variance needed for a not too powerful F-test.

If we examine the relative size of the different trend components in the three studies, it can be seen how these numerical results reflect what otherwise were only subjective feelings. In study A, the largest component is the fifth order one. Since that is the highest order possible when n = 6, the sum of squares (and variance, since the d.f. = 1) actually represents what's left over after the other effects have been removed and to consider it other than random error without other solid supporting reasons would be unrealistic and naive. In study B, the major component is the linear component. The slightly larger quartic effect (which can hardly be taken as a serious effect were this real human factors data) is actually visible in the mean plots, as it would be the wavy line that would result were the mean values connected by a line.

In study C, the one that was most difficult to interpret subjectively, it can be seen that the linear component accounts for relatively little of the observed variability and that if one discounts the validity (and reliability) of the higher order effects, it must be assumed either random error is high or some other unknown factor is operating.

While not specific to the discussion on trends (or sequence effects in general), it should be noted that although the effects due to trials are orthogonal to those due to subjects and to treatments when Latin square types of balanced designs are used, trial effects will be obtained only if there are no interactions among subjects, treatments, and trials. When data is variable and cannot be explained with one of these designs, it is not unreasonable to assume that these interactions effects are not negligible. In the case of the means representing the trials, in fact, they are a composite effect of the effects due to subject-by-treatment interactions and trials, and will represent trials alone only when the SxT interaction is negligible.

Orthogonal polynomials were used here to estimate trend effects in a series of trial or block means. As shall be shown later, they can also be used to determine the sequence for experimental conditions that is completely or reasonably orthogonal to some trend components. They can also be used to evaluate the robustness of designs to trend effects.

Isolating Trend Effects with "Angular Randomization" Designs

When different levels of a quantitative variable are presented sequentially to the same subject in a randomized order, the possibility of a trend effect entering the data is considerable. While orthogonal polynomials can be used to examine the trend of any set of data taken over many trials, unless the order has been truly randomized (which generally require many measures to be adequate), the degree of independence between treatment and trend effects may not be sufficient. G. E. P. Box (1952) proposed a technique called "angular randomization" for selecting the presentation order and the factor levels to orthogonalize these effects in such an experiment. This technique will insure that the experimental effects are orthogonal to third-order time trends, that the polynomial coefficients

are as reliable as possible, and that the error term satisfies the requirements for randomization.

With any set of N trial means, it is mathematically possible to fit the data exactly with a polynomial of order (N - 1). Box however proposes that if the trend can be approximated by a third-order polynomial, then any remaining degrees of freedom could be used to estimate the coefficients of a first order model (main effects) of the experimental factors and what remains could be used to estimate the error variance.

Box and Hunter (1958, pp. 155-162) describe very clearly how angular randomization is used to construct an experimental design. It would be inappropirate in this report to merely reproduce their example; the interested reader should refer to the original paper. However an explanation of the steps that they go through will be given here to help the reader with a limited mathematical background to understand and follow their example.

To illustrate angular randomization, they posed the following problem:

> "Let us see how an expirement requiring $N = 12$ runs is planned in order to measure the first order effects of $k = 2$ continuous independent variables x_1 and x_2 upon some response y, using the coil-heated reactor mentioned above. Our object then is to construct an optimum first order design having random orientation, yet still providing estimates that are orthogonal to <u>any</u> time trend which can be represented by a polynomial of third order. Since four of the original 12 degrees of freedom will be used up in estimating a third order trend, we begin by constructing a first order design from among the eight degrees of freedom that will remain after the cubic trend is eliminated." (p. 158)

To determine the factor levels of the experimental conditions the following steps are to be used. They will best be understood while reading the steps of the original article.

Step 1. Obtain a table of random normal deviates. These can be found in Wold (1948), Dixon and Massey (1957), and Beyer (1966) to name a few sources. Random normal deviates are merely tables of random numbers with a mean of zero and a standard deviation of one.

Step 2. For each factor, select from the table of random normal deviates the same number of values* as there are degrees of freedom left after three have been used to estimate the components of a third order trend. (In the published example, in a 12 "run" design, i.e. one with 12 observations, there are a total of 11 degrees of freedom. When three are used to estimate the linear, quadratic and cubic trend effects, 8 are left to estimate the main effects and error. Therefore for each factor, eight values are selected from the table of random normal deviates. These are shown on page 159 in Box and Hunter's (1958) article as columns Z_1 and Z_2.)

Step 3. Although step 2 provides sets of numbers (vector), one for each factor, that are randomly oriented, they are still not orthogonal to one another. They are made so by partialling out from the values of each succeeding set of values the effects of the preceeding set.

Step 4. The vectors for each factor are now standardized by dividing each element in a vector by the sum of squares of the elements comprising the vector. (In the example on page 160 of Box and Hunter's (1958) article the results of steps 3 and 4 are shown in Columns U_1 and U_{2-1}.)

Step 5. To guarantee that the estimates provided by the design are orthogonal to any cubic trend, the actual factor levels to be used in the experiment are constructed from a linear combination of the orthogonal polynomials for N equal to the total number of observations in the design (in this example, twelve), starting

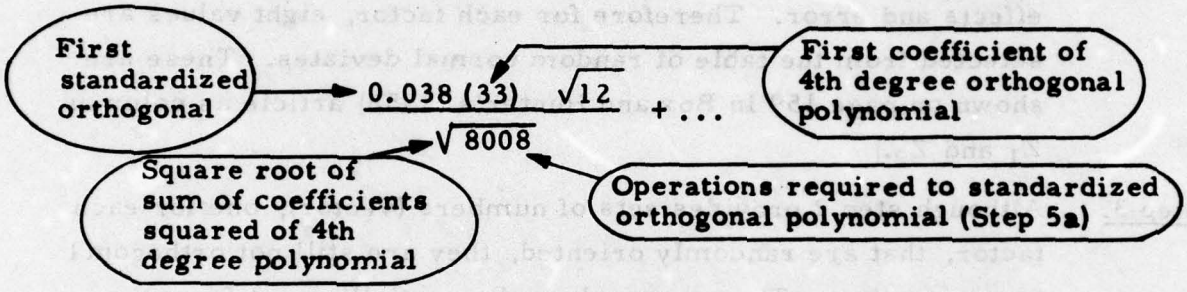
*A set of values for each factor is referred to as a "vector".

with the 4th order polynomial (i. e. one more than the degree of trend being isolated, in this example up to and including the cubic trend) and continuing to the (N-1)th.

a) The orthogonal polynomials are standardized so that the sum of the coefficients squared equal N = 12.

b) The orthogonal polynomials from the 4th to the 11th (in this example) are linearly combined using the elements of the standardized orthogonal vectors as coefficients.

(In the example, this process is illustrated quite clearly. However, to help the reader identify the numbers on page 161 of the Box and Hunter (1958) article note that



(In the example, the results of Step 5 are shown on page 161 of Box and Hunter's (1958) article as columns x_1 and x_2, which represent the actual factor levels used during each observation).

Thus on the first trial (e. g. Time order 1), factor 1 was set at 0. 4 and factor 2 was set at -0. 5. The performance, y, obtained on that trial was 100. 3.

From this data, following the method of analyzing time trend data using orthogonal polynomials described earlier, it is possible to:

1. Write a polynomial that fits the linear, quadratic, and cubic trend components in the data;
2. Isolate the trend data, the main effects of the factors, and an estimate of error, all orthogonal to one another.

Thus we have been able to evaluate multiple main effects in the presence of a sizeable trend effect. When this trend effect is artificially created by experimental conditions, it must be isolated. When tests of significance are made, a failure to isolate the trend effect from the estimate of error can seriously jeopardize the validity of such hypothesis testing. For many human factors problems, however this technique of angular randomization has a number of limitations. For example:

1. Although many of the factors studied in human factors experiments are quantitative (i. e. capable of being measured on an ordered, continuous scale), it is not always convenient to select from any part of the entire scale as the technique demands. Quite often, equipment constraints enable only a few fixed values out of the entire scale. Hill (1960) deals indirectly with this problem when he shows how Box's randomally oriented designs can be combined with Cox's (1951; 1952) systematic designs for quantitative and qualitative variables.

2. Although the effects of multiple factors can be studied, trend free, with designs developed with this technique, the assumption is made that interaction effects are negligible. In these "first order" designs, only main effects can be estimated. This may prove untenable in many human factors problems.*

3. The calculations required to determine which factor levels are to be used depend upon having the sequence (albeit randomized) completely determined first, limiting a primary value of randomized sequences -- their sponteneity.

ORDERING TREATMENTS IN SEQUENCES ROBUST AGAINST TREND EFFECTS

If the experimenter can control the order in which experimental conditions are to be presented to a subject (who is to be tested on a series of conditions),

*However with quantitative variables, extrinsic interactions may be eliminated by selecting the proper scale of measurement for each factor.

he can select a sequence that will enable trend and treatment effects to be essentially unconfounded. Certain trends can run through the data without materially affecting estimates of treatment effects. Both trend and treatment effects can be estimated in the same experiment.

The techniques that are available differ in the degree of orthogonality (independence) that can be obtained between treatment and trend effects. They also differ in regard to the number of experimental factors (i. e. one or more than one) and the order of the trend (i. e. linear, quadratic, or higher) that can be isolated. The designs in this section are of two major types, i. e.:

1. One way sequences that are balanced for linear and/or quadratic trends. (Phillips, 1968)
2. 2^{t-q} factorial plans sequenced in a way that is robust* against linear and quadratic trends. (Daniels & Wilcoxon, 1966)

One-Way Sequences that are Balanced for Linear Trend

In order to create a sequence that is balanced for a trend effect, each condition must be tested more than once. In the designs considered here, it is always assumed that the number of replications, m, is the same for all conditions. Phillips (1968) also specifies other conditions that must exist for a sequence to be balanced against a linear trend. These are:

1. The number of replications, m, must be even, or the number of replications, m, and the number of treatments, t, both must be odd.
2. Each sum of the integers representing the trial position associated with each experimental condition must equal

$$\frac{m\,(mt + 1)}{2}\ .$$

*The term "robust" is used here as it was in the original paper. It means that with these sequences any estimates of treatment effects will be negligibly distorted by the presence of selected trends.

For example, here are three treatments, A, B, and C, replicated twice and put into a six trial sequence:

Trial number	1	2	3	4	5	6
Exper. Cond.	A	B	C	C	B	A

This sequence is orthogonal to any linear trend effect that might be running through the six observations. This is the case since:

$$A = 1 + 6; B = 2 + 5; C = 3 + 4$$

which are all equal to each other and to:

$$\frac{2(2 \times 3 + 1)}{2} = 7$$

Sequences balanced for linear trends are constructed differently, depending on whether the number of replications, m, equal 2, 3, or more than 3. The construction procedures supplied by Phillips (1968) are explained below.

Construction when m = 2. When only two replications are to be used, the simplest way to balance a sequence for linear trend is to write the treatments -- even or odd number -- once in some random order and then to repeat that order a second time in reverse.

Construction when m = 3. When there are to be three replications, according to the condition cited earlier, the number of treatment must be odd. Preparation of a sequence balanced for linear trend in this case is slightly more complicated. Phillips (1968, p. 163) supplies the following method, which shall be illustrated here for the case of m = 3 and t = 5. The following steps are to be followed to construct the sequence when m is odd:

1. Write down the treatments in a random order.

C A E D B

2. Write down in consecutive order the trial numbers beneath each treatment so that each has a number associated with it once. That's replication one.

1 2 3 4 5 } Repl. 1

3. For the second replication, write down the next set of trial numbers consecutively, beginning with the condition just to the right of center. Write clockwise and upon reaching the end, cycle again from the beginning.

6 7
8 9 10 } Repl. 2

4. Start the third replication at the middle condition, and write down the next set of trial numbers cyclically, skipping every other condition until all receive a third replication number.

11 _ 12
- 13 _ 14 _
15 } Repl. 3

If the three replicates of the five treatments are now written out serially ordered according to the trial position values assigned to each, the sequence would be:

Trial No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Treatment	C	A	E	D	B	D	B	C	A	E	E	B	A	D	C

According to the rule, for a sequence to be orthogonal to a linear trend, the sum of the trial integers for each experimental condition must all be equal (in this example) to:

$$\frac{3\,(3\times5 + 1)}{2} = 24\,.$$

A check of the numbers in the example shows that this is the case.

Construction when $m > 3$. With more than three replications per treatment, then the number would have to be divided randomally into sets of two if m is even, or sets of two and three, if m is odd. Once this is done, then the appropriate

technique for constructing sequences for two or three replications would be used to balance each subset and then combined to balance the treatments in the complete sequence for linear trend.

One Way Sequences Balanced for Linear and Quadratic Trends

Additional observations are needed to balance a sequence for quadratic as well as linear trends. The conditions necessary for this balance are:

1. The number of replications, m, must be equal to two times the number of experimental treatments, t.
2. Both m and t must be divisible by 3 or both must be prime to 3. Prime to three means that neither can be divided evenly by 3.
3. The sum of the trial numbers squared for each experimental condition must equal

$$\frac{m\,(mt + 1)\,(2\,mt + 1)}{6}$$

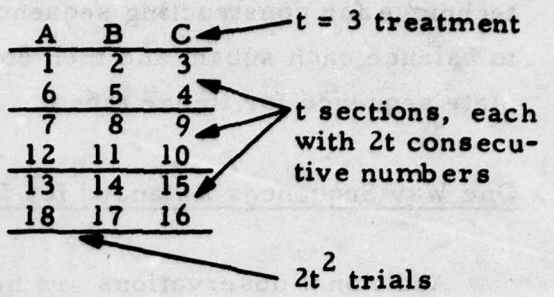
4. The sum of the trial numbers for each experimental condition must equal

$$\frac{m\,(mt + 1)}{2}$$

To satisfy these conditions, Phillips (1968) provides a method of constructing these sequences.

Construction. To construct a sequence balanced for both linear and quadratic trend requires $2t^2$ observations. To illustrate how this is done, a balanced sequence will be constructed for three treatments, A, B, and C, replicated six times. The sequence will be 18 trials long. With m = 6, and t = 3, the first two of the above conditions are satisfield. The following steps must be performed to achieve conditions 3 and 4:

1. Divide the total number of $2t^2$ trials into t sections each containing $2t$ consecutive numbers. The numbers within each section must be paired so that the sum of every pair within (but not between sections) is equal. This can be accomplished by numbering from left to right and then folding back as shown in the example.



2. Next the order of the pairs are to be rearranged by *cyclically shifting the* pairs within each section, one pair at a time after *the first*

A	B	C	
1	2	3	1st section
6	5	4	*remains fixed*
8	9	7	2nd section
11	10	12	cycles one to left
15	13	14	3rd section
16	18	17	cycles two to left

The sequence thus obtained is:

Trial Nos.:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Treatments:	A	B	C	C	B	A	C	A	B	B	A	C	B	C	A	A	C	B

with treatments effects being orthogonal to both linear and quadratic trends. This can be proven by determining if the sums of the trial numbers and the sums of the trial numbers squared are each equal for the three treatments, A, B and C. They are, with the sums of each being 57 and the sums of the integers squared each being 703.

Analysis. The analysis is rather straight forward when treatment and trend effects are orthogonal. Thus if eight conditions had been balanced for a linear trend, the 16 observations would provide 15 degrees of freedom that could be partitioned:

Linear trend component	1
Treatment effects	7
Residual	7

If the treatment conditions were in fact a factorial, then factor and interaction effects could be partitioned from the 7 degrees of freedom. Each trend component represents one degree of freedom.

In selecting a design, an implicit assumption is usually made that some higher-order trend components will not be present in the data. In practice, for the data on treatments to be unbiased, the assumption must be valid. The question can arise as to whether any test of the presence of a higher trend component should be made and what to do about it after the data has been collected. As to the latter, it is the experimenter's responsibility not to take the easy way out before the experimental data has been collected; if he has doubts about the presence of a significantly large higher-order trend component, he should not use a design that ignores that possibility. Still the most sincere intentions can go astray. Rather than merely accept a prior assumption, the investigator should at least inspect his data for higher-order trend components. To aid a visual inspection, he can even remove from mean values the effects of those trend components that are orthogonal to the remainder of the data in the design. If a higher-order trend can be observed that's much larger than the treatment effects, even though it had not been planned for, the experimenter may wish to partial out (orthogonal polynomials) the effect before analyzing the treatments. On the other hand, if the higher-order trend, although visible, appears small relative to the treatment effects, it might be ignored. In either case, some bias will occur. Inspection however at least enables the experimenter to make a better interpretation of his data. This approach of questioning the data should be used whenever assumptions have been made in selecting a design.

Factorial Designs from One-Way Designs Balanced for Trend Effects

Phillips (1968) points out that factorial designs may be constructed as special cases of one-way designs. This is done merely by treating each experimental condition, i.e. each combination of factor levels, as if it were a level of a single variable and have the set of conditions obey the same requirements set forth for the one-way case. Thus to balance a 2 x 3 factorial design for linear trend, the six experimental conditions are ordered according to the rules for a one-way balance for linear trend: first, that there must be an equal number of

observations per cell, and second, that either this number of observations per cell must be even, or both it and the total number of cells must be odd. Thus:

Conditions:	A1B1	A2B2	A1B3	A2B1	A1B2	A3B3
Periods:	4	2	1	6	3	5
	9	11	12	7	10	8

Note that the experimental conditions are assigned at random to the trial periods for the first replication in each cell, and that each second observation is fixed to meet the requirement that the sum of the periods should be equal.

When "screening" designs are employed (Simon, 1973), a block of 8 can be made robust to linear and quadratic trends with 128 observations, and to linear only, with 16 observations. A block of sixteen can be made robust to linear and quadratic trends with 512 observations, and to linear only, with 32 observations. Here again the experimenter is faced with the trade-off between higher costs and better control. The factorial 2^{t-q} described in the next section may be a better alternative because they require fewer observations.

Phillips (1964; 1968) also shows how "magic squares" and "magic rectangles" can be used as two-way designs to balance linear and sometimes quadratic trend effects for experimented main effects and some lower-order interactions. A magic square is an $\underline{n} \times \underline{n}$ array of all the integers from 1 to n^2 such that the sum of the integers in every row, in every column, and in the two principal diagonals is the same, i.e. equal to $n(n^2+1)/2$. With magic rectangles, the sum of integers in all rows are equal and the sum of integers in all columns are equal. Sources and uses of these devices are explained in the two articles. More than one factor may be handled by either of the two dimensions of the figures in the same way that several factors are handled in the one-way designs. This permits them to be used for multifactor experiments. At this time, the advantages of using these two-way designs for human factors experiments over one-way designs is not immediately apparent.

Factorial 2^{t-q} Plans Robust Against Linear and Quadratic Trends

Daniel and Wilcoxon (1966) provide plans for sequencing the experimental conditions of two-level factorial and fractional factorial designs so that the treatment effects are robust against trend effects. These sequences differ in two important ways from those proposed by Phillips in the previous section:

1. These designs will require fewer observations.
2. The independence between treatment and trends in these designs will not be as complete.

Sequences for 2^{t-q} designs are of particular interest since they can be used for screening large numbers of factors economically (Simon, 1973).

