Paper 3335
Experimental Approaches to Marketing and Pricing Research

Dawit Mulugeta, Joseph Thurman, Yuhan Jia, Maxx Friedauer, Jason Greenfield, Lisa Conley
Advanced Analytics, Pharmaceutical Distribution Strategic Pricing and Analytics, Cardinal Health, Dublin, Ohio 43017

ABSTRACT

Design of Experiments (DOE) is an essential component of laboratory, greenhouse and field research in the natural sciences. It has also been an integral part of scientific enquiry in such diverse social science fields such as education, psychology, marketing, pricing and social works. The principle and practices of DOE are among the oldest and the most advanced tools within the realm of statistics. DOE classification schemes, however, are diverse and at times confusing. In this presentation we will provide a simple conceptual classification framework in which experimental methods are grouped into classical and statistical approaches. The classical approach is further divided into pre-, quasi- and true-experiments. The statistical approach is also divided into one-, two-, and more than two factor experiments. Within these broad categories, we will review several contemporary and widely used designs and their applications. The optimal use of Base and Stat SAS to analyze, summarize and report these diverse designs will also be demonstrated. The prospects and challenges of such diverse and critically important analytics tools on business insight extraction in marketing and pricing research will be discussed.

INTRODUCTION

Experimentation is nothing but a quest for knowledge. It can be viewed as a component of the scientific method where a given hypothesis is tested to verify or refute previously held belief, compare and contrast various possibilities and establish new benchmarks in a systematic way. Past experiences and findings, observations and curiosity, theoretical considerations and the desire to push the frontiers of knowledge all contribute to the formulation of a hypothesis. An experiment is carried out when one or more independent variables are deliberately manipulated by the experimenter in a planned fashion and the effect on the dependent (response) variable is measured [9]. Design of Experiments is a branch of statistics that deals with experimentation and aims to determine the cause and effect relationships between or among variables. It is one of the major areas of course works in many academic establishments and is a subject of extensive analysis and reviews in many textbooks and articles both from theoretical and applied perspectives [4, 5, 9, 12, 13, 14, 16, 21].

DOE has also become central in almost every research and enquires including laboratory, greenhouse, field and natural experiments; small and large-scale studies; diverse fields of humanities and natural sciences as well as in small and big sciences

such the human genome project and the Hadron Collider particle accelerator studies that seek the God particles. In addition many successful firms use DOE to test their business strategies in the market; gauge the receptivity of their product and service offers; measure customer perceptions, loyalty and retention; identify new opportunities; evaluate the efficacy of their marketing programs including pricing, promotions, communication channels and productivity of sales force and business sectors [1, 6, 7, 9].

The purpose of this presentation is to provide a brief review of the contemporary experimental designs that are commonly used or can potentially be used for marketing research. The practical use of several designs in the behavioral as well as the natural sciences, their salient features, and the role that they play in our understanding of marketing dynamics are examined.

MARKETING RESEARCH

The basic principles and practices of DOE was first developed, consolidated and popularized by Ronald A. Fisher during his tenure at the Rothamsted experimental station while doing extensive field experiments on wheat in 1918 to 1933 [12]. His seminal paper of 1926 [11] outlined the three fundamental foundations of experiments that include the local control of field conditions for error reduction, replication as a means to estimate error and randomization for valid estimation of error variance.

Since Fisher's work, the principles and practices of DOE has become of great importance in wide ranges of applications including marketing research. Market research is a systematic and objective collection and analysis of data about a particular target market, competition, customer or business [9, 19, 20]. It begun in the United States in the late 1920's by Daniel Starch with the intent of evaluating the effectiveness of advertisements in the form of interviews and survey works [8]. Gradually such an approach was extended to include the gathering of customer response through telephones, call centers and eventually the web. Any of the five P marketing mixes (product, price, promotion, place and people) are potential targets of market research by a marketing research organizations, academic institutions and business firms. Experimentation in marketing research has become increasingly popular since the 1960's [8]. Market tests to establish the sales potential of new products are also becoming standard practices for many companies, and is being used more and more to determine the effectiveness of expected changes of any elements of the market mix [6, 9, 20].

Market experimentation is a costly venture. The research cost may include planning, designing, conducting, data collection, and analysis. Additional costs may include manufacturing, deployment and disposal of products, labor, logistics, and overheads. Therefore an in depth understanding of the available designs, ease of execution, and circumstances of use is helpful to increase the ROI of market research.

A typical marketing manager may conduct exploratory research such as examination of marketing literatures, site visits, meeting with customers or focus group studies where the interest lies in establishing a general picture or in clarifying concepts or in addressing business issues on a product or service that the firm offers [6]. The manager may also conduct descriptive research to measure market share, forecast profitability, evaluate the efficacy of promotions, and optimize inventories or prices. The marketing manager may also leverage transactional or panel data where the same or different repeated measurements are taken on customers, product, sales, etc. to conduct longitudinal descriptive studies [6, 9, 19, 20]. Cross-sectional studies where the interest lies on gauging customer perception on product offers, price and services in a form of survey are quite common. Such studies are a one time or recurrent undertakings where representative random samples are drawn from the population of interest. In such studies, quantitative and qualitative data are collected, tabulated often in a cross-classification table to draw conclusion [6].