Construction. A preferred sequence is determined empirically by examining the sequences of signs in a conventionally ordered sign matrix* that are prepared for 2^t designs and finding the ones with orders that are least correlated with trend effects. As Daniel and Wilcoxon (1966, p. 261) explain: "...certain of the ordered contrasts appearing in the 2^p systems are orthogonal to linear and to quadratic trends. Some other contrasts are nearly orthogonal and some are rather heavily correlated with first and second order trend. The design problem is, then, to choose those sets of ordered contrasts that provide efficient estimation of all desired effects and interactions." (p. 261) The emphasis here is on the word "efficient", which emphasizes the fact that some trade-off between the total number of observations and degree of orthogonalization is acceptable in practice. For example, with Phillips' one-way designs that are completely orthogonal to linear and quadratic trends, 512 observations are needed to study the 16 treatments of a 2^4 design. With the designs of Daniel and Wilcoxon, a

*Those readers who are unfamiliar with the construction and use of sign matrices of 2^{t-q} designs should refer to such references as Davies (1967, p. 262), Cochran and Cox (1957, p. 154), Simon (1973, p. 64), and almost any text that discusses the design and analysis of two-level factorial and fractional factorial designs. The signs of the tables are the "contrasts". A sign matrix is also referred to as a transformation matrix.

sequence for a 2^4 design can be written that is reasonably balanced for linear and quadratic trends with only 16 observations and the assumption there are no three-factor interaction effects.

An example provided by Daniel and Wilcoxon (1966) will be used to illustrate how a particular sequence is selected. First a 2^t factorial design sign (transformation) matrix is constructed that has the same number of independent observations as the number to be sequenced robust to linear and quadratic trend effects. Since they wished to order the experimental conditions of a replicated 2^2 factorial (written 2^{2+1}) requiring eight observations, they began by writing down the conventional sign matrix for a 2^3 factorial with eight independent observations capable of estimating a mean and seven independent effects. This is shown in Table 15. The experimental conditions have been arranged in the "standard order" proposed by Yates (see Davies, 1967, p. 261). Although the matrix is made up of plus and minus signs, it should be remembered that these are actually coefficients of plus and minus ones, with the digit being dropped for convenience. To the right of the matrix are the linear and quadratic

Table 15. 2^3 sign matrix tested for trend effects

Trial No.	Expt. Cond. In Standard Order	Sign (Transformation) Matrix								Orthog. Polynom. Coefficients	
		M	A	B	AB	C	AC	BC	ABC	L	Q
1	(1)	+	−	−	+	−	+	+	−	−7	+7
2	a	+	+	−	−	−	−	+	+	−5	+1
3	b	+	−	+	−	−	+	−	+	−3	−3
4	ab	+	+	+	+	−	−	−	−	−1	−5
5	c	+	−	−	+	+	−	−	+	+1	−5
6	ac	+	+	−	−	+	+	−	−	+3	−3
7	bc	+	−	+	−	+	−	+	−	+5	+1
8	abc	+	+	+	+	+	+	+	+	+7	+7
Inner Products	LX	0	8	16	0	32	0	0	0	168	
	QX	0	0	0	8	0	16	32	0		168
Design Coefficients	$21\rho^2(L,X)$		1	4	0	16	0	0	0		
	$21\rho^2(Q,X)$		0	0	1	0	4	16	0		

coefficients of a set of orthogonal polynomials for N = 8 (obtained from the published tables by Fisher and Yates, 1970).

The next step is to find out which effect (column) or effects are the least correlated (i.e., most robust) with the linear and quadratic coefficients. To illustrate how this is done, let us determine the correlation between the column of coefficients ±1 of Main Effect B and the coefficients of the linear polynomial in column L.

Step 1. Obtain the sum of the cross products between the coefficients in Column B and the coefficients in Column L. This number is called the "inner product" and the results are shown in the row marked (LX) below the appropriate column of the sign matrix. For LB, the inner product is 16 and was obtained as follows:

$$LB = (-1)(-7) + (-1)(-5) + (+1)(-3) + (+1)(-1)$$
$$+ (-1)(+1) + (-1)(+3) + (+1)(+5) + (+1)(+7) =$$
$$+7 \quad +5 \quad -3 \quad -1 \quad -1 \quad -3 \quad +5 \quad +7 \quad = \quad +16.$$

Note that for this operation, the digit 1 associated with each sign in Column B is introduced into the multiplication to show what is actually happening.

Step 2. Determine the sum of the coefficients squared in Column L. This also is given in Table 15. It is 168.

Step 3. Determine the sum of the coefficients squared in effects Column B. Since all of the numbers in any of the columns of the transformation matrix are either +1 or -1, which equal +1 when squared, the sum of the numbers of the eight observations squared will be 8 (in this example).

Step 4. Use the numbers obtained in the preceding steps to complete the following equation for the correlation squared:

$$\rho^2_{LX} = \frac{(\text{Inner product for coefficient and column})^2}{\text{Sum of coefficients squared} \quad X \quad \text{Sum of effects column coefficients squared}}$$

In this example, the correlation squared between Column B and the Linear trend is:

$$\rho^2_{LB} = \frac{16^2}{(168)(8)} = \frac{256}{1344} = .1905$$

The correlation between Column B and the linear trend would be the square root of that number, or .436.

The squared correlations, which the authors call the "design coefficients", are calculated for all of the relationships between the linear and quadratic coefficients and each of the columns of the transformation matrix. For this example, these are given in the last two rows of Table 15, designated $21\rho^2$ (L, X) and (Q, X). The 21 in this example, is used as a multiplier to the correlation squared to avoid awkward fractions and decimals. Since the numbers are used only comparatively, this modification suffices. (In the case of Column B and the linear polynomials, the correlation squared was)

$$.1905 = \frac{256}{1344} = \frac{4}{21}.$$

The number 4 therefore is used to represent the design coefficient for $21\rho^2$ (L, B), rather than 4/21 or .1905 for ρ^2 (L, B). The same modifier applies to all of the design coefficients in that Table. When a design coefficients is 0 or absent, it means that the two elements are uncorrelated and that that sequence is completely orthogonal to the particular trend.

Next let us take a very simple example to show how the design coefficients can be used to determine a sequence robust against trends. Upon examining the design coefficients in Table 15, the presence of a 0 shows that effects AB, AC, BC, and ABC are all ordered orthogonally to the linear trend and effects A, B, C, and ABC are all ordered orthogonal to the quadratic trend. If we wished to order the experimental conditions of a 2^{1+2} designs -- that's a single factor at two levels, replicated four times -- then we must select

one column (for one factor) that will be robust to linear and quadratic trends. In this case, effect ABC would be selected since its sequence is the only one not correlated with either trend. This sequence of eight conditions is used to order the high and low levels of the single factor of the 2^{1+2} design by making the + the high level, a, and - the low level, (1). This sequence would be:

Trial Nos.:	1	2	3	4	5	6	7	8
Treatments:	(1)	a	a	(1)	a	(1)	(1)	a

For a 2^{2+1} (two-factors at two levels, replicated twice) design, which also requires eight observations, the order that will eliminate or minimize linear and quadratic trend effects is more difficult to determine and will in fact require some compromise on complete orthogonality. The basic procedure however is the same. In Table 15, we must find two columns to represent each of the two factors that correlate little or none with both the L or Q coefficient. Furthermore, when these two columns are multiplied together (representing an interaction) this must result in a third column in the table that also correlates little or none with the L and Q coefficients. If we select Column ABC for one and AB for the other (both of which are acceptable columns with low design coefficients), we would be forced to use Column C for the third, which has a very high linear component and would not be acceptable. If we chose Columns ABC and A, both acceptable, we would be forced to use BC for the third, which has no linear component but has too high a quadratic component to be acceptable. Columns ABC and AC are not too bad, and the resultant third column B is also not too bad. The sum of all three design coefficients in Columns ABC, AC, and B is 8/21.

Daniel and Wilcoxon (1966) note that a slightly lower sum for three design coefficients can be achieved if Columns A, B and AB are used. This overall estimate of design efficiency would be 6/21. For the 2^{2+1} design, they named the two factors, Factors F and G, assigning Factor F to Column A and Factor G to Column AB, in order to keep the main effects the ones least correlated with L and/or Q effects. In addition, they reversed the signs of column AB so

*The reader could check this sequence for treatment independence from linear and quadratic trend effects by using the test described by Phillips in the previous section. Linear sums of trial numbers for both conditions are 18 and Quadratic sums of trial numbers squared for both are 102.

that the first condition in the sequence could be composed of the low levels of both factors. This caused column B with its sign reversed to become interaction FG of the new design. Reversal of signs of an entire column does not change the estimates of the factor effects nor their correlation with the trends. The results of these steps can be seen by comparing the appropriate columns in Tables 15 and 16, and the complete 2^{2+1} design is shown in the desired order in Table 16.

From the signs now in the new F and G columns (originally A and -AB respectively), the experimental conditions are renamed and ordered as shown in Table 16.

Table 16. Revised sign matrix of a 2^{2+1} design optimally ordered for trend

Trial No.	Original: A -AB -B Current: F G FG	Experimental Conditions
1	- - +	(1)
2	+ + +	fg
3	- + -	g
4	+ - -	f
5	- - +	(1)
6	+ + +	fg
7	- + -	g
8	+ - -	f

The experimental conditions are in the best sequence for minimizing the degree of overall nonorthogonality between main and interaction effects and linear and quadratic trends. The absolute correlations among the effects and the trends are:

	F	G	FG
L	.22	0	.44
Q	0	.22	0

Designs. Sequences prepared by Daniel and Wilcoxon (1966) for the following designs can be found in Appendix III:

Design	2^{2+1}	2^4	2^{6-2}	2^5	2^{14-9}
No. Obs.	8	16	16	32	32

Of particular interest are the 2^4 and 2^5 designs since these can be used in conjunction with economical designs for the screening of up to 15 or 31 factors (Simon, 1973).

If the reader is interested in other designs, the construction technique just described can be used for that prupose. The reader is encouraged to read the original paper; the material presented here should facilitate the understanding of that material.

Analysis. Given the specific sequence selected by the technique described above and the performance data associated with each experimental condition, Yates' analysis of variance algorithm (Yates, 1937; Davies, 1967, p. 263) can be used to estimate the effects. A number of examples illustrating this technique have been worked out by Daniels and Wilcoxon; one numerical example worked out for the order determined in Table 16 of this paper is shown here in Table 17.

Table 17. Numerical example of analysis by Daniel & Wilcoxon

Run	Spec.	"data"	−60	1	2	3	Name	+8	Corr.	Regr. Coeff.	Symbol
1.	(1)	63	3	6	16	0					
2.	*fg*	63	3	10	−16	−48	(*F*)	−6	$-\hat{L} = +1$	−5	$\hat{F}$
3.	*g*	73	13	6	−16	−24	−(*FG*)	+3	$2\hat{L} = -2$	+1	$-\widehat{FG}$
4.	*f*	57	−3	−22	−32	−32	−(*G*)	+4	$\hat{Q} = -1$	+3	$-\hat{G}$
5.	(1)	67	7	0	4	−32	$= 32\hat{L}$*			−1	$\hat{L}$
6.	*fg*	59	−1	−16	−28	−16	$(AC)_0$**			−1	$\hat{Q}$
7.	*g*	61	1	−8	−16	−32	$(BC)_0$**				
8.	*f*	37	−23	−24	−16	0	error			0	

* $32\hat{L} = -32$; $\hat{L} = -1$.

** $80\hat{Q} = (AC)_0 + 2(BC)_0 = -16 - 2 \times 32 = -80$; $\hat{Q} = -1$.

In the first column of Table 17 the trial number (position or period) for each of the experimental conditions in the second column is shown. These conditions are in the correct sequence to be robust against linear and quadratic trends starting from the top. In the third column, the "data" or performance scores corresponding to the corresponding experimental conditions are given. In the fourth column, the same scores are repeated after being reduced by subtracting 60 from every score to simplify the calculations.

In the next three columns (headed by the numbers 1, 2, and 3), the numbers are given as they are derived by using Yate's algorithm to estimate main and interaction effects. Since this is based on order of experimental conditions in the table that is not the "standard" order, a reidentification of the effects must be made. This correct identification of the effects is given in the "Names" column, which equates F, FG, and G in the new design with A, B, and AB respectively of the original design. Similarly, it relates 32L, AC_o, and BC_o, and error, with C, AC, BC, and ABC respectively. However, because there is some correlation between the main and interaction effects and the linear and quadratic trend components, a correction must be made. This is shown in the "*Corr.*" *double column, adjusting the effects* from the previous column (which were the results of taking the values from Yates' third column and dividing by N). How these corrections were obtained is shown on page 266 of Daniels and Wilcoxon's (1966) paper. Since each effect is associated with a single degree of freedom, we have obtained the regression coefficient for each by the above process. This is the weight to be assigned to these terms were they written in multiple regression form. The final column shows the symbol for the corrected values, namely for the two main effects, one interaction effect, linear and quadratic trend effects, and error. The reader may understand this process by referring to Daniel and Wilcoxon's paper for a more complete explanation. Yates' algorithm for systematically analyzing a 2^n design is described in Davies (1967, 263-265).

CHAPTER VI

SEQUENCES THAT REDUCE THE WORK REQUIRED TO RUN AN EXPERIMENT

In running a human factors engineering experiment, it may be difficult, disruptive, or time consuming to change the level of one or more factors every few trials. Some situations in which this might be so are when:

1. Complex electro-mechanical connections make it difficult to change the hand-control in an aircraft simulator.
2. Many dials have to be adjusted each time the factor level is changed, an action that is not only time consuming but also can increase the opportunity for erroneous settings.
3. Time consuming verification measurements have to be made at the face of a display each time display resolution is changed.
4. An electronic device requires considerable time to return to a steady state after each parametric change.
5. Time must elapse to permit the observer's eyes to dark adapt after each change in the illumination - particularly downward.

In some circumstances, these problems can be handled by employing "hierarchical", "nested", or "split-plot" experimental designs (Cochran and Cox, 1957 ; Myers, 1972; Winer, 1971) that presents experimental conditions in groups that don't require certain factors to be changed except after large blocks have been run. In other circumstances, when such blocked designs are not desirable, it becomes necessary to order experimental conditions in such a way that the number of factor-level changes required from trial to trial is minimized.

The problem of selecting the best sequence of experimental conditions can be complicated however, by the possible presence of time trends. In that case

experimental conditions should be presented in a sequence that requires as few factor-level changes as possible while remaining as robust as possible against temporally-related trend effects.

Two methods are given here for sequencing experimental conditions in 2^{n-p} designs to minimize the number of factor-level changes. The two methods develop sequences that:

1. Restrict the number of factor-levels that must be changed in adjacent trials.
2. Restrict the total number of factor-level changes to be made in the complete experiment while keeping the treatment effects robust against linear trend effects.

Standardized sequences for 8, 16, and 32 observations are provided in Appendix IV. These were obtained from computer programs that examined a great many alternatives to find those meeting the constraints imposed. Dickerson (1974, p. 33) provides a simplified flow diagram for the search algorithm he used in his program should the reader wish to develop his own sequences.

LIMITING THE NUMBER OF FACTOR-LEVEL CHANGES BETWEEN ADJACENT TRIALS

Tiahrt and Weeks (1970) proposed a method of "constrained randomization" for ordering experimental conditions of a 2^{n-p} factorial design that would restrict the number of factor-level changes between adjacent treatments. This method meets the requirements for a statistical analysis based on the randomization model when the usual technique of full randomization of order cannot be carried out. The method results in unbiased estimates of treatment effects and of experimental error.

Sequence Plans

In Appendix IV-A, standardized sequences that meet the above criteria are given for designing experiments involving 8, 16, or 32 experimental conditions.

All are for two-level factorial or fractional factorial designs that would restrict the number of factor-level changes between adjacent conditions to from one to four. The designs for which sequences are given in the appendix are shown in Table 18.

An example of one of the sequences in Appendix IV-A is:

$TG(2^{6-2}, \Delta \leq 3)$, $I = ABCD = CDEF$
1, 7, 4, 3, 6, 8, 2, 5, 13, 15, 9, 11, 14, 10, 12, 16

This would be used to create a one-fourth fractional factorial design for six factors (2^{6-2}) and would restrict the number of factor level changes between adjacent trials to three or less ($\Delta \leq 3$).

The (I = ABCD = CDEF) is the defining generator of the 2^{6-2} fractional factorial. This identifies the 16 experimental conditions out of the total 64 conditions in the complete 2^6 factorial that are to be ordered according to the above sequence. How defining generators are used to develop the conditions of fractional factorial designs will not be discussed here since they are treated thoroughly in numerous other documents (e.g., Simon, 1973; Davies, 1967; Cochran and Cox, 1957).

Construction

The following steps are used to transform the sequence of numerical values given in Appendix IV-A into a sequence of actual experimental conditions having some degree of randomization and providing the prescribed constraint on the number of factor level changes between adjacent trials. The 2^{6-2}, $\Delta \leq 3$ fractional factorial sequence just presented will be used to illustrate the method.

Step 1. Using the defining generator, specify the experimental conditions that are to be included in the design. Use the conventional symbology to designate each condition, i.e. by using a lower case letter when the high level of a factor

Table 18. Sequences given in Appendix IV-A that limit the number of factor-level changes (Δ) between adjacent trials

Experimental Designs	Δ = 1	Δ = 2	Δ = 3	Number of conditions
2^3	X	X		8
2^{4-1}		X	X	
2^{5-2}		X	X	
2^{6-3}		X	X	
2^4	X	X		16
2^{5-1}		X		
2^{6-2}		X	X	
2^{7-3}		X	X	
2^5	X	X		32
2^{6-1}		X		
2^{7-2}			X	

X = sequence is available

is used and no letter when the low level of a factor is used. If the design is a factorial, no special problem should be encountered in listing the experimental conditions. It it is a fractional factorial, then it is necessary to generate from the defining contrasts supplied with each plan the experimental conditions in the principal block (e.i., the one that contains the condition made up of the low level for all factors). An alternative method would be to find these conditions already listed in some publication on 2^{n-p} fractional factorials (e.g., Simon, 1973; Statistical Engineering Lab, 1957; Beyer, 1966).

The experimental conditions for the 2^{6-2} design (I = ABCD = CDEF) are listed in column I of Table 19.

Step 2. Rewrite the conditions of the experimental design in numerical terms in the conventional way by designating the low level of each factor as zero and the high level as 1, ordered from left to right for factor 1 through n respectively. This method of designating experimental conditions is described by Davies (1967), Cochran and Cox (1957), Simon (1973), and others.

The numerical representation of the experimental conditions in the 2^{6-2} design are found in column II, Table 19.

Step 3. Arrange the experimental conditions in the "standard" order as used with the Yates' algorithm for systematically computing factorial effects in 2^n designs (see Davies, 1967, p. 261; Cochran and Cox, 1957, p. 158; and others). This order starts with the experimental condition in which all factors are at their lowest level and develops progressively as follows:

(1) / a / b, ab / c, ac, bc, abc / d, ad, Bd...

When the design is a 2^n factorial, this ordering is easy to do.

Table 19. Transformation and ordering of the conditions in a 2^{6-2} design

I Experimental Conditions	II Numerical Representation A B C D E F	III Digital Equivalents of Binary Values	IV Standard Order Position Values	V New Order When $\Delta \leq 3$
(1)	0 0 0 0 0 0	0	1	1 (1)
ab	1 1 0 0 0 0	3	2	7 ade
cd	0 0 1 1 0 0	12	3	4 abcd
abcd	1 1 1 1 0 0	15	4	3 cd
ace	1 0 1 0 1 0	20	5	8 bde
bce	0 1 1 0 1 0	22	6	5 ace
ade	1 0 0 1 1 0	25	7	2 ab
bde	0 1 0 1 1 0	26	8	6 bce
acf	1 0 1 0 0 1	37	9	11 adf
bcf	0 1 1 0 0 1	38	10	14 abef
adf	1 0 0 1 0 1	41	11	12 bdf
bdf	0 1 0 1 0 1	42	12	15 cdef
ef	0 0 0 0 1 1	48	13	9 acf
abef	1 1 0 0 1 1	51	14	13 ef
cdef	0 0 1 1 1 1	60	15	10 bcf
abcdef	1 1 1 1 1 1	63	16	16 abcdef
Fractional factorial design No. 4.6.4 in Statistical Engineering Laboratory (1957) reference.	Numerical representation of experimental conditions, where 0 = low and 1 = high level of Factors A through F respectively	The digital equivalent of binary numbers that are written as the reverse of each row of the numerical rep. column II	The positions of the experimental conditions when sequenced according to the "standard" order	Standard order positions (col. IV) and corresponding experimental conditions (col. I) arranged so as to limit the factor-level changes between adjacent trials to three or less (Appendix IV-A, Table IX, TG 1)
I	II	III	IV	V

When the design is a fractional factorial, it may be easier to derive the digital equivalent of the numerical representation of the experimental condition to assist with the ordering. This is accomplished by writing the numerical representation (column II, Table 19) in reverse order and reading the resulting binary number as a digital number. Thus, the numerical representation of experimental condition ade is 100110, which reversed is 011001, which is the binary number for 25.

These digital values are arranged in an increasing series. This arrangement is shown in column III, Table 19.

Step 4. Each digital equivalent is assigned the number of its position in the standard order and these standard order position values (column IV, Table 19) are the ones referred to in the sequences presented in Appendix IV-A.

Step 5. Rearrange the experimental conditions according to the new sequence order of "standard order" position values. For example, the sequence for the 2^{6-2} design began with 1, 7, 4, 3, 6, and so forth. These numbers are found in column IV, Table 19, and are associated with the experimental conditions in column I, thus: (1), ade, abcd, cd, bce, ... respectively. That order is indicated in column V, Table 19.

Step 6. If the numerical representations are also listed in the same order, e.g.,

A	B	C	D	E	F	
0	0	0	0	0	0	(1)
1	0	0	1	1	0	ade
1	1	1	1	0	0	abcd
0	0	1	1	0	0	cd
0	1	1	0	1	0	bce

etc.

a check will show that the delta constraint is met, that there are never more and possibly less than three factor-level changes between adjacent trials.

To review the steps up to this point let us look at only one experimental condition that would be tested in a 2^{6-2} fractional factorial, condition ab (column I, Table 19). Its binary representation would be 110000 (column 2), which means that the condition would be a composite of the high levels of factors A and B and the low levels of factors C, D, E, and F. When those numbers are reversed, they become the binary number, 000011, equal to the digital number, 3. (column 3). This is second condition from the top of the list, and assigned the standard position value, 2, when the conditions of the 2^{6-2} design are put in a standard order (column 4). This second condition (ab) is placed in the 7th position (column 5) of the new sequence that limits factor-level changes between adjacent conditions to three or less.

Constrained randomization. But the procedure does not stop here. While a sequence has been developed that restricts the number of level changes between adjacent trials to a prescribed number, randomization has not yet been introduced. The process is therefore continued in order to introduce a condition of "constrained randomization" into the sequence derived above in Step 6.

Step 7. To introduce randomization into this constrained sequence, the names of the factors are randomally reassigned to the columns. Instead of the conventional order of factors A, B, C, D, E, F, for the respective columns, a random reassignment might, for example, produce C, F, A, B, E, D. The names of experimental conditions would now be changed to:

	C	F	A	B	E	D	
Digital equivalents arranged in the order given in column V	0	0	0	0	0	0	(1)
	1	0	0	1	1	0	cbe
	1	1	1	1	0	0	cfab
	0	0	1	1	0	0	ab
	0	1	1	0	1	0	fae
	etc.						

Experimental conditions rewritten according to factor reassignment

Step 8. The columns would now be rearranged so that the Factor designations are in alphabetical order. This in turn rearranges the letter designations of the experimental conditions alphabetically as well, thus:

A	B	C	D	E	F	
0	0	0	0	0	0	(1)
0	1	1	0	1	0	bce
1	1	1	0	0	1	abcf
1	1	0	0	0	0	ab
1	0	0	0	1	1	aef

etc.

Step 9. Further randomization is introduced by selecting at random one of the conditions in the new sequence and "multiplying" every condition in the series by it, thereby creating a new sequence made up of the same experimental conditions. For example, if we had randomally selected the condition bce of the series created in Step 8 and multiplied every condition by it, we would get the new and final sequence; bce, (1), aef, ace, abcf, and so forth.

The method of multiplication is a simple one; for example, multiplying abcf by bce would result in:

$$ab^2c^2ef = aef.$$

This follows the usual rules for algebraic multiplication, except that in this case any squared term equals one, meaning when other terms are present it is cancelled out. Thus

$$bce \times bce = b^2c^2e^2 = 1 \times 1 \times 1 = (1)$$

$$bce \times aef = abce^2f = abcf$$

and so forth.

The restraint of delta equals three or fewer changes will continue to hold with the new sequence. The column changes (Steps 7 and 8) and multiplication process (Step 9) introduce some randomization into an otherwise highly constrained sequence.

This procedure of constrained randomization should be used to construct all of the sequences shown in Appendix IV-A.

Multiple Subject Designs.

When more than one subject is to be tested in an experiment, different subjects are usually presented with the experimental conditions in different orders. Given one of the sequences generated here, the sequences for the other subjects may be generated by the multiplication process used in Step 9, using a different condition to create a new sequence for each different subject. The design thus created for t conditions and t subjects would represent a balanced square in which each sequence has the prescribed delta constraint on the maximum number of level changes between adjacent trials.

Balancing for Linear Trend Effects.

Tiahrt and Weeks did not select their sequences so that they would be balanced for trend effects. It is worth noting however that if an investigator can afford the time, if the same experimental conditions are run by the same subject a second time, but in the reversed sequence, the combined sequences provide a replication of data on which an error estimate can be made and the entire sequence is now completely balanced for a linear trend effect. This can be isolated from treatment effects in the manner described in Chapter V.

SEQUENCES ROBUST AGAINST LINEAR TRENDS THAT LIMIT FACTOR-LEVEL CHANGES

Draper and Stoneman (1968) and Dickinson (1974) provide standardized sequences for eight, sixteen, and thirty-two experimental conditions of factorial and fractional two-level designs that will keep the total number of factor-level changes small and the estimates of main effects free or nearly free of a linear time trend. These plans differ in the following ways from the previous plans by Tiahrt and Weeks (1970): 1) whereas the TW plans limited factor-level changes between adjacent trials to a fixed maximum, these DS&D plans minimize the total number of factor-level changes in the entire sequence; 2) a single sequence of the TW plans is not optimized for linear trend, while those of the DS&D plans are as well as possible within the factor-change constraints. The DS&D designs

when used as intended are limited however in two ways: 1) they are robust only against linear time trends and 2) this robustness is only for main effects and not for the interaction effects. Some ways of modifying these constraints will be described later.

Sequence Selection Criteria.

The sequences included in this section are intended to: 1) limit the number of factor level changes and 2) minimize the correlations ("time count") between main effects and linear time trends. The discussion below supplies the criteria for selecting a reasonable number of factor-level changes and for measuring the overall time count for any particular presentation order.

Factor-level changes. The minimum number of factor-level changes possible for any 2^n sequence consisting of N distinct experimental conditions is N - 1. This would be achieved by changing just one factor at a time between adjacent conditions starting with the second trial. However, the expected number of factor level changes for a 2^n factorial design* in a randomized sequence will be

$$n2^{n-1}$$

Thus for a 2^4 design, the minimum number of factor-level changes would be 15, but the number expected from a random sequence would be 32. Efficient designs from the standpoint of few factor-level changes should lie between these two values, and the designs offered here do.

Time count. An experimental design that is least affected by a linear time trend running through the experiment is one that minimizes the correlation between all experimental effects and the presentation order. Instead of calculating the correlation between each effect and the presentation order however a reduced calculation called a "time count" is ordinarily sufficient to select the presentation order that is best.

When a 2^n factorial or 2^{n-p} fractional factorial design is expressed in the form of a "sign matrix", e.g.

		FACTORS A	B	AB		ONE POSSIBLE ORDER OF PRESENTATION
CONDITIONS	(1)	-	-	+		1
	a	+	-	-	sign	2
	b	-	+	-	matrix	3
	ab	+	+	+		4

the sum of the cross-products between any main effect and the numbers representing the position of each condition in the presentation order is referred to as the time count. By way of illustration, let us determine the time count for Factor A and the presentation order in the above example. That order, of course, is only one of 24 possible orders for presenting four conditions. To obtain the time count we would obtain the sum of the cross-products in the following manner:

FACTOR A	PRESENTATION ORDER	CROSS-PRODUCT*
-	1	-1
+	2	+2
-	3	-3
+	4	+4
	Sum of cross-products	+2 = "time count"

Time counts can also be made for the interaction effects as well. Dickerson, upon calculating in several designs the time count between a linear trend and the two-factor interactions, concluded: "In all cases, these time

*It should be remembered that whereas for convenience we only show a sign that represents the high (+) or low (-) level of each factor, there is actually the number one associated with each sign. Thus, we are actually multiplying -1 times 1 = -1, +1 times 2 = +2 and so forth. The time count for any effect in a two-level factorial design can be converted to a correlation coefficient by dividing it by

$$N\left[(N^2 - 1)/12\right]^{1/2}$$

where N is the number of experimental conditions in the design. This denominator is 18.32, 73.76, and 295.68 respectively for 2^3, 2^4, and 2^5 factorial designs or fractional factorial with the same number of experimental conditions.

counts [for two-factor interactions] were large. Thus none of the run orders listed in [Dickinson's tables] is capable of providing estimates of the two-factor interactions which are not seriously biased by a linear time trend. The corresponding experimental designs should be used only as main-effect plans" (p. 34).

In the 2^4 and 2^5 designs, Dickinson not only gives the maximum time-count and the factor-level change for each design but also R^2, the squared multiple correlation between the four main effects and a linear time trend. This quantity was not used as a criterion in ranking the generated order runs but was computed separately. For the 2^5 sequences, Dickinson (1974) writes: "...the magnitudes of the squared multiple correlations are all small enough so that, for practical purposes, the five main effects are nearly orthogonal to linear time trend." (p. 35)

Sequences for Designs Containing 8, 16, or 32 Treatments

In Appendix IV-B, sequences for eight-run designs by Draper and Stoneman and sixteen and thirty-two runs by Dickinson are given. With all of these sequences, main effects are robust to linear trend effects and the total number of factor-level changes are limited. Characteristics of these sequences are listed in Table 20.

Here is an example of the kind of information supplied in the appendix. For an 8-observation, 2^3 design, the following four kinds of information would be supplied:

I	II	III	IV
14865732	3 4 2	9	-2 0 0

I is the sequence robust to linear trends while minimizing the factor level changes; the integers refer to the positions of the experimental conditions of a 2^3 design when they are arranged in the Standard order (see Davies, 1967, p. 261). II shows the number of changes in factor levels for variables A, B, and C respectively during 8 observations. III is the total number of factor level changes, i.e., the sum of II. IV gives the time count for variables A, B, and C respectively, which in this example, shows that only the first factor is correlated at all with the linear trend. The -2 time count when divided by

Table 20. Characteristics of sequences for 8, 16, or 32 treatments minimizing factor level changes to be found in Appendix IV-B

Design	Number of Main Effects	Number of Conditions	Total Number of Factor-Level Changes	Maximum Time Counts for any Main Effect
2^3	3	8	9 to 11	0 to 6
2^{4-1}	4	8	14 to 18	0 to 6
2^{5-2}	5	8	16 to 21	0 to 8
2^{6-3}	6	8	21 to 27	8 to 8
2^{7-4}	7	8	28	0 to 16
2^4	4	16	15	16 to 28
2^5	5	32	31	36 to 44

the 18.32 value indicated earlier would yield a correlation of -.11, which is not too great. Based on the number of level changes and the time counts, the experimenter would select the sequence that best meets his needs.

Using These Sequences

The design sequences for a particular factorial or fractional factorial can be selected on the basis of the maximum time count and total number of factor-level changes for any main effect. However, in practice, when an investigator can make valid assumptions about how important these criteria might be for the individual effects, he may employ the designs somewhat differently in order to increase their usefulness. Some of the alternatives ways of using the sequences in the appendix are discussed below.

1. If an investigator can anticipate which few interaction effects are most likely to be important in a factorial design, he may use a fractional factorial that handles more main effects than he intends to study to

find a sequence that is robust against linear trend effects for both main effects and selected interactions of a factorial design. The extra main effects would be redesignated interactions.

This can best be illustrated with an example. If in a three factor factorial design only one two-factor interaction is expected to be important, an optimized sequence for a 2^{4-1} fractional factorial can be used to optimize for all main effects and the single interaction. Since the fourth factor of a 2^{4-1} design has actually been confounded with an interaction of a 2^3 design, factor names can be assigned to the 2^{4-1} design in such a way that what might have been the fourth factor main effect in the 2^{4-1} design (for which the sequence was optimized) becomes the suspected interaction effect of the 2^3 design (for which the sequence is now also optimized). *

2. If the investigator expects to have trouble changing the factor-levels of only one of several factors, then the criteria used to select a sequence may be changed. Instead of selecting a sequence on the basis of the maximum number of factor level changes over all factors, it could be selected on the basis of the smallest number for a single factor.

By way of example, here are two sequences and their characteristics taken from Appendix IV-B.

	Sequences for a 2^3 Design	Changes in Each Factor			Total Changes	Time Count for each Factor		
		A	B	C		A	B	C
I)	14865732	2	4	2	9	-2	0	0
II)	14856732	5	4	2	11	0	0	0

Sequence II, on the whole, requires a few more factor level changes than Sequence I. In Sequence II, none of the main effects are affected by the linear time trend, while in Sequence I Factor A and the linear

* These relationships between factorial and fractional factorial effects are explained in Simon (1973) and Davies (1967) among others.

time trend are correlated slightly. To obtain the slight advantage in regard to the time trend, Sequence II might be selected if changing the factor levels is difficult for only one factor. If that one factor were Factor C in Sequence II, only two changes would be required, which is the same as for Sequence I.

Dickerson (1974, p. 36) proposes that when the cost of making a factor level change is not the same for all factors the experimenter might assign a specific cost for changing the levels of each factor, and developing sequences so weighted. This is essentially a refinement of the point made in the previous paragraph.

There are other alternatives open to the investigator who can use the sequencing concepts proposed in this chapter but is unable to find a design in the appendix that fits his particular needs. For example:

1. Dickinson mentions that a run order exists for a 2^4 experiment in which all four main effects and six two-factor interactions are completely uncorrelated with a linear time trend, but that it requires a total of 33 factor level changes. He did not indicate however how these are distributed among the main effects. Furthermore, as has already been noted, if there is only one factor that is difficult to change, a design that distributes total factor-level changes in such a way to have few changes for one or two factors and many for the others would be acceptable.

   An investigator might find sequences robust to trend effects by using the designs proposed by Daniel and Wilcoxin (discussed in the chapter on trends) as a basis. These designs are robust for linear and quadratic trends on main and interaction effects, but no information is supplied on the number of factor-level changes. An interested investigator could calculate these himself for selected designs.

2. An investigator might develop a computer program that will find those sequences that are optimum for linear and quadratic trends on main and critical interaction effects, and require minimum factor-level changes. In all probability, these sequences would have to be for longer runs (e.g., 2^5 perhaps). Dickerson (1974) supplies some clues as to how such a program could be developed.

Extending the Sequences

If blocks or replications of a design are to be run consecutively, Tiahrt and Weeks (1970, pp 479-480) explain how to order the sequence of experimental conditions in a 2^t factorial so that the number of factor level changes between adjacent treatments are restricted to a specified number.

Tiahrt and Weeks (1970) also point out that the sequences restricting the number of adjacent factor-level changes for two-level factorials can be immediately extended to symmetrical factorials with more than two levels per factor "provided any change of a factor is counted as one change even though the factor may be varied by more than just one level" (p. 480).

The eight run sequence for the 2^{7-4} fractional factorial design proposed by Draper and Stoneman (1968), and reproduced in Appendix IV-B can be used for ordering the experimental conditions in a seven factor "screening" experiment (see Simon, 1973; Box and Hunter, 1961). In that experimental design, only eight conditions are needed to estimate seven main effects (each being confounded with separate sets of interaction effects), which Box and Hunter suggest should be presented in a randomized order. The use of Draper and Stoneman's sequence, however, might be better for it would not only reduce the number of factor-level changes required in the experiment but would also make the main effects robust to any linear trends. If main effects are to be isolated from two-factor interactions in this screening designs, it is necessary to run eight additional conditions (composed of levels exactly opposite of those found in the original set). The sequence in which the second set of conditions are run can be the same as that in which the first set were run since both maximum time count and number of factor level changes are invariant with respect to a reversal of signs of all elements in any or all effects columns.

Analysis

The method of constrained randomization proposed by Tiahrt and Weeks (1970, 475-479) provides a basis for a statistical analysis utilizing a randomization model, that results in unbiased estimates of main and interaction effects

and of experimental error. The reader must refer to the original paper for this exposition. If an experimenter can assume that the errors are distributed normally and independently as in the randomization model rather than being correlated, then the usual routine analysis of a factorial experiment may be done. If these assumptions are wrong, then estimates of main and interaction effects may not be the best linear estimates and estimates of the variances will be incorrect. These statements however are not unique to these designs, but are true for all human factors experiments in which the treatments are applied sequentially.

For the designs that limit factor-level changes and are robust to linear trends, Dickerson (1974) writes that "significance testing can be carried out by pooling all of the interactions to form an error term providing there are no appreciable errors in setting the levels of the factors" (p. 36). Any reader who believes there would be errors in setting the levels of his factors should read Dickerson's discussion on how this would affect the analysis. He does point out the fact that even when standard significance tests are not strictly valid -- a situation that Simon (1971) suggests might be common in human factors engineering experiments -- the experiment may still serve to indicate the directions and magnitudes of the principal effects.