Causal research establishes the cause and effect of relationships between or among a number of factors, and it has the following three distinct features: 1) Results are consistent across populations in direction and to a certain extent magnitude of change, 2) The dependent variable responses predictably upon manipulation of the independent variable, that is, there is no strong competing explanation for the outcome, and 3) there is a step-by-step mechanism leading from cause to effect, that is, the action precedes the outcome [1, 6, 7]. The marketing manager may be interested to determine the effect of price points of selected products on volume, sales and profit. He may also study the effects of types, frequency and channels of promotions on sales and market share. Such undertakings are experiments or invariably called tests. A given firm can conduct studies on many of the marketing programs including new product and service offers, store design and locations, sales force effectiveness, bundling of products, delivery frequency and routes, and other customer services.

While there are distinctions among exploratory, descriptive and causal researches, there is also a great deal of continuity and complementarity among each other. For instance if the frequency of customer attrition is increasing or customer traffic / visit is down, then the manger may begin with exploratory research by gathering facts to identify the main drivers. He may identify customer service or pricing among many marketing programs to be the culprits. Then he would conduct some form of descriptive study using historical data to measure metrics such as frequency and magnitude of price change on profitability, sales trend and customer's sensitivity to prices. Based on findings from such studies he may proceed to execute a well-designed price experiment in the market to measure the impact of price changes on several metrics of interest.

EXPERIMENTAL VALIDITY

Experimental validity refers to how a study is well executed in establishing a causal relationship between the dependent and independent factors as well as how well the

results can be generalized and applied in settings outside of the study. Proper planning and execution, proper definition of the population for which conclusions are drawn, minimization of interference from non-treatment factors, holding non-treatment factors constant, adjusting the results to remove the effects of non-treatment factors, the use of proper data collection, analysis and interpretation of results are among the known important steps that help achieve experimental validity [2, 4, 6, 9].

Experimental validity is closely tied with experimental error which is a measure of the variation that exists among observations on experimental units treated alike. Experimental units are things (customers, products, etc.) to which treatments are applied. Variation is the fodder of experimentation and is of two types: the inherent variability that exists in the experimental material and any variability attributed to the conduct of the experiment [20].

Experimental validities are of two types (Table 1). The first is internal validity and it refers to the execution of a test in such a way that results are attributed solely due to treatments [1, 5, 6, 7, 9, 20]. It is the extent to which a given study establishes the desired causal relationships between or among the variables of interest. An internally valid study is unbiased, has a good error control and responses of the dependent variable are to a large extent caused by the independent variable. The second is external validity which indicates the extent to which results are extrapolated to the real world [1, 5, 6, 7, 9, 20]. It is the measure of generalizability, relevance and value of experimental results outside of the study setup. Artificially controlled, constricted and highly manipulated laboratory or greenhouse studies have high internal validity whereas field and natural experiments are much better in external validity. External validity is enhanced when conditions in the setup of the study are similar to the real world.

Table 1. Sources of Internal and External Experimental Invalidity using Customers are an Experimental Units*.

Internal Validity	
History	External events that affect customer response.
Maturation	Change in customer through time (age, response).
Mortality	Customer dropping out of the study.
Testing	Pre-test measurement effect on post-test.
Selection Bias	Systemic differences in selection of experimental units.
Extreme Groups	Performance groups customers are from (high or low).
Instrumentation	Status of measurement device and procedure.
Contamination	Treatment protocol integrity.
External Validity	
Interaction	Nonrandom selection and its effect on treatments.
Hawthorne Effect	Subjects awareness that they are in a study.
Carryover Effect	Multiple treatments influence on customer response.
Pretest Effect	Increase sensitivity due to pre-exposure.

*Adopted from [1, 6, 7, 9] with some modifications.

Internal Validity

History: Refers to factors that affect the results of a study or the response of the dependent variable and are beyond the control of the investigator [1, 5, 6, 7, 9, 20]. These may include large scale or local events that occur during the course of the experiment such as inclement weather, holidays, political events, disasters, competition in the market, and seasonal changes.

Maturation: Refers to changes that occur through time within the test units that may include customers, products, services or markets due to effects other than treatments [6, 7, 9]. Such changes may include customer age, consumption pattern or purchase behavior, product lifecycle, market expansion or contraction. Maturation of test subjects may also include behavioral (boredom, neglect and inattention) and life style (income, education, social status) changes during the course of the study period.

Mortality: Experimental units may be lost during the course of an experiment. Customers may close their stores or move somewhere else, products may be pulled out of the market, test prices may be changed by personnel other than the experimenter. All these change the composition of experimental units thereby affecting results [6, 9].

Testing: Exposure to test protocols and pre-treatment measurement likely influences the outcome of subsequent measurements in studies consisting of customers [6, 7, 9, 19]. This is likely due to increased awareness of test objectives, increase sensitivity to repeated measurements, increased conditioning to treatments, and change in attitude attributed to the need to compare, standardize and calibrate responses across time.