CHAPTER VII

DESIGNS FOR COMPARING DIFFERENT PRESENTATION ORDERS

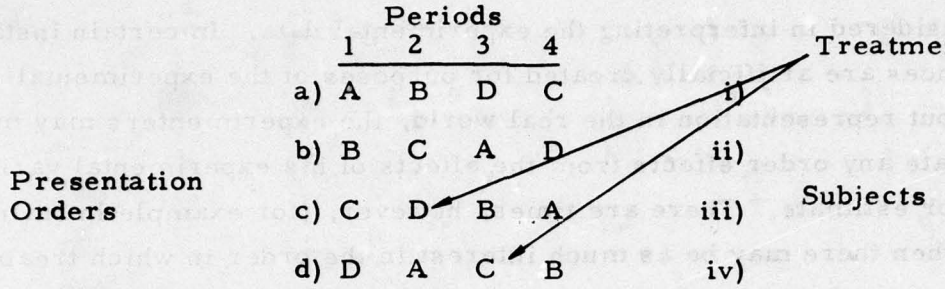
When treatments are presented in different orders to different groups of subjects, those differences among orders become a source of variance that must be considered in interpreting the experimental data. In certain instances, when sequences are artificially created for purposes of the experimental design without representation in the real world, the experimenters may merely wish to isolate any order effects from the effects of his experimental variables and his error estimate. There are times, however, (for example in training research) when there may be as much interest in the order in which treatments are to be presented to a trainee as there is in differences among the treatments themselves. In fact, when the treatments represent different courses in a study program, the order of presentation may be all important. Then it becomes important to know which order will result in the highest performance level when the sequence of study is completed. In addition to presentation order, other aspects of scheduling may also affect training effectiveness.

The reader should be warned that the development of systematic and economical ways of determining optimized orders and schedules that would meet the needs of applied training research has been totally inadequate. The various considerations can make the problem an enormously large and complex one. This section therefore is a limited attempt to define some of the problems and suggest some possible approaches. It is not within the scope of this report to develop new methods although the potential is apparent.

Three types of designs will be considered in this chapter. In the first section, a design is described that can isolate the average effect of presentation orders from treatment effects. In the second section, a design is described that might be used when there is an interest in the effects of different presentation orders but not in the effects of the different treatments. In the third section, designs are discussed that might be used to evaluate the effect on performance of different ways of scheduling different treatments.

DESIGNS THAT ISOLATE ORDER EFFECTS FROM TREATMENTS AND SUBJECTS

The following Latin square exemplifies an experimental design in which subjects are tested sequentially on the experimental conditions:



If this design were analyzed, the 15 degrees of freedom could be partitioned as follows (without considering residual cross-over effects):

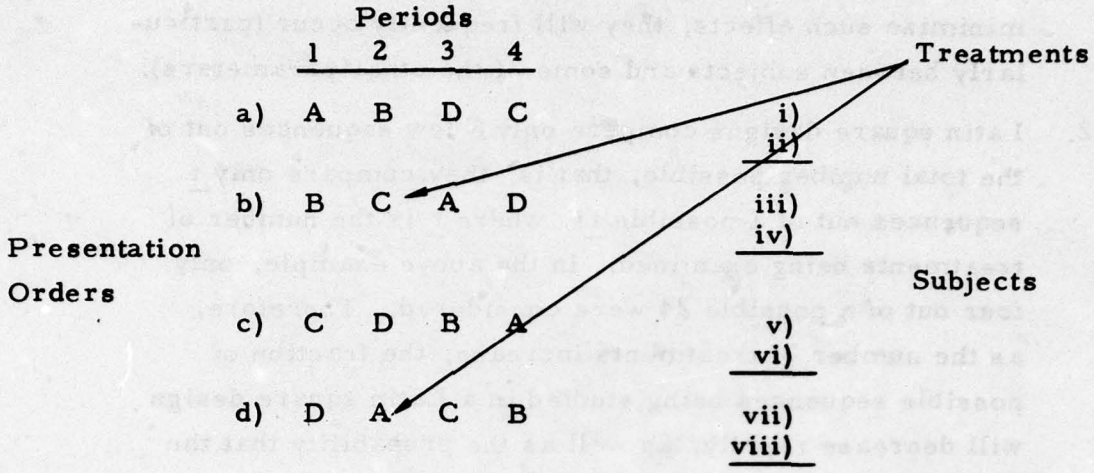
Periods (columns)	3
Orders/Subjects (rows)	3
Treatments (letters)	3
Error	6

The use of the Latin square design presupposes that there are no interactions between Treatments and Periods, Treatments and Subjects (or Orders), and Periods and Subjects (or Orders). If in fact there are any such interactions, then the results will be distorted since the effects of these interactions will not be isolated from main effects or error estimates.

But more germane to the present paper, in the above design, subject and order effects are totally confounded because each subject is given the treatments in a different order.

If subjects and orders are to be isolated, it is necessary to use the same presentation order with more than one subject. Preferably the same number of

subjects should be used with each order, and each of these matched on characteristics critical for the task. For example, testing two subjects on each order, the design would appear as:



and the total 31 degrees of freedom could be partitioned as follows:

Source of Variance	dif.
Between subjects	7
Between orders	3
Between subjects with the same order (E1)	4
Within subjects	24
Between trials	3
Between treatments	3
Error (E2)	18

This design and its analysis are treated fully in most advanced textbooks on statistics (e.g. Cochran and Cox, 1957; Davies, 1967: Winer, 1971).

The following shortcomings of these design should be noted:

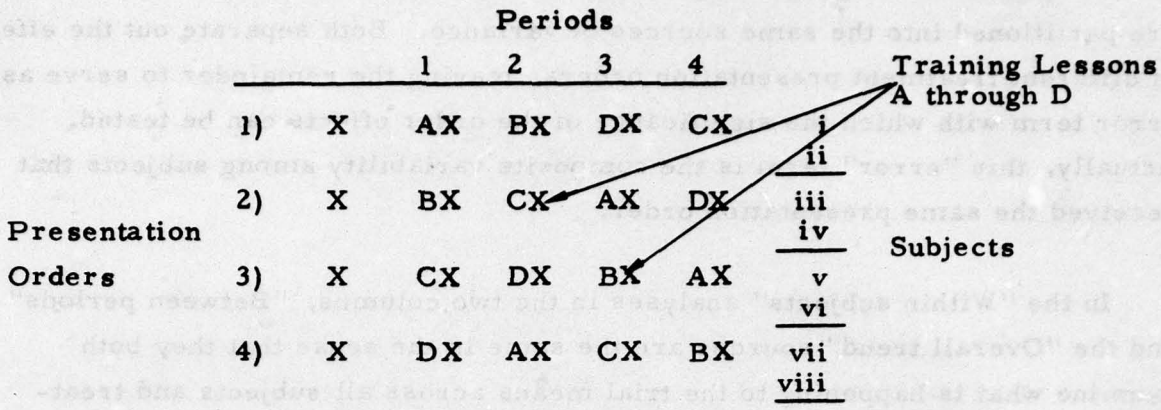
1. The assumption of no interactions between subject, trials, and treatments still holds. In human factors research, unless considerable care has gone into the planning to minimize such effects, they will frequently occur (particularly between subjects and some of the other parameters).

2. Latin square designs compare only a few sequences out of the total number possible, that is, they compare only t sequences out of a possible $t!$, where t is the number of treatments being examined. In the above example, only four out of a possible 24 were considered. Therefore, as the number of treatments increase, the fraction of possible sequences being studied in a Latin square design will decrease rapidly, as well as the probability that the best sequence has been included.

3. In so far as the above analysis is concerned, the only comparison being made is in the average performance achieved with each sequence. If sequences do have different effects, this could be reflected in different starting and ending points and a different rate of performance growth; the average could conceivably remain constant. While inspection of the raw data should confirm the actual situation, this analysis per se does not provide a quantitative description or comparison of learning rates as a function of presentation order.

DESIGNS THAT STUDY ORDER EFFECTS WHEN TREATMENT EFFECTS ARE NEGLIGIBLE

How might we study what effect presenting a series of training lessons in different orders will have on the rate and extent of learning? Actually a modified version of the previous design can be used and analyzed to provide this information. Here is a modified version of the design:



This time, the capital letters, A, B, C, and D, represent four different lessons of a training program that precede a constant test of performance, X. It is the performance on the constant test following the training session that will be analyzed. The First "X" condition in the design indicates that a performance test will also be made before any training begins. Thus, one can conceptualize this as an experiment in which a group of subjects are tested repeatedly on the same performance task. Sub-groups have all been given the same lessons, presented in different orders to each group. The data can be analyzed to answer the following questions:

1. Do different presentation orders result in different average performance levels?
2. What sort of performance change is there, if any, from period to period when the data among the different groups are combined?
3. What differences are there, if any, in the rate of change in performance among the groups given different presentation orders?
4. Which presentation order results in the highest final performance level?

To answer these questions, the 31 degrees of freedom associated with the foregoing experimental design will be partitioned as shown in Column B in Table 21. To show how this partitioning compares with that used in the analysis described on page 149, the earlier partitioning is repeated in Column A.

In spite of the slightly different labels, both "Between subjects" analyses are partitioned into the same sources of variance. Both separate out the effects of different treatment presentation orders, leaving the remainder to serve as an error term with which the significance of the order effects can be tested. Actually, this "error" term is the composite variability among subjects that received the same presentation order.

In the "Within subjects" analyses in the two columns, "Between periods" and the "Overall trend" sources are the same in the sense that they both examine what is happening to the trial means across all subjects and treatments (lessons). They differ however in the kind of information that is being extracted. In Column A, the typical analysis of variance merely indicates whether or not there is a reliable difference among the trial means. In Column B, using the orthogonal polynomials technique described in Chapter V, the trial (period) means are further analyzed to determine what overall trend components are present, if any, across all subjects and conditions. For P periods, (P - 1) trend components of one degree of freedom each can be estimated, although the presence of higher-than-cubic trends in significant amounts should be a signal to question the quality of the data.

The primary difference between these two analyses is in the partitioning of the final 21 degrees of freedom of the "Within subjects" variance. In a design of the Latin square type, one may <u>either</u> isolate the effects due to treatments <u>or</u> the effects of the interactions between periods and presentation order (when this order has been isolated from subjects). The differences between analyses A and B therefore occur because in A, treatment effects (3 degrees of freedom) were isolated while in B, a form of the periods-by-order interaction (9 degrees of freedom) is isolated. In the case of the B analysis, there is no interest in treatment differences (since they were considered to be different lessons of a complete course of study) and the question to be answered is whether or not there is a difference in how trial means change depending on the order in which the lessons are presented. Referring to this interaction as "Between order trend" indicates the intention of examining the trend characteristics exhibited

Table 21. Two ways of partitioning training data in a replicated Latin square design.

(A) Source of Variance		d. f.	(B) Source of Variance		d. f.
Between subjects		7	Between subjects		7
Between orders	3		Between orders (means)	3	
Between Ss in the same order (E_{bs})	4		Between individual means (E_{bs})	4	
Within Subjects		24	Within Subjects		24
Between periods	3		Overall trend	3	
			Linear	1	
			Quadratic	1	
			Cubic	1	
Between treatments	3		Between order trends	9	
			Linear	3	
			Quadratic	3	
			Cubic	3	
Residual (E_{ws})	18		Residual (E_{ws})	12	
TOTAL		31	TOTAL		31

by each presentation order. Orthogonal polynomials are used to determine the linear, quadratic, and cubic trend components for each order of presentation group. Then each trend component is compared among orders to see if they differ.

Since in all likelihood, it will be the linear component that is the largest, what we are essentially doing is comparing the slopes of the four sequences, that is, the rate of learning as a function of the order in which the lessons are

presented. The remaining components, quadratic and cubic, are comparing differences in non-linearity among the curves of the four sequences. If these differ significantly, it could indicate that with the lessons ordered in certain sequences performance will increase quickly in some cases and more slowly in others. While inspection of the raw data can give hints to all these results, the design and analysis presented here permit a test of the significance (reliability) of the observations.

In both analyses, what's left over after all other sources of variance have been removed is referred to as a residual and is used as the error estimate (E_{ws}) to test all of the Within Subject sources of variance.

This type of multiple trend analysis was suggested by Alexander (1946) and later elaborated upon by Grant (1956). The main difference between those two papers is that in Alexander's, only trends were partitioned into their "slope" (linear component) and "deviations from linearity" (a combination of all higher-order trend components), while Grant isolated every trend component as was done in the above example. Although there were more groups (orders) and subjects in Grant's design, the pattern of the analysis is identical to the example presented here. Both Alexander and Grant provide numerical examples and explanations of the complete analysis of these designs, as does Winer (1971, p. 590) who uses Grant's data for this purpose.

Additional questions could be asked of the design analyzed in Column B. An analysis of variance could be applied only to the final performance test results to determine which presentation order resulted in the highest final performance level. An analysis could also have been made on the pre-training performance measures (which would not have been used in the analysis in Column B) to see if the presentation order groups started at the same performance level. If not, the original performance test scores could be used as the covariant in a test of the final performance scores using the analysis of covariance described in Chapter V.

Considerations in the use of this design

This design was never intended to study the trends of different orders when the lessons have significantly different effects on performance. To study order-by-trials interactions (which is what the order trends are), as proposed originally, treatment effects must be negligible. In fact, if different lessons do have greatly different impacts on performance, then no study need be done to know what the trend would be for different presentation orders. Performance could be expected to rise fastest when the lessons having the greatest impact are presented first, and least when they are presented last. Of course, if the effects are additive, then in the end after all lessons had been presented, the final level in both cases would be the same.

But in spite of some rather predictable differences in the rate of learning when lessons have different effects, a study of the effect of order on trends might still be made on the premise that an entire order of presentation may have an effect that is greater than the sum of its parts. Even with non-negligible differences among the treatments, an examination of some trend effects appears reasonable. Certainly an examination of the slopes (linear trend) associated with the different orders is a realistic test to make, and possibly, if there are enough treatments involved, a quadratic trend test might also provide valid information. All higher-order trend effects could be considered as irrelevant variences that should be isolatated from those of interest, and being artificially introduced into the design, should not be combined with an estimate of error.

DESIGNS FOR STUDYING THE EFFECTIVENESS OF VARYING SCHEDULING PLANS

In the development of a training program, proper scheduling of the lessons that make up the complete course may be as important as the content of the specific lessons. Deciding how this best be done may require the aid of a systematic research plan.

Components of Scheduling

Several parameters must be selected in order to optimize the scheduling of training lessons. The more important ones are discussed below along with some preliminary considerations for the design of experiments suitable for answering the questions raised. The components of scheduling considered here include: order of presentation, lesson content selection, length of lessons, and repetition patterns.

Order of presentation. Which order of lessons A, B, and C will result in the best ultimate performance in the test "T" situation ?

Seq. 1: A B C T
Seq. 2: B C A T
Seq. 3: C A B T

Quite obviously, the three sequences do not cover all of the possible orders and it is apparent that as the number of lessons increase, the number of sequences required to cover all alternatives is (t!). As was discussed in the previous section, some counterbalanced design might be employed to reduce the number. Studying carry-over effects from trial-to-trial in a preliminary order study might give enough information about the relative transfer power of different lessons, that along with practical and rational consideration, a few nearly optimized orders could be proposed and tested for the final selection.

Lesson content selection. What lessons (A, B, C, or D) should be included in certain phases of the training session to optimize eventual test (T) results ? Examples of possible comparisons might be:

Seq. 1 A B C T_1 T_2
Seq. 2 A C D T_1 T_2
Seq. 3 B D T_1 T_2
Seq. 4 A T_1 T_2

In pilot training, for example, certain lessons in the simulator may not contribute very much to eventual pilot skill in the air and need not be included. Evaluating effective content may require two evaluation tests, in the simulator (T_1) and in the air (T_2).

Once again the total number of possible alternatives becomes excessive, namely (2^t), as soon as the number of lessons get very large. The quick and dirty use of "coactive" designs (DeGray, 1968) and multiple regression analyses (Hoerl and Kennard, 1970) may help to quickly identify the lessons and combinations of lessons (independent of order) that might be most profitable to investigate more carefully.

Length of lessons. While related to requirements set by the content, the length of time to be devoted to particular lessons may also be a sequencing question when all lessons are to be given, and when they are known to vary in the strength of their effect on the final tests. It may be desirable that those materials that result in greater transfer be practiced for longer times than others. Examples of possible comparisons among treatments A and B for different unit lengths are:

Seq. 1	AAAA	T
Seq. 2	AAA	T
Seq. 3	AA	T
Seq. 4	AABB	T
Seq. 5	ABBB	T
Seq. 6	AAAB	T

Once again, the potential number of alternatives that might possibly be tested, would be too great a task to seriously consider doing them all.

Length of sequence components might also take into consideration of what combination of practice time -- in a simulator(S) and in flight (F) -- would bring a pilot to a desired level of training a) the quickest, b) the cheapest, or c) some optimum combination of both of these. Examples of possible sequences might be:

Seq. 1	SSSSF
Seq. 2	SSSFF
Seq. 3	SSFFF
Seq. 4	SFFFF
Seq. 5	FFFFF

Repetition patterns. The sequences discussed up to this point presented each treatment only once, albeit for shorter or longer periods. There may be scheduling advantages in repeating a lesson after some other lesson has intervened

Thus a study of the repetition pattern may be a parameter of some importance. Sequences of the following types might be investigated:

Seq. 1 A B C B T
Seq. 2 A B C A B C T
Seq. 3 A T A B T A B C T

Innumerable combinations are of course possible.

Combinations. Order, content selection, lesson length, and repetition patterns are all scheduling variables which, when optimized, might enhance the rate of learning. If an experiment were planned that included all of these considerations, the possible alternatives to be studied in any unsystematic effort could quickly become astronomical.

Behavioral scientists facing this class of problem have failed to develop methods of approaching these problems systematically and economically. While analogous problems have been found in other disciplines, the methods and experimental designs employed by them have never been applied directly to psychological research. In particular, the rotation experiments for agricultural research appear particularly applicable.

Rotation Experiments

This class of experimental designs was first employed to study the long-term effects of different agricultural crop rotations. Different crops and different chemicals were systematically applied to different plots of ground in various orders for a number of years in order to compare the effect of different rotations on yield. While a primary function of rotation experiments has been to discover the order of chemical application that produces the highest yield, they can also be employed to investigate direct and residual sequential effects as well as cumulative effects (from the repeated application of the same treatment) and limiting effects (when a cumulative effect begins to stabilize).

It is beyond the scope of this paper to show how these marginally related rotation designs employed in agriculture might be applied to human training research. For an introduction to this class of design, the reader is referred to papers by Cochran (1939), Yates (1963), and Petterson (1964; 1965).

CHAPTER VIII
FUTURE NEEDS

Three important steps are required next to optimize the usefulness of this report. These have to do with

1. Integrating into a single multi -purpose plan the designs described in this report that handle different sequence effects.

2. Integrating the designs in this report with economical multi-factor designs (Simon, 1973).

3. Developing a more uniform technique of analysis for the different plans.

INTEGRATING THE DESIGNS IN THIS REPORT

Designs have been presented in this report that were intended to isolate trends, measure residual effects, and examine the effects of different presentations orders. While there are similarities among these designs, ordinarily only one or the other sequence effect would be examined. Ideally, these plans should be combined into a single design to handle several sequence effects within the same experiment. How this can best be accomplished needs to be examined.

INTEGRATING THESE DESIGNS WITH THOSE USED TO SELECT EXPERIMENTAL DATA POINTS

Designs that can be used to examine large multifactor experimental spaces by the economical selection of a relatively few data points (Simon, 1973) do not consider the sequence in which each subject will be tested on each experimental condition. The designs in this report however supply the sequences. Since economical multifactor research is developed around the use of fractional factorial designs made up of 8 and 16 (and higher if desired) experimental conditions, many of the plans discussed in this report could be used to

handle sequence effects during the data collection phase. Just how this should best be done and how it can best be accomplished without totally destroying the economy of data collection must be examined.

Integrating these two groups of designs is even more important when there is an interest both in treatment and sequence effects, as might be the case when equipment design and training research are combined. Designs that serve both purposes must be worked out in more detail.

DEVELOPING A MORE UNIVERSAL TECHNIQUE OF ANALYSIS

Methods of analyzing the results of the designs presented in this report have been dealt with sparingly, and in most cases, the only discussion can be found by referring back to the original documents. There, in some cases, clear explanations plus completely worked out examples have been provided. In other cases, only the theoretical basis for the analysis has been supplied. The aid of a statistician will be required.

The similarities among designs, along with the problems of non-orthogonal sources of variance, suggest that a general purpose computer program could be developed at moderate expense that would be suitable for conveniently analyzing all (or most) of the designs proposed. That this has not been done is primarily due to the fact that each innovator of a new technique was primarily concerned with ways of analyzing his particular technique. Also since the use of a general purpose analysis model would only be practical in conjunction with a computer facility, it should be noted that many of these techniques were developed before computers were as readily available as they are today.

The steps for developing a general purpose mode of analysis for these problems are as follows:

1. Examine each class of experimental designs to determine which designated sources of variance are independent and which are not, and in the latter case, the degree of non-orthogonality. Designated sources of variance include the more common items such as: Periods, Subjects, Treatments,

Direct (first and second), Residual (first and second), Linear Direct X Linear Residual, Order, and possibly some other interactions. A matrix of independence would be determined.

2. A multiple regression model would be developed to analyze each of these designs. Computer programs already in existence for handling correlated variables could be used (e.g. stepwise regression programs with the capability of specifying the order in which the terms are partialled out). Multiple analyses would be made, by permutating the order of designated correlated variables. (This is what was done in the analysis of the change-over design shown in Table 2 for example).

3. Sums of squares of each term would be recorded, so that appropriate terms of the regression equation (each with one degree of freedom) can be combined to permit tests of statistical significance.

There is very little new in the above procedure. If an investigator can identify those sources of variance that are independent of one another and those that are not, and can establish the rationale for ordering the partialling out of the correlated terms, a program such as BMD Biomedical Computer Program #BMD02R (Dixon, 1973) could be used immediately for the analysis.

REFERENCES

Alexander, H. W. A general test for trend. Psychological Bulletin, 1946, 43, 533-557.

Anderson, R. L., and Houseman, E. E. Tables of orthogonal polynomial values extended to N = 104. Iowa Agricultural Experimental Station Research Bulletin No. 297, 1942.

Atkinson, G. F. Designs for sequences of treatments with carry-over effects. Biometrics, 1966, 22, 292-308.

Bakan, D. The test of significance in psychological research. Psychological Bulletin, 1966, 66, 423-437.

Berenblut, I. I. Change-over designs with complete balance for first residual effects. Biometrics, 1964, 20, 707-712.

Berenblut, I. I. A change-over design for testing a treatment factor at four equally spaced levels. J. Royal Statistical Society, Series B, 1967, 29, 370-373.

Berenblut, I. I. Change-over designs balanced for the linear component of first residual effects. Biometrika, 1968, 55, 297-303.

Bose, R. C., Clatworthy, W. H., and Shrikhande, S. S. Tables of partially balanced designs with two associate classes. North Carolina Agricultural Experimental Station, Technical Bulletin 107, 1954.

Box, G. E. P. Multi-factor designs of first order. Biometrika, 1952, 39, 49-57.

Box, G. E. P., and Hay, W. A. A statistical design for the efficient removal of trends occurring in a comparative experiment with an application in biological assay. Biometrics, 1953, IX, 304-319.

Box, G. E. P., and Hunter, J. S. Experimental designs for the exploration and exploitation of response surfaces. In Chew, V. (Ed.) Experimental Designs in Industry. New York: Wiley, 1958, pp. 138-190.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Cochran, W. G. Long term agricultural experiments. Journal of the Royal Statistical Society, Series B, 1939, 6, 104-148.

Cochran, W. G. Analysis of covariance: its nature and uses. Biometrics, 1957, 13, 261-281.

Cochran, W. G., Autrey, K. M. and Cannon, C. Y. A double change-over design for daily cattle feeding experiments. Journal of Dairy Science, 1941, 24, 937-951.

Cochran, W. G. and Cox, G. M. Experimental Designs (2nd ed), New York: Wiley, 1957.

Cox, D. R. Some systematic experimental designs. Biometrika, 1951, 38, 312-323.

Cox, D. R. Some recent work on systematic experimental designs. Journal of the Royal Statistical Society, Series B, 1952, 14, 211-219.

Cox, D. R. Planning of Experiments. New York: John Wiley and Sons, Inc., 1958.

Daniel, C., and Wilcoxon, F. Factorial 2^{p-q} plans robust against linear and quadratic trends. Technometrics, 1966, 8, 259-278.

Davis, O. L. Design and analysis of industrial experiments. (2nd ed.) New York: Hafner. 1967.

Davis, A. W. and Hall, W. B. Cyclic change-over designs. Biometrika, 1969, 56, 283-292.

DeLury, D. B. Values of the intergrals of the orthogonal polynomials up to n= 26. University of Toronto Press. 1950.

DeGray, R. J. Design for interactions. Technometrics, 1968, 10, 389-391.

Dickinson, A. W. Some run orders requiring a minimum number of factors level changes for the 2^4 and 2^5 main effect plans. Technometrics, 1974, 16, 31-37.

Dixon, W. J. (Ed.) BMD biomedical computer programs. Berkeley, Calif.: University of California Press, 1973.

Dixon, W. J., and Massey, F. Introduction to statistical analysis. (2nd ed.) New York: McGraw-Hill, 1957.

Draper, R. N. and Stoneman, D. M. Factor changes and linear trends in eight-run two-level factorial designs. Technometrics, 1968, 10, 301-311.

Draper, N. R., and Smith, H. Applied regression analysis. New York: Wiley, 1968.

Dudeney, H. E. Amusements in mathematics. New York: Dover, 1943, pp 205-206.

Edwards, A. L. Experimental design in psychological research. New York: Rinehart, 1950.

Federer, W. T. Experimental design: theory and application. New York Macmillan, 1955.

Federer, W. T. and Atkinson, G. F. Tied-double-change-over designs. Biometrics, 1964, 20, 168-181.

Finney, D. J. Cross-over designs in bioassay. Procedures of the Royal Society of London, Series B., 1956, 145, 42-60.

Finney, D. J. and Outhwaite, Anne D. Serially balanced sequences. Nature, 1955, 176, 748.

Finney, D. J. and Outhwaite, Anne D. Serially balanced sequence in bioassay. Procedures of the Royal Society of London, Series B. 1956, 145, 493-507.

Fisher, R. A. The design of experiments. (5th ed.) New York: Hafner, 1949

Fisher, R. A., and Yates, F. Statistical tables for biological, agricultural and medical research. (6th ed.) London: Oliver and Boyd, 1963.

Geisser, S. A method for testing treatment effects in the presence of learning. Biometrics, 1959, 15, 389-395.

Guilford, J. P. Psychometric methods. (1st ed.) New York: McGraw Hill, 1936.

Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. Psychological Bulletin, 1956, 53, 141-153.

Hays, W. L. Statistics. New York: Holt, Rinehard, and Winston, 1963.

Hill, H. M. Experimental designs to adjust for time trends. Technometrics, 1960, 2, 67-81.

Hoerl, A. E., and Kennard, R. W. Ridge regression: applications to nonorthogonal problems. Technometrics, 1970, 12, 69-82.

Li, C. C. Introduction to experimental statistics. New York: McGraw-Hill, 1964.

Lindquist, E. F. Design and analysis of experiments in education and psychology. New York: Houghton Mifflin, 1953.

Lucas, H. L. Switchback trials for more than two treatments. Journal of Dairy Science, 1956, 39, 146-154.

Lucas, H. L. Extra-period latin-square change-over designs. Journal of Dairy Science, 1957, 40, 225-239.

Mason, J. M., and Hinkelmann, K. Change-over designs for testing different treatment factors at several levels. Biometrics, 1971, 27, 430-435.

Myers, J.L. Fundamentals of experimental design (2nd ed.) Boston: Allyn and Bacon, 1972.

Nair, C.R. Sequences balanced for pairs of residual effects. Journal of American Statistical Association, 1967, 62, 205-225.

Namboodiri, N.K. Experimental designs in which each subject is used repeatedly. Psychological Bulletin, 1972, 77, 54-64.

Patterson, H.D. The analysis of change-over trials. Journal of Agricultural Science, 1950, 40, 375-380.

Patterson, H.D. Change-over trials. Journal of the Royal Statistical Society, Series B, 1951, 13, 256-271.

Patterson, H.D. The construction of balanced designs for experiments involving sequences of treatments. Biometrika, 1952, 39, 32-48.

Patterson, H.D. and Lucas, H.L. Change-over designs, Technical Bulletin No. 147, North Carolina Agricultural Experiment Station and United States Department of Agriculture, September 1962.

Patterson, H.D. Theory of cyclic rotation experiments. Journal of the Royal Statistical Society, Series B, 1964, 26, 1-45.

Patterson, H.D. The factorial combination of treatments in rotation experiments. Journal of Agricultural Science, 1965, 65, 171-182.

Patterson, H.D. Nonadditivity in change-over designs for a quantitative factor at four levels. Biometrika, 1970, 57, 537-549.

Patterson, H.D. Quenouille's changeover designs. Biometrika, 1973, 60, 33-45.

Patterson, H.D. and Lucas, H.L. Extra-period change-over designs. Biometrics, 1959, 15, 116-131.

Phillips, J.P.N. The use of magic squares for balancing and assessing order effects in some analysis of variable designs. Applied Statistics, 1964, 13, 67-73.

Phillips, J.P.N. Methods of constructing one-way and factorial designs balanced for trends. Applied Statistics, 1968, 17, 162-170.

Poulton, E.C. Bias in experimental comparisons between equipments due to the order of testing. Ergonomics, 1969, 12, 675-687.

Poulton, E.C. Unwanted range effects from using within-subject experimental designs. Psychological Bulletin, 1973, 80, 113-121.

Sampford, M.R. Methods of construction and analysis of serially balanced sequences. Journal of the Royal Statistical Society, Series B, 1957, 19, 286-304.

Sheehe, P.R. and Bross, I.D.N. Latin squares to balance immediate residual, and other order effects. Biometrics, 1961, 17, 405-413.

Simon, C.W. Reducing irrelevant variance through the use of blocked experimental designs. Culver City, Calif.: Equipment Engineering Division, Hughes Aircraft Company, Tech. Report No. AFOSR-70-5, November 1970.

Simon, C.W. The use of central-composite designs in human factors engineering experiments. Culver City, Calif.: Equipment Engineering Division, Hughes Aircraft Company, Tech. Report No. AFOSR-70-6, December 1970

Simon, C.W. Consideration for the proper design and interpretation of human factors engineering experiments. Hughes Aircraft Company, Tech. Report No. P73-325, December 1971, 152 pp.

Simon, C.W. Economic multifactor designs for human factors engineering experiment. Hughes Aircraft Company, Tech. Report No. P73-326, June 1973, 171 pp.

Simon, C.W. An analysis of 239 experiments published in the journal, human factors, 1958 - 1972. Part I. Design and data collection characteristics. Part II. Nature of human factors data. Part III. Results of a survey on the degree to which this data has been applied to real systems. Hughes Aircraft Company, 1974. (IN PREPARATION)

Statistical Engineering Laboratory, Fractional factorial experiment designs for factors at two levels. Washington: Government Printing Office, National Bureau of Standards Applied Mathematical Series No. 48, 1957.

Stevens, W.L. The completely orthogonalized Latin square. Annuals of Eugenics, 1939, 9, 82.

Tiahrt, K.J. and Weeks, D.L. A method of constrained randomization for 2^n factorials. Technometrics, 1970, 12, 471-486.

Williams, E.J. Experimental designs balanced for the estimation of residual effects of treatments. Australian Journal of Scientific Research, 1949, 2, 149-168.

Williams, E.J. Experimental designs balanced for pairs of residual effects. Australian Journal of Scientific Research. Series A. 1950, 3, 351-363.

Winer, B.J. Statistical principles in experimental designs (2nd ed). McGraw-Hill Book Company, 1971.

Wold, H. Tracts for computors, no. 25. Cambridge University Press, 1948

Yates, F. The design and analysis of factorial experiments. Harpenden, England: Imperial Bureau of Soil Science, Technical Communication No. 35, 1937.

Yates, F. The Design of Rotation Experiments. Commonwealth Bureau of Soil Services, Technical Communication No. 46, 1949.

Youden, W.J. Randomization and experimentation. Technometrics, 1972, 14, 13-22.

APPENDICES

I. *CHANGE-OVER DESIGNS* (Chapter III)

- A. SINGLY-BALANCED
- B. DOUBLY-BALANCED
- C. DIRECT-BY-RESIDUAL (LINEAR)

II. SERIALLY-BALANCED SEQUENCES (Chapter IV)

- A. TYPE I, $k = 1$, $r = 1$
- B. TYPE I, $k = 2$, $r = 1$
- C. TYPE II, $k = 1$, $r = 1$
- D. TYPE II, $k = 1$, $r = 2$
- E. DUDENEY'S ROUND TABLE SOLUTIONS

III. SEQUENCING 2^{t-q} FACTORIALS TO BE ROBUST AGAINST TREND EFFECTS (Chapter V)

IV. SEQUENCES THAT MINIMIZE THE NUMBER OF FACTOR-LEVEL CHANGES

- A. BETWEEN ADJACENT TRIALS
- B. IN THE TOTAL EXPERIMENT WHILE ROBUST AGAINST LINEAR TREND EFFECTS

For each design, the reader should ascertain whether the rows or columns are periods or sequences. In some cases, an arrow will be used at the beginning to indicate the direction of the sequence over periods.

APPENDIX I - A

SINGLY-BALANCED CHANGE-OVER DESIGNS*

Either a complete design or the first sequences are provided. With all completed designs, the rows are the periods and the columns are the sequences and subjects. When only first sequences are provided, multisequences are given when multiple blocks are required to complete the balance. The additional information needed to construct PBIB change-over designs is also supplied. Type of design and characteristics are given with each plan. Total number of observations is N = pb. Construction techniques are described in Chapter III.

t = k = 4, p = 4, b = 1, n = 4 (Completed LS-CO design)*

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

*This design was not obtained using a cyclic solution. See footnote on page .

t = k = 5, p = 5, b = 2, n = 10 (Completed LS-CO design)

1	2	3	4	5
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2
4	5	1	2	3

1	2	3	4	5
5	1	2	3	4
2	3	4	5	1
4	5	1	2	3
3	4	5	1	2

t = k = 6, p = 6, b = 1, n = 10 (Completed LS-CO design)

1	2	3	4	5	6
2	3	4	5	6	1
6	1	2	3	4	5
3	4	5	6	1	2
5	6	1	2	3	4
4	5	6	1	2	3

t = k = 7, p = 4, b = 2, n = 14 (First sequences BIB-CO design, cyclic solution)

Block 1:	1	Block 2:	1
	2		7
	4		5
	7		2

(Appendix I-A, singly-balanced change-over designs, continued)

t = k = 8, p = 8, b = 1, n = 8 (First sequence LS-CO design, cyclic solution)

1
2
8
3
7
4
6
5

t = k = 9, p = 9, b = 2, n = 18
(First sequences LS-CO, cyclic)

	1		6
	2		5
Block 1	9	Block 2	7
	3		4
	8		8
	4		3
	7		9
	5		2
	6		1

t = k = 9, p = 5, b = 1, n = 9
(First sequence CCO design, cyclic)

1
2
4
3
6

t = k = 10, p = 10, b = 1, n = 10
(First sequence LS-CO, cyclic)

1
2
10
3
9
4
8
5
7
6

t = k = 10, p = 5, b = 1, n = 10
(First sequence CCO, cyclic)

1
4
2
9
8

t = k = 11, p = 6, b = 2, n = 22
(First sequences BIB-CO, cyclic)

	1		1
	2		11
Block 1	4	Block 2	9
	7		6
	11		2
	5		8

t = k = 11, p = 5, b = 1, n = 11
(First sequence CCO, cyclic)

1
5
8
2
3

t = k = 12, p = 12, b = 1, n = 12
(First sequence LS-CO, cyclic)

1
2
12
3
11
4
10
5
9
6
8
7

t = k = 12, p = 5, b = 1, n = 12
(First sequence CCO, cyclic)

1
2
6
5
8

t = k = 13, p = 5, b = 1, n = 13 (First sequence CCO design, cyclic)

1
3
4
8
5

t = k = 14, p = 14, b = 1, n = 14
(First sequence LS-CO, cyclic)

1
2
14
3
13
4
12
5
11
6
10
5
9
6
8
7

t = k = 14, p = 5, b = 1, n = 14
(First sequence CCO, cyclic)

1
4
5
12
6

(Appendix I-A, singly-balanced change-over designs, continued)

t = k = 15, p = 5, b = 1, n = 15 (First sequence CCO, cyclic)

1
2
6
5
8

t = k = 16, p = 16, b = 1, n = 16
(First sequence LS-CO, cyclic)

1
2
16
3
15
4
14
5
13
6
12
7
11
8
10
9

t = k = 16, p = 5, b = 1, n = 16
(First sequence CCO, cyclic)

1
9
6
7
2

t = 16, p = 4, k = 4, b = 8, n = 32 (BIB and model designs, PBCO)

BIB design *

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Model design **

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

(Eight blocks are obtained by using each row and each column of the BIB design to substitute in the first sequence (column) of the model design to form the first sequence of eight blocks of the new design. Each block is completed by treatment symbol substitution.)

*Bose et al Design 129
**Patterson and Lucas Design 5

(Sources from which the designs in Appendix I-A came are shown in Table 1 of the main text.)