Selection Bias: Refers to an arbitrary selection of test units based on convenience, prior knowledge or biased information so that the selected group fails to adequately represent the population [4, 6, 9, 12]. It reflects systemic error due to lack of randomization and proper sampling.

Extreme Groups: Extreme cases of observations may occur due to the intrinsic characteristics of experimental units. Such extreme observations tend to shift positions in that subjects with extreme values in earlier measurements tend to move towards the average value in subsequent measurements [1].

Contamination (Confounding): An extraneous variable correlated with the independent variable affecting the outcome of the experiment (response of the dependent variable) [12, 14]. It may also occur when the effect of one factor or treatment cannot be distinguished from that of another factor or treatment [16].

Instrumentation: Refers to changes in devices and experimental procedures that affect the response of the dependent variable during the course of the study [1, 7, 9]. These

may include changes in setting and calibration of mechanical, electrical and mobile devices; modification in survey methods that include questionnaire types and interview methods (person to person, phone, online); consistency in administering treatments, and experimental protocols including data gathering and management techniques through time.

External Validity

Interaction: This limits generalizability of experimental results and occurred when treatments are assigned to subjects selected non-randomly. When subjects are not randomly selected then their characteristics may bias their performance, and the study results may not be applicable to the population where subjects came from [9].

Hawthorne Effect: This refers to unintentional change in research subjects response due to increased self-awareness and special attention received from researchers rather than due to manipulation of independent variables [6, 9].

Carryover Effect: Multiple treatments applied successively tend to have limited generalizability simply because the earlier treatments may have variable influences on the outcome of the latter treatments. The difficulty of isolating the impacts of separate treatments render multiple and successive treatments less generalizable [9].

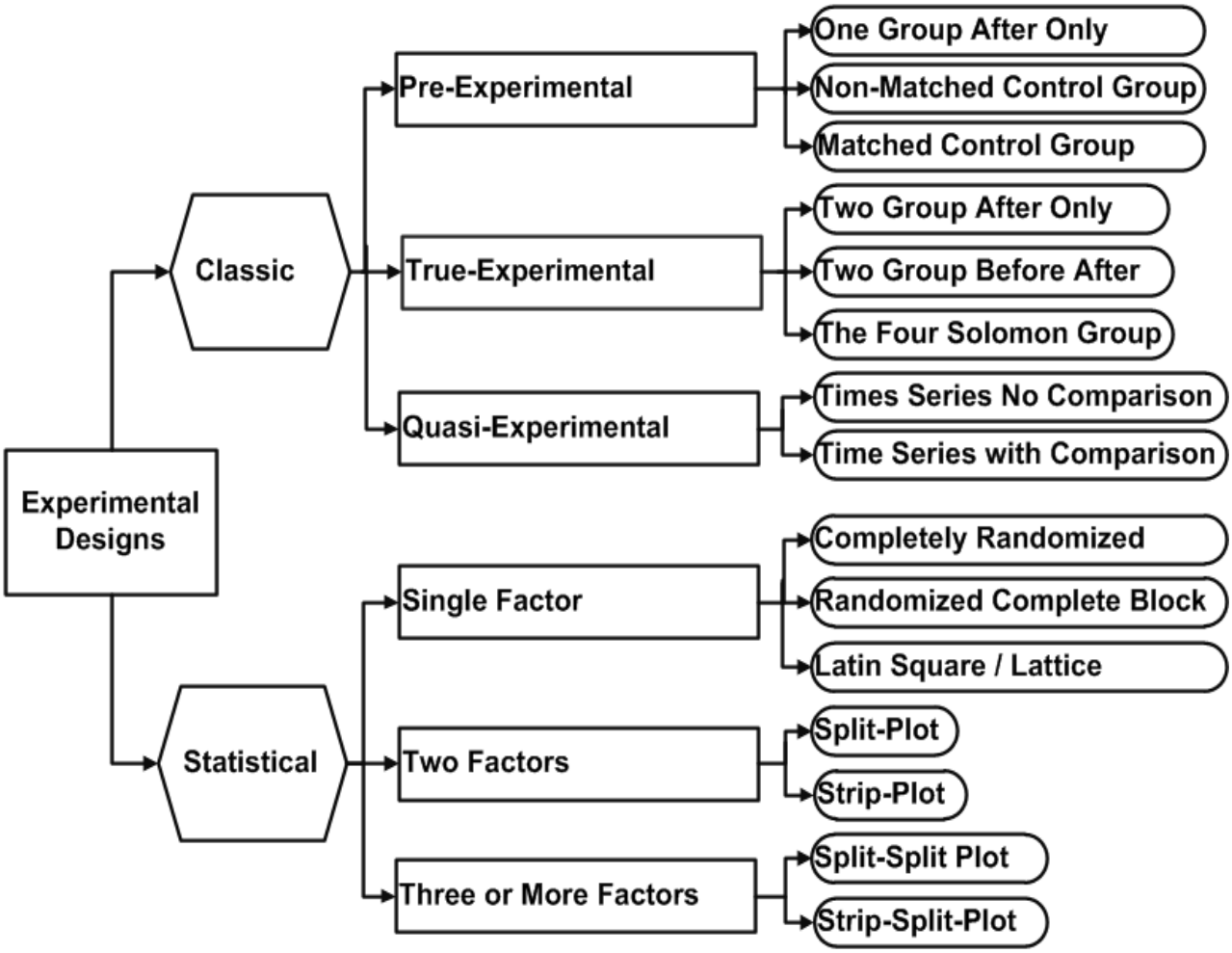
Pretesting Effect: Exposure of subjects to treatments during the pretest period influences the outcome of treatments applied latter. This is due to sensitization of subjects and their corresponding reaction to the subsequent treatment in the post period. This effect is as much a threat to external as it is to internal validity [1, 6. 9].