APPENDIX I - B

DOUBLY-BALANCED CHANGE-OVER DESIGNS*

With these completed designs, rows represent periods and columns are sequences and subjects.

t = 4

I

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

II

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

III

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

t = 5

I

1	2	3	4	5
2	3	4	5	1
3	4	5	1	2
4	5	1	2	3
5	1	2	3	4

II

1	2	3	4	5
3	4	5	1	2
5	1	2	3	4
2	3	4	5	1
4	5	1	2	3

III

1	2	3	4	5
4	5	1	2	3
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2

IV

1	2	3	4	5
5	1	2	3	4
4	5	1	2	3
3	4	5	1	2
2	3	4	5	1

t = 6

I

1	2	3	4	5	6
2	3	4	5	6	1
3	4	5	6	1	2
6	1	2	3	4	5
4	5	6	1	2	3
5	6	1	2	3	4

II

1	2	3	4	5	6
3	4	5	6	1	2
4	5	6	1	2	3
6	1	2	3	4	5
5	6	1	2	3	4
2	3	4	5	6	1

III

1	2	3	4	5	6
4	5	6	1	2	3
5	6	1	2	3	4
3	4	5	6	1	2
2	3	4	5	6	1
6	1	2	3	4	5

IV

1	2	3	4	5	6
5	6	1	2	3	4
3	4	5	6	1	2
6	1	2	3	4	5
2	3	4	5	6	1
4	5	6	1	2	3

V

1	2	3	4	5	6
6	1	2	3	4	5
2	3	4	5	6	1
5	6	1	2	3	4
4	5	6	1	2	3
3	4	5	6	1	2

t = 7

I

1	2	3	4	5	6	7
2	3	4	5	6	7	1
3	4	5	6	7	1	2
4	5	6	7	1	2	3
5	6	7	1	2	3	4
6	7	1	2	3	4	5
7	1	2	3	4	5	6

II

1	2	3	4	5	6	7
3	4	5	6	7	1	2
5	6	7	1	2	3	4
7	1	2	3	4	5	6
2	3	4	5	6	7	1
4	5	6	7	1	2	3
6	7	1	2	3	4	5

III

1	2	3	4	5	6	7
4	5	6	7	1	2	3
7	1	2	3	4	5	6
3	4	5	6	7	1	2
6	7	1	2	3	4	5
2	3	4	5	6	7	1
5	6	7	1	2	3	4

IV

1	2	3	4	5	6	7
5	6	7	1	2	3	4
2	3	4	5	6	7	1
6	7	1	2	3	4	5
3	4	5	6	7	1	2
7	1	2	3	4	5	6
4	5	6	7	1	2	3

V

1	2	3	4	5	6	7
6	7	1	2	3	4	5
4	5	6	7	1	2	3
2	3	4	5	6	7	1
7	1	2	3	4	5	6
5	6	7	1	2	3	4
3	4	5	6	7	1	2

VI

1	2	3	4	5	6	7
7	1	2	3	4	5	6
6	7	1	2	3	4	5
5	6	7	1	2	3	4
4	5	6	7	1	2	3
3	4	5	6	7	1	2
2	3	4	5	6	7	1

t = 8

I

1	2	3	4	5	6	7	8
2	1	4	3	6	5	8	7
3	4	1	2	7	8	5	6
4	3	2	1	8	7	6	5
5	6	7	8	1	2	3	4
6	5	8	7	2	1	4	3
7	8	5	6	3	4	1	2
8	7	6	5	4	3	2	1

II

1	2	3	4	5	6	7	8
5	6	7	8	1	2	3	4
2	1	4	3	6	5	8	7
6	5	8	7	2	1	4	3
7	8	5	6	3	4	1	2
3	4	1	2	7	8	5	6
8	7	6	5	4	3	2	1
4	3	2	1	8	7	6	5

III

1	2	3	4	5	6	7	8
7	8	5	6	3	4	1	2
5	6	7	8	1	2	3	4
3	4	1	2	7	8	5	6
8	7	6	5	4	3	2	1
2	1	4	3	6	5	8	7
4	3	2	1	8	7	6	5
6	5	8	7	2	1	4	3

IV

1	2	3	4	5	6	7	8
8	7	6	5	4	3	2	1
7	8	5	6	3	4	1	2
2	1	4	3	6	5	8	7
4	3	2	1	8	7	6	5
5	6	7	8	1	2	3	4
6	5	8	7	2	1	4	3
3	4	1	2	7	8	5	6

V

1	2	3	4	5	6	7	8
4	3	2	1	8	7	6	5
8	7	6	5	4	3	2	1
5	6	7	8	1	2	3	4
6	5	8	7	2	1	4	3
7	8	5	6	3	4	1	2
3	4	1	2	7	8	5	6
2	1	4	3	6	5	8	7

VI

1	2	3	4	5	6	7	8
6	5	8	7	2	1	4	3
4	3	2	1	8	7	6	5
7	8	5	6	3	4	1	2
3	4	1	2	7	8	5	6
8	7	6	5	4	3	2	1
2	1	4	3	6	5	8	7
5	6	7	8	1	2	3	4

VII

1	2	3	4	5	6	7	8
3	4	1	2	7	8	5	6
6	5	8	7	2	1	4	3
8	7	6	5	4	3	2	1
2	1	4	3	6	5	8	7
4	3	2	1	8	7	6	5
5	6	7	8	1	2	3	4
7	8	5	6	3	4	1	2

t = 9

I

1	2	3	4	5	6	7	8	9
2	3	1	5	6	4	8	9	7
3	1	2	6	4	5	9	7	8
4	5	6	7	8	9	1	2	3
5	6	4	8	9	7	2	3	1
6	4	5	9	7	8	3	1	2
7	8	9	1	2	3	4	5	6
8	9	7	2	3	1	5	6	4
9	7	8	3	1	2	6	4	5

II

1	2	3	4	5	6	7	8	9
7	8	9	1	2	3	4	5	6
4	5	6	7	8	9	1	2	3
2	3	1	5	6	4	8	9	7
8	9	7	2	3	1	5	6	4
5	6	4	8	9	7	2	3	1
3	1	2	6	4	5	9	7	8
9	7	8	3	1	2	6	4	5
6	4	5	9	7	8	3	1	2

III

1	2	3	4	5	6	7	8	9
9	7	8	3	1	2	6	4	5
5	6	4	8	9	7	2	3	1
6	4	5	9	7	8	3	1	2
2	3	1	5	6	4	8	9	7
7	8	9	1	2	3	4	5	6
8	9	7	2	3	1	5	6	4
4	5	6	7	8	9	1	2	3
3	1	2	6	4	5	9	7	8

IV

1	2	3	4	5	6	7	8	9
8	9	7	2	3	1	5	6	4
6	4	5	9	7	8	3	1	2
9	7	8	3	1	2	6	4	5
4	5	6	7	8	9	1	2	3
2	3	1	5	6	4	8	9	7
5	6	4	8	9	7	2	3	1
3	1	2	6	4	5	9	7	8
7	8	9	1	2	3	4	5	6

V

1	2	3	4	5	6	7	8	9
3	1	2	6	4	5	9	7	8
2	3	1	5	6	4	8	9	7
7	8	9	1	2	3	4	5	6
9	7	8	3	1	2	6	4	5
8	9	7	2	3	1	5	6	4
4	5	6	7	8	9	1	2	3
6	4	5	9	7	8	3	1	2
5	6	4	8	9	7	2	3	1

VI

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
3	1	2	6	4	5	9	7	8
6	4	5	9	7	8	3	1	2
9	7	8	3	1	2	6	4	5
2	3	1	5	6	4	8	9	7
5	6	4	8	9	7	2	3	1
8	9	7	2	3	1	5	6	4

VII

1	2	3	4	5	6	7	8	9
5	6	4	8	9	7	2	3	1
9	7	8	3	1	2	6	4	5
8	9	7	2	3	1	5	6	4
3	1	2	6	4	5	9	7	8
4	5	6	7	8	9	1	2	3
6	4	5	9	7	8	3	1	2
7	8	9	1	2	3	4	5	6
2	3	1	5	6	4	8	9	7

VIII

1	2	3	4	5	6	7	8	9
6	4	5	9	7	8	3	1	2
8	9	7	2	3	1	5	6	4
5	6	4	8	9	7	2	3	1
7	8	9	1	2	3	4	5	6
3	1	2	6	4	5	9	7	8
9	7	8	3	1	2	6	4	5
2	3	1	5	6	4	8	9	7
4	5	6	7	8	9	1	2	3

---

*Designs for t = 4, 5, 7, 8, and 9 are complete sets of orthogonal Latin squares from Fisher and Yates (1963). Design for t = 6 was from Williams (1950, p. 354), which is isomorphic with solution given by Dudeney (1943).

APPENDIX I - C

DIRECT-BY-RESIDUAL INTERACTION CHANGE-OVER DESIGNS

i. Values of (X, Y) giving difference solutions from which designs can be constructed. Solutions for t = 5 through 10 were taken from Patterson (1973), while the others were derived using his method. Construction techniques are described in Chapter III.

t = 5

X 0 3 2

Y 1 4

t = 6

X 0 2 3

Y 1 5 4

t = 7

X 0 2 4 6

Y 1 3 5

t = 8

X 0 2 4 6

Y 1 3 5 7

t = 9

X 0 2 4 5 7

Y 1 3 6 8

t = 10

X 0 2 3 6 8

Y 1 5 4 7 9

t = 11

X 0 2 4 6 8 10

Y 1 3 5 7 9

t = 12

X 0 2 4 5 6 9

Y 1 3 7 8 11 10

t = 13

X 0 2 4 6 7 10 12

Y 1 3 5 8 9 11

t = 14

X 0 2 4 6 8 10 12

Y 1 3 5 7 9 11 13

t = 15

X 0 1 4 5 8 10 11 14

Y 2 3 7 6 9 13 12

t = 16

X 0 2 4 6 8 10 12 14

Y 1 3 5 7 9 11 13 15

(Appendix I - C, Direct-By-Residual Interaction Change-Over Designs continued)

ii. Completed designs for two and three factors provided by Mason and Hinkelmann (1971) that handle the D&R interaction residuals. Rows are periods and sequences and subjects are columns.

DESIGN FOR TWO TREATMENT FACTORS EACH AT TWO LEVELS

		Subject							
		I	II	III	IV	V	VI	VII	VIII
Period	1	A_1	A_1	A_2	A_2	B_1	B_1	B_2	B_2
	2	B_1	B_2	B_1	B_2	A_1	A_2	A_1	A_2
	3	A_2	A_2	A_1	A_1	B_2	B_2	B_1	B_1
	4	B_2	B_1	B_2	B_1	A_2	A_1	A_2	A_1

DESIGN FOR TWO TREATMENT FACTORS EACH AT THREE LEVELS

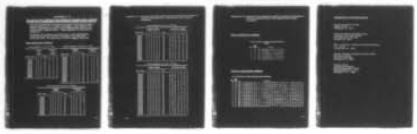
		Subject																
		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
Period	1	A_1	A_1	A_1	A_2	A_2	A_2	A_3	A_3	A_3	B_1	B_1	B_1	B_2	B_2	B_2	B_3	B_3	B_3
	2	B_1	B_2	B_3	B_2	B_3	B_1	B_3	B_1	B_2	A_1	A_2	A_3	A_2	A_3	A_1	A_3	A_1	A_2
	3	A_3	A_3	A_3	A_1	A_1	A_1	A_2	A_2	A_2	B_2	B_2	B_2	B_3	B_3	B_3	B_1	B_1	B_1
	4	B_2	B_3	B_1	B_3	B_1	B_2	B_1	B_2	B_3	A_2	A_3	A_1	A_3	A_1	A_2	A_1	A_2	A_3
	5	A_2	A_2	A_2	A_3	A_3	A_3	A_1	A_1	A_1	B_3	B_3	B_3	B_1	B_1	B_1	B_2	B_2	B_2
	6	B_3	B_1	B_2	B_1	B_2	B_3	B_2	B_3	B_1	A_3	A_1	A_2	A_1	A_2	A_3	A_2	A_3	A_1

DESIGN FOR THREE TREATMENT FACTORS EACH AT TWO LEVELS

		Subject																						
		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII	XXIII	XXIV
Period	1	A_1	A_1	A_1	A_1	A_2	A_2	A_2	A_2	B_1	B_1	B_1	B_1	B_2	B_2	B_2	B_2	C_1	C_1	C_1	C_1	C_2	C_2	C_2	C_2
	2	B_1	B_2	C_1	C_2	B_1	B_2	C_1	C_2	A_1	A_2	C_1	C_2	A_1	A_2	C_1	C_2	A_1	A_2	B_1	B_2	A_1	A_2	B_1	B_2
	3	C_1	C_2	B_1	B_2	C_2	C_1	B_2	B_1	C_1	C_2	A_1	A_2	C_2	C_1	A_2	A_1	B_1	B_2	A_1	A_2	B_2	B_1	A_2	A_1
	4	B_2	B_1	C_2	C_1	B_2	B_1	C_2	C_1	A_2	A_1	C_2	C_1	A_2	A_1	C_2	C_1	A_2	A_1	B_2	B_1	A_2	A_1	B_2	B_1
	5	C_2	C_1	B_2	B_1	C_1	C_2	B_1	B_2	C_2	C_1	A_2	A_1	C_1	C_2	A_1	A_2	B_2	B_1	A_2	A_1	B_1	B_2	A_1	A_2
	6	A_2	A_2	A_2	A_2	A_1	A_1	A_1	A_1	B_2	B_2	B_2	B_2	B_1	B_1	B_1	B_1	C_2	C_2	C_2	C_2	C_1	C_1	C_1	C_1

AD-A035 109 HUGHES AIRCRAFT CO CULVER CITY ... SPLAY SYSTEMS--ETC F/G 5/10
METHODS OF HANDLING SEQUENCE EFFECTS IN HUMAN FACTORS ENGINEERI--ETC(U)
DEC 74 C W SIMON F44620-72-C-0086
UNCLASSIFIED HAC-P74-541A NL

3 OF 3
AD
A035109



END

DATE
FILMED
3-77

APPENDIX II - A

SERIALLY BALANCED SEQUENCES (TYPE I, k = 1, r = 1)

Rows should be read in order as blocks. Blocks follow sequentially after one another. Each design is the sequence for a single subject. All of the designs were taken from Sampford (1957).

t = 6

(					A;)
A	F	B	E	C	D;
D	F	C	E	A	B;
B	A	C	F	D	E;
E	B	F	A	D	C;
C	B	D	A	E	F;
F	E	D	B	C	A;

t = 7

(						A;)
A	G	B	D	E	F	C;
C	A	B	F	D	G	E;
E	B	A	C	D	F	G;
G	F	E	A	D	C	B;
B	G	C	F	A	E	D;
D	B	C	E	G	A	F;
F	B	E	C	G	D	A;

t = 8

(							A;)
A	H	B	G	C	F	D	E;
E	C	G	D	F	H	A	B;
B	A	C	H	D	G	E	F;
F	B	E	G	A	D	H	C;
C	B	D	A	E	H	F	G;
G	B	H	E	A	F	C	D;
D	C	E	B	F	A	G	H;
H	G	F	E	D	B	C	A;

t = 9

(								A;)
A	I	E	D	F	C	H	G	B;
B	G	F	H	D	E	I	A	C;
C	I	F	E	G	H	B	A	D;
D	H	F	G	I	B	C	A	E;
E	H	I	G	C	B	D	A	F;
F	I	H	C	D	B	E	A	G;
G	D	I	C	E	B	F	A	H;
H	E	F	D	C	G	A	B	I;
I	D	G	E	C	F	B	H	A;

t = 10

(									A;)
A	J	B	I	C	H	D	G	E	F;
F	J	C	I	D	H	E	G	A	B;
B	A	C	J	D	I	E	H	F	G;
G	D	B	J	E	I	A	F	H	C;
C	B	D	A	E	J	F	I	G	H;
H	J	G	I	F	B	E	C	A	D;
D	C	E	B	F	A	G	J	H	I;
I	B	C	D	J	A	H	G	F	E;
E	D	F	C	G	B	H	A	I	J;
J	I	H	B	G	C	F	D	E	A;

t = 11

(										A;)
A	K	E	G	F	D	H	C	J	I	B;
B	I	G	D	J	F	H	E	K	A	C;
C	K	G	H	I	E	F	J	B	A	D;
D	I	H	F	G	J	K	B	C	A	E;
E	J	H	G	K	I	C	B	D	A	F;
F	I	K	H	J	C	D	B	E	A	G;
G	I	J	D	K	C	E	B	F	A	H;
H	K	J	E	D	C	F	B	G	A	I;
I	F	K	D	E	C	G	B	H	A	J;
J	G	E	H	D	F	C	I	A	B	K;
K	F	E	I	D	G	C	H	B	J	A;

t = 14

(..E;)	A	H	B	I	C	J	D	K	E	L	F	M	G	N;	N	L	J	I	K	M	H	E	D	B	G	A	C	F
	B	J	C	K	D	L	E	M	F	N	G	H	A	I;	I	H	L	N	M	K	J	F	E	B	C	D	A	G
	C	L	D	M	E	N	F	H	G	I	A	J	B	K;	K	N	J	H	M	I	L	G	E	C	B	D	F	A
	D	N	E	H	F	I	G	J	A	K	B	L	C	M;	M	L	I	J	K	H	N	A	D	E	F	C	G	B
	E	I	F	J	G	K	A	L	B	M	C	N	D	H;	H	K	L	M	J	N	I	B	A	E	G	F	D	C
	F	K	G	L	A	M	B	N	C	H	D	I	E	J;	J	L	H	I	M	N	K	C	E	A	B	F	G	D
	G	M	A	N	B	H	C	I	D	J	E	K	F	L;	L	K	I	N	H	J	M	D	G	C	A	F	B	E

APPENDIX II - B

SERIALLY BALANCED SEQUENCES (TYPE I, k = 2, r = 1)

Rows should be read in order as blocks. Blocks follow sequentially after one another. Each design is the sequence for a single subject. All of the designs were taken from Sampford (1957).

t = 4

(See page 77 in Chapter IV of main text.)

t = 5

(See Design B, Figure 4, Chapter IV in main text.)

t = 7

(See page 76 in Chapter IV of main text.)

t = 8

5; 5 6 4 7 3 8 2 1; 1 8 6 2 5 4 7 3; 3 4 2 5 1 6 8 7; 7 5 3 6 4 1 2 8;
8 1 7 2 6 3 5 4; 4 3 2 7 8 5 1 6; 6 7 5 8 4 1 3 2; 2 4 8 3 1 7 6 5;
5 6 7 1 3 8 4 2; 2 3 1 4 8 5 7 6; 6 1 5 8 7 2 3 4; 4 5 3 6 2 7 1 8;
8 2 1 4 6 3 5 7; 7 8 6 1 5 2 4 3; 3 7 4 5 2 6 8 1; 1 2 8 3 7 4 6 5;

APPENDIX II - C

SERIALLY BALANCED SEQUENCES (TYPE II, k = 1, r = 1)

Generating sequences are provided for this type of design. Two alternatives are given for each t values. To complete the design, remaining blocks in order are obtained by keeping fixed the position of the letter with the asterisk and permuting cyclically, in alphabetical order, the remaining (t-1) letters. Example of this is shown in text, Chapter IV.

t = 5	(i)	A	B	D	E*	C				
	(ii)	A	B	D	C	E*				
t = 6	(i)	A	F*	B	C	E	D			
	(ii)	A	B	F*	C	E	D			
t = 7	(i)	A	B	G*	F	C	E	D		
	(ii)	A	B	F	C	E	D	G*		
t = 8	(i)	A	H*	B	G	C	D	F	E	
	(ii)	A	B	G	C	H*	D	F	E	
t = 9	(i)	A	B	H	C	G	I*	D	E	E
	(ii)	A	B	H	C	G	D	F	E	I*
t = 10	(i)	A	J*	B	I	C	D	H	E	G	F
	(ii)	A	B	I	C	J*	D	H	E	G	F

Those sequences with the asterisk on the letter at the end of the block will all yield completely reversible designs. Other kinds of Type II, k = 1 designs are described in the text, Chapter IV. These designs came from Sampford (1957).

APPENDIX II - D

SERIALLY BALANCED SEQUENCES (TYPE II, k = 1, r = 2)

Rows are to be read in order as blocks. Blocks follow sequentially after one another. Each design is a sequence for a single subject. All of the designs were taken from Nair (1967).

(1) $t = 4$ ⟶

4 1; 2 4 3 1; 4 3 2 1; 3 4 2 1; 4 2 3 1; 2 3 4 1; 3 2 4 1;

(2) $t = 5$

5 1; 2 4 5 3 1; 4 3 5 2 1; 5 4 3 2 1; 3 5 4 2 1; 4 5 2 3 1;
5 2 4 3 1; 2 5 3 4 1; 3 2 5 4 1; 5 3 2 4 1; 2 3 4 5 1;
3 4 2 5 1; 4 2 3 5 1;

(3) $t = 6$

6 1; 2 4 5 6 3 1; 4 5 3 6 2 1; 5 4 6 3 2 1; 6 3 4 5 2 1;
3 6 5 4 2 1; 4 2 6 5 3 1; 5 6 4 2 3 1; 6 5 2 4 3 1;
2 6 3 5 4 1; 3 5 6 2 4 1; 5 3 2 6 4 1; 6 2 5 3 4 1;
2 3 6 4 5 1; 3 2 4 6 5 1; 4 6 2 3 5 1; 6 4 3 2 5 1;
2 5 4 3 6 1; 3 4 2 5 6 1; 4 3 5 2 6 1; 5 2 3 4 6 1;

(4) $t = 7$

7 1; 2 7 4 5 6 3 1; 4 6 3 7 5 2 1; 5 7 4 3 6 2 1; 6 7 5 4 3 2 1;
7 3 5 6 4 2 1; 3 6 5 4 7 2 1; 4 7 6 2 5 3 1; 5 4 6 7 2 3 1;
6 5 2 4 7 3 1; 7 5 6 2 4 3 1; 2 5 7 3 6 4 1; 3 5 2 6 7 4 1;
5 6 7 3 2 4 1; 6 2 7 5 3 4 1; 7 6 3 2 5 4 1; 2 6 3 4 7 5 1;
3 2 7 6 4 5 1; 4 2 3 7 6 5 1; 6 4 3 7 2 5 1; 7 4 6 2 3 5 1;
2 3 4 5 7 6 1; 3 7 4 2 5 6 1; 4 3 5 7 2 6 1; 5 2 7 3 4 6 1;
7 2 4 5 3 6 1; 2 4 6 5 3 7 1; 3 4 2 6 5 7 1; 4 5 2 3 6 7 1;
5 3 2 6 4 7 1; 6 3 5 4 2 7 1;

(5) $t = 8$

8 1; 2 4 7 6 5 8 3 1; 4 7 8 3 5 6 2 1; 5 6 4 7 3 8 2 1;
6 5 7 4 8 3 2 1; 7 4 3 8 6 5 2 1; 8 3 6 5 4 7 2 1;
3 8 5 6 7 4 2 1; 4 2 6 7 8 5 3 1; 5 8 2 4 6 7 3 1;
6 7 5 8 4 2 3 1; 7 6 8 5 2 4 3 1; 8 5 4 2 7 6 3 1;
2 6 5 3 8 7 4 1; 3 5 8 7 6 2 4 1; 5 3 7 8 2 6 4 1;
6 2 3 5 7 8 4 1; 7 8 6 2 5 3 4 1; 8 7 2 6 3 5 4 1;
2 8 3 7 4 6 5 1; 3 7 6 4 2 8 5 1; 4 6 2 8 7 3 5 1;
6 4 8 2 3 7 5 1; 7 3 4 6 8 2 5 1; 8 2 7 3 6 4 5 1;
2 3 8 4 7 5 6 1; 3 2 4 8 5 7 6 1; 4 8 7 5 3 2 6 1;
5 7 3 2 8 4 6 1; 7 5 2 3 4 8 6 1; 8 4 5 7 2 3 6 1;
2 5 6 8 3 4 7 1; 3 4 2 5 8 6 7 1; 4 3 5 2 6 8 7 1;
5 2 8 6 4 3 7 1; 6 8 4 3 2 5 7 1; 8 6 3 4 5 2 7 1;
2 7 4 5 6 3 8 1; 3 6 7 2 4 5 8 1; 4 5 3 6 2 7 8 1;
5 4 6 3 7 2 8 1; 6 3 2 7 5 4 8 1; 7 2 5 4 3 6 8 1;

APPENDIX II - E

DUDENEY'S ROUND TABLE SOLUTIONS

The following material was taken from Dudeney's (1943, pp. 205-206) "Amusements in Mathematics". They are the solutions to a permutation problem of finding the ways of arranging n people -- "Knights of the Round Table" -- in a ring so that no one has the same two neighbors more than once but has each neighbor an equal number of times. These solutions have been used by Williams (1950) to form doubly-balanced change-over designs and Nair (1967) to form serially balanced sequences for pairs of residuals.

Three kinds of information will be provided for each number of treatments:

1) The first row of each block
2) Each t-value that is a repeater
3) The sets of numbers that are cycled

For example, for t = 4, there is only one block. The first row is

1 2 3 4

and 1 is the repeater and 2, 3, and 4 are cycled down the columns, thus:

1 2 3 4
1 3 4 2
1 4 2 3

The three rows each represent circles or arrangements in which four persons can be seated so that no one has the same two partners on either side twice. In some solutions, there will be more than one block, or one repeater, or one cycle.

For t = 5 to 12, the first row of each block, each repeater, and each cycle will be given here:

(Appendix II-E, Dudeney's round table solutions - Continued)

t	First row of blocks	Repeaters	Cycles
5	12345 13254	1, 2	(3, 4, 5)
6	123645 124563	1	(2, 3, 4, 5, 6)
7	1234576 1627534 1352674 1574362 1527346	1	(2, 3, 4) and (5, 6, 7)
8	18634527 18457236 18273645	1	(2, 3, 4, 5, 6, 7, 8)
9	219745638 295168347 293184756 291564783	1, 2	(3, 4, 5, 6, 7, 8, 9)
10	1X83654729 1X65297438 1X29386574 1X74832956 (X = 10)	1	(2, 3, 4, 5, 6, 7, 8, 9, 10)
11	2E94765183X 21E763X8549 2EX39485176 2E5813X6794 2E1X3496758 (E = 11, X = 10)	1, 2	(3, 4, 5, 6, 7, 8, 9, 10, 11)
12	123T4E5X6978 124E6987X5T3 125X87E43T69 1269X53T78E4 1278T369E45X (T = 12, E = 11, X = 10)	1	(2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)

Since each row does represent a circle, the solution is not changed if the row is cycled; thus 2134 could be written 1342 or 3421 or 4213.

APPENDIX III

SEQUENCING 2^{t-q} FACTORIALS TO BE ROBUST AGAINST TREND EFFECTS

Sequences developed by Daniel and Wilcoxin (1966) are reproduced below for specified 2^{t-q} factorial designs involving 8, 16, or 32 experimental conditions. All are robust against linear and quadratic trend effects. The construction and analysis of these designs are discussed in Chapter V.

Columnar headings in these tables are defined as follows:

Run: These are the trial numbers for the corresponding experimental conditions, ordered robust against the trend effects.

Spec.: These are the experimental conditions. In accordance with standard practice, the presence of a letter means that the high level of a two-level factor is used to make up the experimental condition; the absence of a letter means that the low level of the missing factor is used.

Alternative Specification: The same experimental design with a different set of letters assigned to each factor.

Ordered Contrasts: These identify the main and interaction effects that would be estimated were the factorial solved using Yates' algorithm for systematically analyzing 2^n designs (Davies, p. 263).

Aliases: These indicate the effects that are confounded for the particular fractional factorial design and cannot be isolated within that design.

Correction for Trends: These are the values needed in the analysis to adjust an experimental effect for linear and quadratic trend effects, as explained in the text in Chapter V.

Efficiencies: This refers to the relative loss of information when an effect must be corrected for trends; information refers to the confidence in the data when compared with effects requiring no correction. The highest possible of course is 1; none are less than .70.

Summary for the ordered 2^{2+1}

Run.	Spec.	Ordered Contrasts	Corrections	Efficiencies	Names
1.	(1)	Total			
2.	fg	(F)	$-\hat{L}$	16/17	$\hat{F}$
3.	g	-(FG)	$+2\hat{F}$	4/5	$\hat{FG}$ #
4.	f	-(G)	$+\hat{Q}$	20/21	$\hat{G}$
5.	(1)	$(C)_o = 32\hat{L}$		16/21	$\hat{L}$ ##
6.	fg	$(AC)_o$*		20/21	$\hat{Q}$
7.	g	$(BC)_o$*			
8.	f	$(ABC)_o$ = error		1	

*$80Q = (AC)_o + 2(BC)_o$

(Appendix III, sequencing 2^{t-q} factorials to be robust against trend effects, continued)

Summary for the Ordered 2^4

Run	Spec.	Ordered Contrasts	Corrections for Trends $\hat{L}$	Corrections for Trends Q	Efficiencies	Alternative Specification
1.	(1)					(1)
2.	*fgj*	$-(HJ)$	$+1^{**}$		64/65	*abd*
3.	*fhj*	$-(GJ)$	+2		16/17	*acd*
4.	*gh*	$+(GH)$		−1	336/337	*bc*
5.	*ghj*	$-(FJ)$	+4		4/5	*bcd* #
6.	*fh*	$+(FH)$		−2	84/85	*ac*
7.	*fg*	$+(FG)$		−4	21/22	*ab*
8.	*j*	$-(FGHJ)$				*d*
9.	*fghj*	$+(FGH^{*} = 128\hat{L}$			64/85	*abcd* ##
10.	*h*	$-(FGJ)^{*}$				*c*
11.	*g*	$-(FHJ)^{*}$				*b*
12.	*fj*	$+(F)$			1	*ad*
13.	*f*	$-(GHJ)^{*}$				*a*
14.	*gj*	$+(G)$			1	*bd*
15.	*hj*	$+(H)$			1	*cd*
16.	*fgh*	$-(J)$			1	*abc*

* Three-factor interactions assumed zero

$1344Q = -(FGJ) - 2(FHJ) - 4(GHJ) = -(AD)_6 - 2(BD)_6 - 4(CD)_6$ 16/17

** Thus: $\widehat{HJ} = (HJ)/16 + \hat{L}$; $\widehat{GJ} = (GJ)/16 + 2\hat{L}$.

*Summary for the Ordered 2^{6-2}, All Three-Factor Interactions Zero**

Run	Spec.	Ordered Contrasts and Aliases	Corrections for Trends $\hat{L}$	Corrections for Trends Q	Efficiencies
1.	(1)				
2.	*abde*	$-(AF + CD)$	−1		64/65
3.	*acdf*	$-(AE + BD)$	-2^{**}		16/17
4.	*bcef*	$(BC + EF)$		−1	256/257
5.	*bcd*	$-(AD + BE + CF)$	−4		4/5 #
6.	*ace*	$(AC + DF)$		−2	8/9 #
7.	*abf*	$(AB + DE)$		−4	16/17
8.	*def*	$-(BF + CE)$			1
9.	*abcdef*	$(ABC) = 128\hat{L}$			64/85 ##
10.	*cf*	$-(E)$		−4	16/17
11.	*be*	$-(F)$		−8	4/5 #
12.	*ad*	(A)			1
13.	*acf*	$-(BCD) = -256Q$			256/257 ##
14.	*bdf*	(B)			1
15.	*cde*	(C)			1
16.	*abc*	$-(D)$			1

* Confounding pattern: $I + ABDE + ACDF + BCEF$

** Thus: Est. $(AE + BD) = (AE + BD)/16 + 2\hat{L} = (AE + BD)/16 + (ABC)/64$.

(Appendix III, sequencing 2^{t-q} factorials to be robust against trend effects, continued)

Summary for the Ordered 2^5

Run	Spec.	Ordered Contrasts	Corrections for Trends $\hat{L}$	Corrections for Trends Q	Efficiency $= k/(k+1)$ e
1.	(1)				
2.	abe	$-(CE)$	+1		320
3.	ace	$-(BE)$	+2		80
4.	bc	$+(BC)$		−1	5696
5.	bcde	$-(AE)$	+4		20
6.	acd	$+(AC)$		−2	1424
7.	abd	$+(AB)$		−4	356
8.	de	$-(ABCE)$			
9.	d	$-(BCDE)$**			–
10.	abde	$+(BD)$		−4	356
11.	acde	$+(CD)$		−8	88
12.	bcd	$-(DE)$			•
13	bce	$+(ABCD)$			–
14.	ac	$-(ABDE)$			•
15.	ab	$-(ACDE)$			•
16.	e	$+(AD)$			•
17.	abcde	$+(ABC)$**			–
18.	cd	$-(ABE)$***			–
19.	bd	$-(ACE)$***			–
20.	ade	$+(A)$			•
21.	a	$-(BCE)$***			–
22.	be	$+(B)$			•
23.	ce	$+(C)$			•
24.	abc	$-(E)$			•
25.	abce	$-(ADE)$***			–
26.	c	$+(ACD)$			•
27.	b	$+(ABD)$			•
28.	ae	$-(ABCDE)$			•
29.	ad	$+(D)$			•
30.	bde	$-(CDE)$			•
31.	cde	$-(BDE)$			•
32.	abcd	$+(BCD)$			•

* These contrasts have efficiency 1

** $1280\hat{L} = -(BCDE) + 2(ABC) = (D)_0 + 2(E)_0$; Eff. $\hat{L} = 320/341$

*** $21760Q = -(ABE) - 2(ACE) - 4(BCE) - 8(ADE)$
$= (AE)_0 + 2(BE)_0 + 4(CE)_0 + 8(DE)_0$. Eff. $Q = 320/341$

Zero subscripts refer to original names of contrasts.

(Appendix III, sequencing 2^{t-q} factorials to be robust against trend effects, continued)

Summary of the Ordered 2^{14-9}, All Three-Factor Interactions Assumed Zero

Run	Spec.	Ordered Contrasts and Aliases	Corrections for Trends $\hat{L}$	Corrections for Trends $\hat{Q}$	Eff. = $k/(k+1)$ / k
1.	(1)				
2.	abdegjlm	$-(AN + BO + CD + FG + HJ + KL)$	+1		256
3.	acdfgkln	$-(AM + BD + CO + EG + HK + JL)$	+2		64
4.	bcefjkmn	$+(BC + DO + EF + HL + JK + MN)$		−1	4096
5.	bcdhjklo	$-(AD + BM + CN + EJ + FK + GL)$	+4		16
6.	acegkmo	$+(AC + DN + EH + FL + GK + MO)$		−2	1024
7.	abfghjno	$+(AB + DM + EL + FH + GJ + NO)$		−4	256
8.	defhlmno	$-(AO + BN + CM + EK + FJ + GH)$			•
9.	efghjkl	$-(AG + BJ + CK + DL + EM + FN + HO)$	+8		4#
10.	abdfhkm	$+(AF + BH + CL + DK + GN + JO)$		−4	256
11.	acdehjm	$+(AE + BL + CH + DJ + GM + KO)$		−8	64
12.	bcghlmn	$-(BK + CJ + DH + EN + FM + LO)$			•
13.	bcdefgo	$+(AL + BE + CF + DG + JM + KN)$		−16	16
14.	acfjlmo	$-(AK + CG + DF + EO + HM + NL)$			•
15.	abeklno	$-(AJ + BG + DE + FO + HN + LM)$			•
16.	dgjkmno	$+(AH + BF + CE + GO + JN + KM)$			•
17.	abcdefghjkimno	$2^9\hat{L}$			••##
18.	cfhkno	$-(M)$		−8	64
19.	behjmo	$-(N)$		−16	16
20.	adghlo	$+(A)$			•
21.	aefgmn	$-(O)$		−32	4#
22.	bdfjln	$+(B)$			•
23.	cdeklm	$+(C)$			•
24.	abcgjk	$-(D)$			•
25.	abcdmno	$2^{11}\hat{Q}$			••••##
26.	cegjlno	$+(E)$			•
27.	bfgklmo	$+(F)$			•
28.	adefjko	$-(G)$			•
29.	ahjklmn	$+(H)$			•
30.	bdeghkn	$-(J)$			•
31.	cdfghjm	$-(K)$			•
32.	abcefhl	$+(L)$			•

* These effects have efficiency 1.

** $2^9\hat{L} = (E)_6$, i.e. standard contrast (E). Eff. $\hat{L}$ = 256/341

*** $2^{11}\hat{Q} = (DE)_6$, i.e. standard contrast (DE). Eff. $\hat{Q}$ = 4096/5797

#Eff. less than .90

##Eff. between .70 and .80

All others higher than .90

APPENDIX IV - A

SEQUENCES THAT MINIMIZE THE NUMBER OF FACTOR-LEVEL CHANGES BETWEEN ADJACENT TRIALS

The following tables present transformation generators for particular 2^{n-p} factorials and for particular compatibility conditions, i.e., values of Δ. The number in parentheses identifies a transformation generator and the sequence 1, 2, ... , N is equivalent to corresponding treatment combinations in the particular set under consideration. The order of the set of treatment combinations is the usual increasing order of base 2 numbers. Thus, for example, in a 2^2 factorial 1 = 00, 2 = 01, 3 = 10, 4 = 11. All of the sequences were taken from Tiahrt and Weeks (1970). Construction techniques are described in Chapter VI.