EXPERIMENTAL DESIGNS

A well planned and executed study aims to establish the cause-and-effect relationships between or among variables [2, 5, 9]. All experimental designs are not created equal. As shown in Figure 1, they can be classified into two broad categories called classical and statistical [1]. While the former considers only one treatment level of an independent variable at a time, the latter deals with measuring the impact of different treatment levels of one or more independent variables. The classic design consists of the pre-, the true- and the quasi-experimental designs [1, 5, 6, 7, 9]. The statistical design can be classified into one- two- and three or more factor experiments [12, 21].

The classic experimental designs have experimental groups and may or may not have a control group. True experiments, some of them belong to the classic design, must have control group, randomization, and manipulation of treatments. In all cases the independent variable is administered to the experimental group and not to the control group, and both groups are measured on the same dependent variable.

Figure 1. Types and Classification of Experimental Designs*.



*Adopted from [1, 12, 21] with some modifications.

Looking into the historical trend of the response variable prior and after application of a treatment is one of the best practices in marketing research. This will provide the experimenter an overview of the impact of factors on the targeted response variable other than treatments. In addition it will provide an unambiguous read on results. Figure 2 shows all possible trends that would happen prior and during the course of a study period.

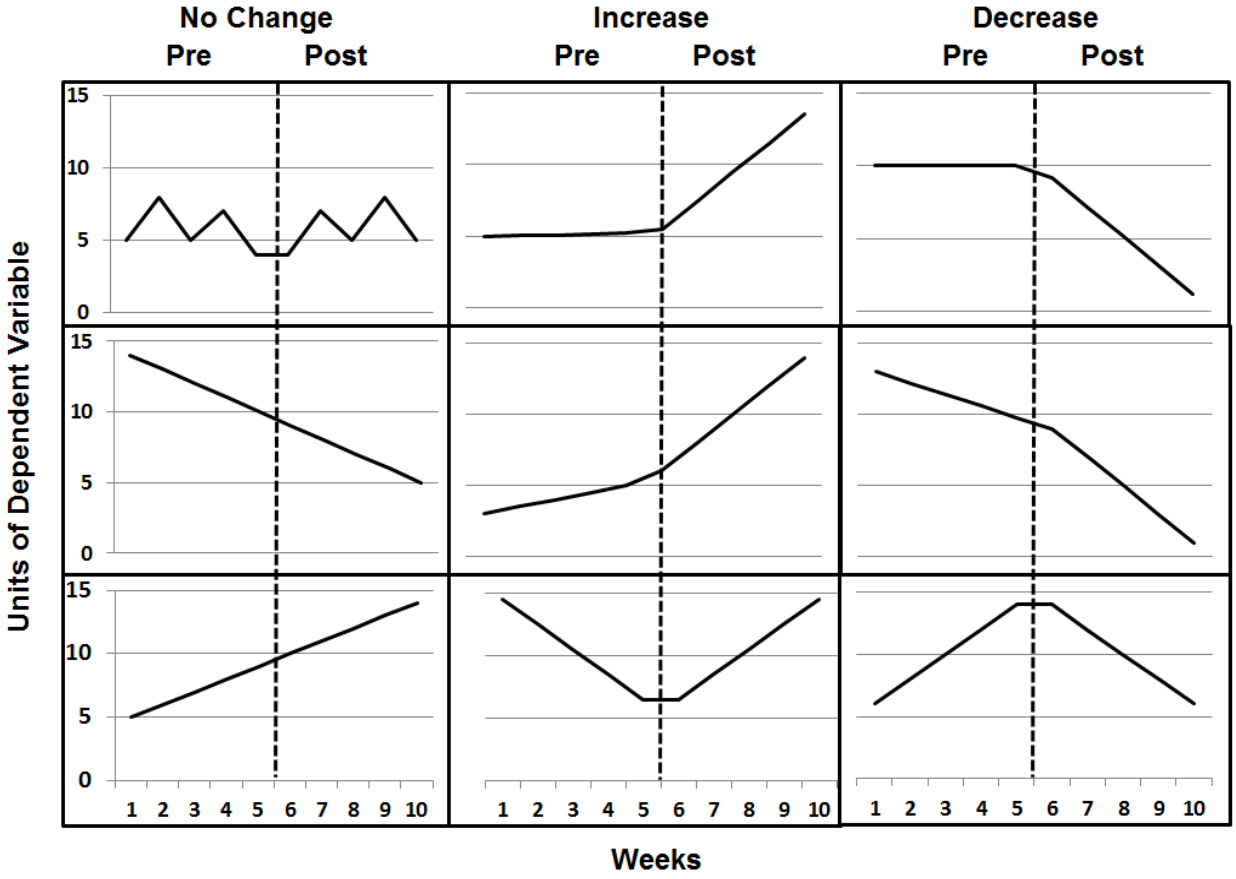
If a marketing manager is measuring the impact of promotion on volume of a specific product, three major outcomes are expected. Sales volume may remain the same, or it may increase or decrease. When volume remains the same the trend may be unchanged or is on a decreasing or increasing trajectory. However in the absence of the complete view of volume trend before and after promotion (treatment application), the latter two may provide a false read on the effectiveness of promotion on sales volume.