TABLE I

$TG(2^2, \Delta = 1)$;
$TG(2^{4-2}, \Delta = 2)$, I = AB = CD
(1) 1, 2, 4, 3

TABLE II

$TG(2^3, \Delta = 1)$;	
$TG(2^{6-3}, \Delta = 2)$, I = AB = CD = EF	
tg Number	Sequence of tc's
(1)	1, 2, 4, 3, 7, 5, 6, 8
(2)	1, 2, 4, 3, 7, 8, 6, 5
(3)	1, 2, 4, 8, 6, 5, 7, 3

TABLE III

$TG(2^3, \Delta \leq 2)$;
$TG(2^{4-1}, \Delta = 2)$, I = ABCD;
(1) 1, 2, 4, 3, 7, 8, 5, 6
(2) 1, 2, 3, 4, 7, 6, 8, 5
(3) 1, 2, 8, 7, 4, 6, 5, 3
(4) 1, 4, 2, 6, 5, 7, 3, 8

TABLE IV

$TG(2^{5-2}, \Delta = 2)$, I = AB = CDE
(1) 1, 4, 3, 2, 6, 8, 5, 7
(2) 1, 3, 7, 8, 4, 2, 6, 5
(3) 1, 3, 2, 6, 7, 5, 8, 4

TABLE V

$TG(2^{5-2}, \Delta \leq 3)$ I = ABC = ADE
(1) 1, 3, 4, 2, 5, 7, 8, 6
(2) 1, 2, 8, 6, 3, 4, 5, 7
(3) 1, 5, 6, 2, 4, 8, 3, 7

TABLE VI

$TG(2^{6-3}, \Delta = 3)$ I = ABC = CDE = ADF
(1) 1, 2, 4, 3, 5, 6, 8, 7
(2) 1, 3, 5, 2, 8, 6, 4, 7
(3) 1, 7, 4, 2, 8, 3, 5, 6

TABLE VII

$TG(2^4, \Delta = 1)$															
$TG(2^{6-2}, \Delta = 2)$, I = ABC = DEF															
$TG(2^{7-3}, \Delta = 2)$, I = ABC = ADE = AFG															
(1)	1,	2,	4,	8,	6,	5,	7,	3,	11,	9,	10,	12,	16,	14,	13,	15
(2)	1,	2,	4,	8,	6,	5,	7,	3,	11,	9,	10,	14,	13,	15,	16,	12
(3)	1,	2,	4,	8,	6,	14,	10,	12,	16,	15,	11,	9,	13,	5,	7,	3
(4)	1,	2,	4,	8,	16,	12,	10,	14,	6,	5,	13,	15,	7,	3,	11,	9

TABLE VIII

$TG(2^4, \Delta \leq 2)$;															
$TG(2^{5-1}, \Delta = 2)$, I = ABCDE;															
(1)	1,	4,	2,	3,	5,	8,	6,	7,	11,	10,	12,	9,	13,	14,	16,	15
(2)	1,	2,	8,	5,	3,	4,	6,	7,	11,	16,	13,	10,	14,	9,	12,	15
(3)	1,	10,	16,	14,	5,	8,	12,	9,	2,	6,	13,	11,	7,	15,	3,	4
(4)	1,	4,	11,	15,	5,	7,	13,	10,	2,	12,	9,	3,	8,	14,	6,	16

TABLE IX

$TG(2^{6-2}, \Delta \leq 3)$, I = ABCD = CDEF															
(1)	1,	7,	4,	3,	6,	8,	2,	5,	13,	15,	9,	11,	14,	10,	12,	16
(2)	1,	12,	8,	2,	7,	16,	5,	3,	11,	13,	6,	14,	9,	4,	10,	15
(3)	1,	12,	15,	5,	7,	13,	9,	2,	4,	11,	14,	16,	6,	8,	3,	10

TABLE X

$TG(2^{7-3}, \Delta = 3)$, I = ABEG = ACFG = BCEF															
(1)	1,	7,	4,	2,	8,	6,	5,	3,	10,	16,	11,	9,	14,	13,	15,	12
(2)	1,	9,	5,	13,	8,	2,	10,	16,	14,	7,	4,	3,	11,	12,	15,	6
(3)	1,	13,	10,	2,	8,	12,	15,	6,	4,	16,	11,	3,	5,	9,	14,	7

TABLE XI

$TG(2^4, \Delta \leq 3)$															
$TG(2^{6-2}, \Delta \leq 4)$, I = ABCD = CDEF															
$TG(2^{7-3}, \Delta \leq 4)$, I = ABEG = ACFG = BCEF															
(1)	1,	8,	5,	4,	2,	7,	3,	6,	12,	15,	10,	13,	9,	16,	11,	14
(2)	1,	6,	9,	15,	4,	8,	10,	3,	7,	12,	2,	5,	16,	14,	11,	13
(3)	1,	12,	7,	4,	3,	6,	13,	16,	2,	5,	10,	15,	8,	11,	14,	9
(4)	1,	15,	16,	9,	7,	2,	3,	10,	4,	14,	11,	5,	6,	13,	12,	8

(Appendix IV - A, Sequences that minimize factor-level changes between adjacent trials, continued)

Table XII

TG(2^5, $\Delta = 1$)

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

Table XIII

TG(2^5, $\Delta \leq 2$);
TG(2^{6-1}, $\Delta = 2$), I = ABCDEF
TG(2^{7-2}, $\Delta \leq 3$), I = ABCD = DEFG

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

Table XIV

TG(2^5, $\Delta \leq 3$)

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

Table XV

TG(2^5, $\Delta \leq 4$
TG(2^{6-1}, $\Delta \leq 4$), I = ABCDEF
TG(2^{7-2}, $\Delta \leq 5$), I = ABCDE = DEFG

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

Table XVI

TG(2^6, $\Delta = 1$)

(1)		(2)		(3)	
1		1		1	
17		33		5	
25		49		6	
9	*	17	**	14	***
13	35	25	61	13	4
29	38	9	45	29	20
31	52	41	37	61	28
23	56	57	53	53	12
7	64	58	21	37	16
15	60	42	5	45	48
11	59	10	13	47	46
27	63	26	15	15	62
19	55	18	47	31	58
3	39	50	63	63	57
4	47	34	55	64	49
2	45	2	51	32	50
18	37	6	19	24	18
20	33	14	20	8	2
24	41	46	28	40	3
8	42	38	12	56	35
16	34	54	44	55	33
14	38	22	43	39	41
6	40	30	59	7	42
5	48	62	60	23	34
21	46	64	52	21	38
22	62	32	36	17	54
30	61	16	35	25	22
32	57	48	39	27	30
28	58	40	7	19	26
26	50	56	8	51	10
10	54	24	4	52	9
12	53	23	3	60	11
44	49	31	11	44	43
43	51	29	27	36	59
to		to		to	
*		**		***	

Table XVII

TG(2^6, $\Delta \leq 2$)
TG(2^{7-1}, $\Delta = 2$), I = ABCDEFG

(1)		(2)		(3)	
1		1		1	
2		41		9	
8		34		26	
16	*	2	**	42	***
24	45	42	19	57	23
7	39	58	31	17	21
31	40	26	55	2	7
28	64	50	52	50	47
18	58	18	49	58	46
10	41	20	59	18	64
6	37	12	25	24	30
4	38	60	28	40	29
20	46	51	11	39	31
27	61	53	9	51	27
9	54	23	3	60	59
14	56	39	17	20	19
22	63	15	57	12	35
32	59	32	33	8	33
12	52	48	45	16	49
3	44	64	5	32	52
19	60	63	38	48	4
25	48	27	14	56	10
11	47	43	29	55	25
15	35	44	62	37	61
29	33	16	56	41	62
30	50	10	8	11	14
21	53	6	36	43	6
13	49	30	4	3	22
5	55	46	24	5	54
23	51	13	40	13	34
17	36	61	47	45	36
26	34	54	7	15	38
57	42	22	35	63	44
62	43	21	37	53	28
to		to		to	
*		**		***	

Table XVIII

TG(2^6, $\Delta \leq 3$)

(1)		(2)		(3)	
1		1		1	
20		23		8	
27		32		22	
15	*	44	**	52	***
14	41	20	17	57	6
12	43	30	59	28	3
4	60	63	35	44	49
8	59	49	41	15	60
10	39	55	38	35	56
18	36	31	2	11	31
22	50	14	19	14	26
21	56	22	50	62	12
9	63	5	60	55	24
29	61	25	36	43	63
31	53	43	11	9	58
24	45	48	3	17	46
5	34	13	6	4	13
13	54	8	56	40	2
30	58	21	40	50	36
25	64	24	33	19	48
28	48	39	10	25	45
3	42	34	16	41	29
2	52	58	29	39	59
17	55	46	53	54	20
7	40	37	51	30	7
16	38	26	7	32	16
32	46	4	15	21	38
26	62	42	64	61	53
6	49	61	62	42	23
23	51	57	45	34	64
19	35	18	9	18	27
11	37	27	28	52	10
57	47	52	12	47	33
33	44	54	47	37	3
to		to		to	
*		**		***	

(From Tiahrt & Weeks, 1970)

APPENDIX IV - B

SEQUENCES THAT MINIMIZE THE NUMBER OF FACTOR-LEVEL CHANGES IN THE TOTAL EXPERIMENT WHILE ROBUST AGAINST TREND EFFECTS

Alternate arrangements (sequences) are given for ordering the experimental conditions in 8, 16, and 32 run designs. Each number in the sequence represents the position of each experimental condition when ordered in the "standard order". These sequences are discussed in Chapter VI.

All of the 8 run sequences were taken from Daper and Stoneman, (1968) while all of the 16 and 32 run sequence were taken from Dickerson (1974). All sequences are robust to linear trends only.

Eight experimental conditions

Alternate arrangements of runs of a 2^3 design

	Changes in Variable				Time Count for Variable		
Order of Runs	1	2	3	Total	1	2	3
14865732	3	4	2	9	−2	0	0
12873564	3	3	4	10	4	4	4
14856237	4	3	3	10	−4	4	4
12837564	3	3	4	10	4	4	6
14865237	4	3	3	10	−6	4	4
14856732	5	4	2	11	0	0	0
14687532	3	6	2	11	−2	0	0
14685732	3	6	2	11	−2	2	0
14865372	3	4	4	11	−2	0	2
14867532	3	6	2	11	−2	−2	0
18465732	3	4	4	11	−2	0	−2
14687352	3	4	4	11	−2	−2	2

Alternate arrangements of the runs of a 2^{4-1} design

	Changes in Variable					Time Count for Variable			
Order of Runs	1	2	3	7	Total	1	2	3	7
12873564	3	3	4	4	14	4	4	4	−4
12837564	3	3	4	4	14	4	4	6	−6
12783564	5	3	4	4	16	6	4	4	−2
12837654	5	3	4	4	16	2	4	6	−4
12873546	3	4	5	4	16	4	2	6	−4
18237564	3	5	4	4	16	4	2	4	−6
12873654	5	3	4	6	18	2	4	4	−2
18237654	5	5	4	4	18	2	2	4	−4
18273564	3	5	6	4	18	4	2	2	−4
12873645	4	4	5	5	18	0	2	6	0
18237645	4	6	5	3	18	0	0	6	−2
18273465	4	4	5	5	18	0	−2	6	0
12783546	5	4	5	4	18	6	2	6	−2
12738654	5	3	4	6	18	6	4	6	0
12836754	5	5	4	4	18	0	6	6	−4
18236574	5	5	4	4	18	0	6	4	−6
18237546	3	6	5	4	18	4	0	6	−6

Alternate arrangements of the runs of a 2^{5-2} design

	Changes in Variable						Time Count for Variable				
Order of Runs	1	2	3	4	5	Total	1	2	3	4	5
12486573	2	3	2	5	4	16	−8	8	8	−8	0
13726845	2	4	5	2	3	16	8	0	8	8	−8
13267584	3	5	2	2	5	17	8	8	8	8	−8
12738654	5	3	4	4	3	19	6	4	6	6	−4
17248653	2	5	4	5	4	20	0	2	4	−2	4
12784563	6	3	4	3	4	20	0	4	4	−4	4
17248635	2	6	5	4	3	20	0	0	6	0	2
12748563	6	3	4	3	4	20	0	4	6	−4	6
12748653	4	3	4	5	4	20	−2	4	6	−2	4
17248563	4	5	4	3	4	20	2	2	4	−4	6
12783654	7	3	4	4	3	21	4	4	4	4	−4
12784536	5	4	5	3	4	21	2	2	6	−4	4

(Appendix IV-B, Sequences that minimize the number of factor-level changes in the total experiment while robust against linear trends, continued)

Alternate arrangements of the runs of a 2^{6-3} design

	Changes in Variable							Time Count for Variable					
Order of Runs	1	2	3	4	5	6	Total	1	2	3	4	5	6
12457863	4	5	2	5	2	3	21	0	8	8	−8	8	−8
14268753	2	5	2	5	4	5	23	−8	6	8	−6	0	−6
12468753	2	5	2	7	4	3	23	−8	8	8	−4	0	−8
14268573	2	5	2	3	4	7	23	−8	8	8	−8	0	−4
12367854	7	5	2	2	5	3	24	4	8	8	8	−8	−8
14627853	4	5	4	3	6	3	25	−6	6	6	−4	0	−4
14627538	3	5	5	4	4	4	25	−2	8	8	−2	2	0
14257683	4	5	2	5	2	7	25	0	8	8	−8	8	−4
14627583	4	5	4	3	4	5	25	−4	8	6	−4	2	−2
14627835	4	4	5	4	5	3	25	−6	4	8	−2	−2	−4
14267583	4	5	2	3	4	7	25	−4	8	8	−4	4	−4
14628735	2	4	5	4	7	3	25	−8	4	8	−4	−4	−4
14267853	4	5	2	3	6	5	25	−6	6	8	−4	2	−6
14628753	2	5	4	5	6	3	25	−8	6	6	−6	−2	−4
12467853	4	5	2	5	6	3	25	−6	8	8	−2	2	−8
14672853	4	7	4	3	6	3	27	−4	4	4	−4	0	−4

Alternate arrangements of the runs of a 2^{7-4} design

	Changes in Variable								Time Count for Variable						
Order of Runs	1	2	3	4	5	6	7	Total	1	2	3	4	5	6	7
12358764	5	5	2	4	7	3	2	28	8	8	8	0	−4	−8	−8
12357648	3	5	3	4	6	4	3	28	10	10	10	4	0	−4	−2
12357846	3	4	3	5	6	3	4	28	10	6	10	0	0	−8	−6
12358674	5	5	2	4	5	5	2	28	6	10	8	0	−6	−6	−8
12357684	3	5	2	4	5	5	4	28	10	10	8	4	−2	−6	−4
12358647	4	5	3	5	5	4	2	28	4	10	10	−2	−6	−4	−8
12357864	3	5	2	6	5	3	4	28	10	8	8	2	−2	−8	−6
12358746	5	4	3	5	6	3	2	28	8	6	10	−2	−2	−8	−8
12356487	4	3	3	5	7	2	4	28	4	12	12	0	−4	0	−4
12356478	5	3	3	6	6	2	3	28	6	12	12	2	−2	0	−2
12356847	4	3	3	5	5	4	4	28	4	12	10	0	−6	−2	−6
12357486	3	4	3	5	4	5	4	28	10	6	12	0	2	−6	−4
12356748	5	3	3	4	6	4	3	28	8	12	10	4	−2	−2	−2
12357468	3	5	3	6	4	4	3	28	10	8	12	2	2	−4	−2
12356784	5	3	2	4	7	3	4	28	8	12	8	4	−4	−4	−4
12358467	4	5	3	3	7	4	2	28	4	8	12	−4	−4	−4	−8
12356874	5	3	2	6	5	3	4	28	6	12	8	2	−6	−4	−6
12358476	5	4	3	3	6	5	2	28	6	6	12	−4	−2	−6	−8
12354678	5	5	3	4	6	2	3	28	6	10	14	0	0	0	−2
12354687	4	5	3	5	5	2	4	28	4	10	14	−2	−2	0	−4
12354768	5	5	3	4	4	4	3	28	8	8	14	0	2	−2	−2
12354876	5	4	3	3	6	3	4	28	6	6	14	−4	0	−4	−6
12354786	5	4	3	5	4	3	4	28	8	6	14	−2	2	−4	−4
12354867	4	5	3	3	5	4	4	28	4	8	14	−4	−2	−2	−6
12345678	7	3	1	4	6	2	5	28	4	8	16	0	0	0	0
12345768	5	5	1	4	4	4	5	28	6	6	16	0	2	−2	0
12345687	6	3	1	5	5	2	6	28	2	8	16	−2	−2	0	−2

(Appendix IV-B, Sequences that minimize the number of factor-level changes in the total experiment while robust against linear trends, continued)

Sixteen experimental conditions

Alternate *Run Orders for 2^4 Designs Requiring 15 Factor Level Changes*

Design	Maximum Time Count	R^2	Run Order
A	16	.090	1 2 10 14 16 15 7 3 11 12 4 8 6 5 13 9
B	16	.141	1 2 10 14 16 15 7 3 4 8 6 5 13 9 11 12
C	24	.141	1 2 6 14 16 15 11 3 4 8 7 5 13 9 10 12
D	24	.141	1 2 6 14 16 15 11 3 7 5 13 9 10 12 4 8
E	24	.224	1 2 10 14 16 15 11 12 4 3 7 8 6 5 13 9
F	24	.224	1 2 10 14 16 12 11 15 7 3 4 8 6 5 13 9
G	28	.300	1 2 10 12 11 15 16 14 6 8 4 3 7 5 13 9

Thirty-two experimental conditions

Alternate *Run Orders for 2^5 Designs Requiring 31 Factor Level Changes*

Design	Max. Time Count	R^2	Run Order
A	36	.036	1 2 18 20 28 32 31 15 11 3 7 5 6 22 21 29 30 26 10 9 13 14 16 12 4 8 24 23 19 27 25 17
B	36	.040	1 2 18 20 28 32 31 15 11 3 7 5 6 22 21 29 30 26 10 9 13 14 16 12 4 8 24 23 19 17 25 27
C	36	.045	1 2 18 20 29 32 31 15 11 3 7 5 6 22 30 29 25 26 10 9 13 14 16 12 4 8 24 23 21 17 19 27
D	36	.045	1 2 18 20 28 32 31 15 11 3 7 5 6 22 30 26 25 29 13 9 10 14 16 12 4 8 24 23 21 17 19 27
E	40	.05[illegible]	1 2 18 20 28 32 31 15 16 14 6 22 24 23 19 27 11 3 7 5 13 9 25 17 21 29 30 26 10 12 4 8
F	44	.028	1 2 18 20 28 32 31 15 11 3 7 5 21 29 13 9 10 26 30 22 6 14 16 12 4 8 24 23 19 27 25 17
G	44	.030	1 2 18 20 28 32 31 15 11 3 7 5 21 29 13 9 10 26 30 22 6 14 16 12 4 8 24 23 19 17 25 27
H	44	.033	1 2 18 20 28 32 31 15 11 3 7 5 21 29 30 26 10 9 13 14 16 12 4 8 6 22 24 23 19 17 25 27
I	44	.033	1 2 18 20 28 32 31 15 11 3 7 5 21 29 30 26 10 9 13 14 16 12 4 8 6 22 24 23 19 27 25 17
J	44	.042	1 2 18 20 28 32 31 15 11 3 7 5 6 22 30 29 13 9 25 26 10 14 16 12 4 8 24 23 21 17 19 27
K	44	.049	1 2 18 20 28 32 31 15 16 8 4 3 7 5 13 9 25 17 21 29 30 26 10 14 6 22 24 23 19 27 11 12
L	44	.063	1 2 18 20 28 32 31 15 16 12 4 3 7 5 13 29 21 22 30 26 25 27 11 9 10 14 6 8 24 23 19 17
M	44	.066	1 2 18 20 28 32 31 15 11 3 7 5 6 22 30 26 10 9 25 29 13 14 16 12 4 8 24 23 21 17 19 27
N	44	.067	1 2 18 20 28 32 31 15 16 12 4 3 7 5 21 22 30 29 13 9 11 27 25 26 10 14 6 8 24 23 19 17
O	44	.070	1 2 18 20 28 32 31 15 16 14 13 29 25 27 19 3 7 5 6 8 4 12 11 9 10 26 30 22 24 23 21 17

DISTRIBUTION LIST FOR REVISED EDITION

Defense Documentation Center
Cameron Station
Alexandria, VA 22314

Education Research Information Center
Processing & Reference Facility
4833 Rugby Ave., Suite 303
Bethesda, MD 20014

NASA - Scientific & Technical Information Facility
P. O. Box 33
College Park, MD 20740

National Technical Information Services (NTIS)
Operations Division
5285 Port Royal Road
Springfield, VA 22151

Executive Editor
Psychological Abstracts
American Psychological Assoc.
1200 Seventeenth St., N. W.
Washington, D. C. 20036