The other two scenarios relate to an increasing or decreasing effect after treatment application. When such post-treatment trends are viewed in conjunction with the pre-treatment period they offer a clear picture on the impact of the treatment (promotion) on

the response variable (sales volume). The effect will be either enhancing volume to an already flat or positive growth or completely reverse the downward trend. On the other hand if promotion contributes to volume reduction irrespective of the pre-treatment volume trend, other non-treatment history factors are the likely culprits.

Some of the classic experimental designs ("Two Group Before After" and "The Four Solomon Group) consider pre-treatment performance and trend of the dependent variable in their analysis and often are the most preferred choices for marketing and many of the social sciences research. The statistical experimental designs focus on the post-period only, that is, performance after application of treatments [4, 12, 21].

Figure 2. Changing Patterns of a Response Variable during a Pre- and a Post-Treatment Application.



Dotted vertical lines indicate time of treatment application.

CLASSIC EXPERIMENTAL DESIGNS

Pre-Experimental Designs

These groups of designs are exploratory by their very nature, lack the rigor that goes with an in-depth scientific enquiry, and have little or no control of external factors that influence the result of the experiment [5, 6, 9]. However they can be the stepping stone

to more elaborate studies and can help formulate hypothesis about causal relationships among factors of interest.

One Group After Only Design (One Shot Study): A single group of test units is exposed to a treatment as shown in Table 2 and its response is observed once [1, 5, 6]. The group is self-selected or arbitrarily assigned, and there is no randomization of experimental units. This design may be appropriate for exploratory research to establish some understanding of a marketing problem and may be useful for its simplicity, ease and low cost. However the total absence of control group or randomization renders it to be of almost no scientific value [1, 5, 6, 9]. A marketing manager may use transaction data to assess the impact of one time current or past event such as price change or promotion on immediate sales and profits. However almost all of the external factors are not controlled and hence the results of the study can be attributed both to the applied treatment and something else.

Table 2. Pre Experimental Designs Layout, Treatment and Measurement*.

Name	Randomization	Control	Pre-Period	Group	Treatment	Post-Period
One-Group After Only	x	x	x	EG	✓	O_1
Nonmatched Control Group	x	✓	x	EG	✓	O_1
				CG	x	O_2
Matched Control Group	x	✓	x	EG (Matched)	✓	O_1
				CG (Matched)	x	O_2
One-Group Before Afer	x	x	O_1	EG	✓	O_2

* x = Not applied; ✓ = Applied; EG = Experimental Group; CG = Control Group.
O_1, O_n = Observation / Measurement.

Non-Matched Control Group Design: Has test and control groups in which only the test group receives treatment (Table 2). No measurement of any kind is taken prior to treatment application. Experimental units are not randomized so there is a lack of representativeness and the assignment of groups is highly susceptible to selection bias [1, 5, 6, 7]. This design could be used for exploratory studies of many of the marketing mixes as long as the test and control groups are similar. One or several variables can be used to ensure that the two groups are similar. Instances of this kind of research may include the comparison of two markets following promotion to one, or the comparison of two groups of customers in which one receives price discounts of a product or enhanced awareness of current marketing activities. One of the major drawbacks of this design is that the control group is selected arbitrarily or such designation was made for convenience and ease. Such biased selection of the control group can contribute to erroneous conclusion.

Matched Control Group Design: This design is similar to the non-matched control group (Table 2) with the exception of having a control group that is matched with the

experimental group using one or several important criteria [1, 6]. This minimizes selection bias and improves upon the non-matched group in precision of the treatment impact measurement. The number of potential control groups that have a good match with the experimental group decreases as the number of matching criteria increases. If a marketing manager is interested to measure the effectiveness of product offers or promotions among several stores he can select manageable numbers of test and control stores following the matching process. He can use stores sales volume, sales trend, profitability, age, customer visit, geography and demography for matching purposes. As there may exist many matching criteria that exert variable influences on treatment performance, some form of prioritization in selecting the most relevant ones is necessary.

One Group Before After Design (One Group Pre-Test Post-Test): This is an upgrade of the one-shot study in that it adds a pre-test observation (Table 2). Unlike the preceding three designs, here a "before" measurement is introduced [5, 6, 9]. In essence this pre measurement is analogs to having a control group. History is uncontrolled. Promotion, customer purchasing behavior, market activities and myriads of factors may vary between the pre and the post periods. If sales or any variable of interest is on an increasing or decreasing trend results can over- or under-estimate the effect of the treatment, respectively. For this reason it is hard to study seasonal products or events using this design. This design is frequently used in clinical, education and pricing research. In addition to its simplicity like the one shot study, the manager will be able to compare measurements of the post with the pre periods through indexing or matched pair t test. Although the experimental effects cannot be generalized, this design can be helpful in cases when a control group is unavailable or is difficult to designate.

True Experimental Designs

Unlike the pre-experimental groups, the true experimental designs implement the good principles of experimentation, randomization, and hence are superior to pre experimental designs in drawing causal inferences [1, 5, 6, 7, 9, 14, 19, 20]. Randomization prevents selection and accidental bias in treatment assignments, balances the groups with respect to many known and unknown confounding variables thereby reducing systemic differences, and forms the basis for sound statistical tests with the basic assumption that treatments are exposed equally to extraneous variables. In addition all true experimental designs have control groups.

Two Group After Only Design (Post Randomized): This is similar to the static group design (Table 3) except that randomization of experimental units both for test and control group is introduced [1, 5, 6, 9]. As long as there are enough sample sizes, the random selection of test and control groups ensures that the two groups are matched well across many extraneous variables that influence the outcome of the study. This design is very popular and is commonly used in agricultural, laboratory, as well as in many scientific experiments. The goal is to measure one or more treatment effects in the post period of

the test group and compare the results with the control group of the same period. Most statistical experimental designs use this approach to measure post performances only without the need to have pre-period measurements. Pre-period measurements for many of the statistical designs may be impractical, costly, unnecessary and non-feasible.

Two Group Before After Design (Pre-Post Randomized Group): Extensively used in marketing research for the experimenters decides the selection and assignment of experimental units [1, 5, 6, 7, 9]. Assignment is not arbitrarily but rather randomly using one or several matching variables between the test and control units. Measurement of treatment effects is straight forward. Extraneous source of variables including history, maturation, mortality and measurement errors influence both the test and the control groups in about the same way (Table 3).

Table 3. True Experimental Designs Layout, Treatment and Measurement*.

Name	Randomization	Control	Pre-Period	Group	Treatment	Post-Period
Two Group After Only	✓	✓	x	EG	✓	O_1
				CG	x	O_2
Two Group Before After	✓	✓	O_1	EG	✓	O_3
			O_2	CG	x	O_4
Solomon Four Group	✓	✓	O_1	EG	✓	O_3
			O_2	CG	x	O_4
			x	EG	✓	O_5
				CG	x	O_6

* x = Not applied; ✓ = Applied; EG = Experimental Group; CG = Control Group.
O_1, O_n = Observation / Measurement.

The randomization process minimizes selection bias. Although this method is popular in marketing research, it suffers from the effect of sensitization in cases when experimental units are human subjects. Sensitization occurs when measurements on human subjects or customers of the control groups are made in the pre period in cases of studies that involve interviews and survey questionnaires. Pre-exposure and increased awareness of the subject can negatively influence the post-period responses of the test and control groups.

The Four Solomon's Group Design: This is hailed as the best experimental design in studies where humans are test and control subjects [1, 5, 6, 7, 9]. It is suitable to overcome the effect of sensitization that hampers the effectiveness of the pre-post randomized group design where humans are the focus of the study. This design is similar to the two group before-after design except that additional sets of test and control groups are added without any measurement in the pre-period so as to avoid any pre-exposure that results in sensitization (Table 3). This is popular design in sociology, education, psychology, but have had limited application in market research as the design

is not cost effective and require large number of samples consisting of two test and two control groups.

Quasi-Experimental Designs

In a quasi-experimental design the experimenter has some degree of control over extraneous factors but randomization may be impractical or beyond the control of the experimenter [1, 5, 6, 7, 9]. So the lack of randomization of experimental units poses a real threat to internal validity, and it may not be possible to clearly show causal relationship between the treatment and the observed outcome. In addition the experimenter may not have control, as in true experiments, when and to whom treatments are applied [6]. Natural experiments are a type of quasi-experimental designs where findings tend to be relevant to the real world. In a natural experiment, no variable is manipulated by the researcher, and no control group is designated or random assignment of a treatment exists. Such a design is useful to study the impact of large scale, sporadic, unusual, sudden or had to manipulate events such as disasters, accidents, earthquakes, and smoking. Overall quasi-experimental designs do not suffer the heavy maladies associated with external validity as much as artificial studies.

Time Series Design with and without Control: These are the most popular quasi-experimental designs. In this design a group of test units are observed successively and measurements of the same variable are taken at different time intervals prior and after treatment [1, 6] (Table 4). Panel data are the most suitable fodder for this design. The design is an important marketing research tool in many setups. Failure to control history is its major weakness.

Table 4. Time Series Design Layout, Treatment and Measurement.

Name	Randomization	Control	Pre-Period	Group	Treatment	Post-Period
Time Series with Control	x	x	O_1, O_2, O_n	EG	✓	O_1, O_2, O_n
	x	✓	O_1, O_2, O_n	CG	x	O_1, O_2, O_n
Time Series without Control	x	x	O_1, O_2, O_n	EG	✓	O_1, O_2, O_n

* x = Not applied; ✓ = Applied; EG = Experimental Group; CG = Control Group.
O_1, O_n = Observation / Measurement.

STATISTICAL EXPERIMENTAL DESIGNS

Statistical experimental designs are used to measure the effect of two or more treatments on the performance of the response variable which can be quantitative or qualitative. The partitioning and measurement of errors attributed to treatments and non-treatment factors using the analysis of variance (ANOVA) technique is one of the main strength and distinguishing feature that separates the statistical from the classic experimental designs. The main drawback of this design is that it focuses only on post treatment measurements without any consideration to pre-treatment performances or historical trends.

Design of an experiment has three components: estimation of an error, control of an error and proper interpretation of results [12, 21]. The difference among experimental units treated alike is called experimental error. This error is the primary basis for deciding whether an observed difference is real or due to chance. Every experiment must be designed to have a measure of experimental error. To measure experimental error replication is needed. In addition randomization of experimental units ensures that each units will have an equal chance of being assigned, measured, and exposed fairly equally to extraneous variables [13, 14]. A well designed experiment must minimize errors in order to increase the chance of detecting differences among treatments. Control of error is achieved through proper blocking technique where all treatments are exposed to similar experimental conditions. Blocking contributes to the reduction of experimental errors because within each block, experimental units are more similar [12, 13, 14, 21].

For each of the statistical experimental designs described below, the following there essential components are provided. 1) the experimental layout that shows how the randomization, blocking/replication and treatments are arranged, 2) the analysis of variance (ANOVA) which is an arithmetic process introduced by Sir Ronald Fisher to partitioning the total sum of squares into components associated with recognized sources of variation [21], and 3) a simple SAS code that would run the respective designs to generate ANOVA table and treatment means. Although SAS has extensive STAT procedures such as mixed, anova, glimmix, genmod, GEE, plan; and suites such as SAS JUMP and SAS QC that can be used to analyze experimental data [3, 18], we used the glm procedure to demonstrate the respective designs using SAS.

One Factor Experiments

Completely Randomized Design: Treatments are assigned completely at random so that each experimental unit has the same chance of receiving any one treatment (Table 5). This design is useful when the experimental units are essentially homogenous and the experimenter has full control of exogenous factors such as in laboratory experiments or in plant or animal experiments where environmental effects are much alike [12]. This design relies on randomization to control for the effects of extraneous variables. The experimenter assumes that, on the average, extraneous factors will affect treatment conditions equally; so any significant differences between conditions can fairly be attributed to the independent variable. Although this is a flexible design that can accommodates large number of treatments and replicates it is considered less efficient to measure variation among treatments due to the unrestricted nature of randomization and breadth of experimental errors (Table 5) that includes the entire variation except that due to treatments [21]. This design may have limited use for marketing research as it is difficult to attain homogenous experimental units in the absence of blocking to have good control of exogenous factors in the market.

Table 5. Completely Randomized Design with Four Replications and Six Treatments - Layout, ANOVA and SAS Procedure.

Layout*

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

ANOVA**

Source of Variation	df
Treatment (T)	T - 1 = 5
Error	Total df - T df = 18
Total	RT - 1 = 23

SAS Procedure

```
proc glm data=mydata ;
      class mytreatment ;
      model y = mytreatment ;
run ;
```

*Numbers within the layout box indicate treatments.
**df = degree of freedom; R=Replications.
Adopted from [12, 21] with some modifications.

Randomized Complete Block Design: This design is used when the experimental units can be meaningfully grouped into blocks or replications (Table 6).The purpose of blocking is to reduce experimental error by eliminating the contribution of known sources of variation among experimental units [12, 21]. This is done by grouping the experimental units into blocks such that variability within each block is minimized and variability among blocks is maximized (Table 6). While the intent of blocking is to attribute most experimental errors to variation within a block, proper blocking requires an in depth knowledge on the source, shape, orientation, gradient and direction of exogenous factors or source of variation. Precision of measurement in this design is superior to the completely randomized design because of blocking. In addition there is no restriction on the number of treatments that can be included in the blocks and it is easy to estimate missing data of experimental units in the event of mishaps [21].

Table 6. Randomized Complete Block Design with Four Replications and Six Treatments - Layout, ANOVA and SAS Procedure.

Layout*

Gradient

Rep1	1	2	5	6	3	4
Rep2	4	6	3	2	5	1
Rep3	5	4	1	3	2	6
Rep4	3	1	2	6	4	5

ANOVA**

Source of Variation	df
Replication (R)	R - 1 = 3
Treatment (T)	T - 1 = 5
Error	(R - 1)(T - 1) = 15
Total	RT - 1 = 23

SAS Procedure***

```
proc glm data=mydata ;
      class block mytreatment ;
      model y = block mytreatment ;
      lsmeans mytreatment / stderr pdiff ;
run ;
```

*Numbers within the layout box indicate treatments, Rep = Replications.
**df = degree of freedom.
***Blocks and treatments are fixed.
Adopted from [12, 21] with some modifications.

While this is one of the most widely used designs in agricultural research [12], it can potentially be used for many marketing researches. For instance if the pricing manager is interested to measure the impact of reducing a given product price by 2, 5, 7 and 10% on volume and profitability, he can assign these prices to randomly selected stores within

each of distinctly different three or four markets (blocks). He can also use other blocking factors such as store size to conduct the study within a single market. While treatment performance assessment for this design is made using the analysis of variance, it is also possible to pull a pre-treatment period data and do some analysis using the pre-post design for each market or store-size groups. One of the drawbacks of this design is that as the number of treatments increases, block size, heterogeneity within blocks and experimental error increases [12, 21].

Latin Square Design: This design is intended to simultaneously handle two known sources of variation among experimental units. The distinguishing feature of this design is that it uses two blocking factors commonly referred as row- and column-blocks [12, 21]. This arrangement is useful when clear gradients exist in opposite direction of the experimental site. In order to ensure this, every treatment occurs only once in each row-block and once in each column-block [12]. This procedure ensures the estimation of variation among row-blocks and among column-blocks and the removal of experimental errors. This design is popular in agricultural research from the 1920's to this date. The technique was introduced into marketing in 1953, and enjoyed a period of great influence and popularity until 1973, when it abruptly disappeared from the publications of the discipline. Careful investigation of the research record of this period revealed that its demise was due to increasingly poor application method that led to compromised results, combined with the emergence of full and fractional factorials with superior capabilities that occurred at approximately the same time [17].

Table 7. Latin Square Design with Five Replications and Five Treatments - Layout, ANOVA and SAS Procedure.

Columns*

Rows					
	1	2	3	4	5
	2	1	5	3	4
	3	4	1	5	2
	4	5	2	1	3
	5	3	4	2	1

ANOVA

Source of Variation	df*
Row (R)	T - 1 = 4
Column (C)	T - 1 = 4
Treatment (T)	T - 4 = 4
Error	(T-1)(T-2) = 12
Total	T^2 - 1 = 24

SAS Procedure

```
proc glm data=mydata ;
      class row column mytreatment ;
      model y = row column mytreatment ;
      lsmeans mytreatment / stderr pdiff ;
run ;
```

*Numbers within the layout box indicate treatments.
**df = Degree of freedom.
Adopted from [12, 21] with some modifications.

The execution of any marketing research using this design is practically difficult. Finding and designation of perpendicular gradients that help measure variation and reduce error will be a logistical nightmare. The chief disadvantage of this design is that the number of rows, columns and treatments must be equal [21]. If there are many treatments, the number of experimental units required becomes impractical. Latin square like the

randomized block design, suffer in that as the block size increases, the experimental error increases [21].

Lattice Design: This is one of the most popular incomplete block designs in agricultural research [12]. The design overcomes the problem of increasing experimental errors associated with an increase in the number of treatments and block size that are common in complete block designs. Blocks may not contain the complete list of treatments, and a reasonably small number of treatments can be maintained in each block even if the number of treatments is large. Although the level of precision can increase as block size decreases there will be an unequal degree of precision in comparison of treatment means [12, 21].

Table 8. Lattice Design with Three Blocks, Three Replications and Nine Treatments - Layout, ANOVA and SAS Procedure.

Layout*

	Replication 1			Replication 2			Replication 3		
Block 1	5	1	4	1	4	2	9	7	1
Block 2	8	3	2	6	9	8	5	6	4
Block 3	7	9	6	3	7	5	2	3	8

ANOVA

Source of Variation	df*
Replication	B = 3
Treatment	B^2 - 1= 8
Block (B)	B^2 - 1= 8
Error	(B - 1)(B^2 - 1) = 16
Total	B^2(B + 1) - 1 = 35

SAS Procedure

```
proc glm data=mydata ;
class row column mytreatment ;
model y = row column mytreatment ;
lsmeans mytreatment / stderr pdiff ;
run ;
```

*Numbers within the layout box indicate treatments.
**df = Degree of freedom.
Adopted from [12, 21] with some modifications.

Data analysis with lattice design can be complex, however, the level of precision associated with its use make it more preferred than complete block designs [12]. This design is divided into balanced and partially balanced types. The basic features include that the number of treatments has to be perfect square (9, 16, 15, etc.), the block size must be equal to the square root of the number of treatments, and the number of replications is one more than the block size. etc. This design will be difficult to leverage for marketing research especially if the marketing manager wants to include many levels of his choice of the marketing mix. In addition to the stringent requirements of the design on number of treatments, blocks and replications the complexity of data analysis and interpretation of results will render this design a less tempting choice for marketing research.

Two Factors Experiments

A multitude of factors influence the outcome of a study on any marketing mix. Price levels, competition intensity, number of substitute products, seasonality, promotion type and frequency all individually or in various combinations may influence product demand,

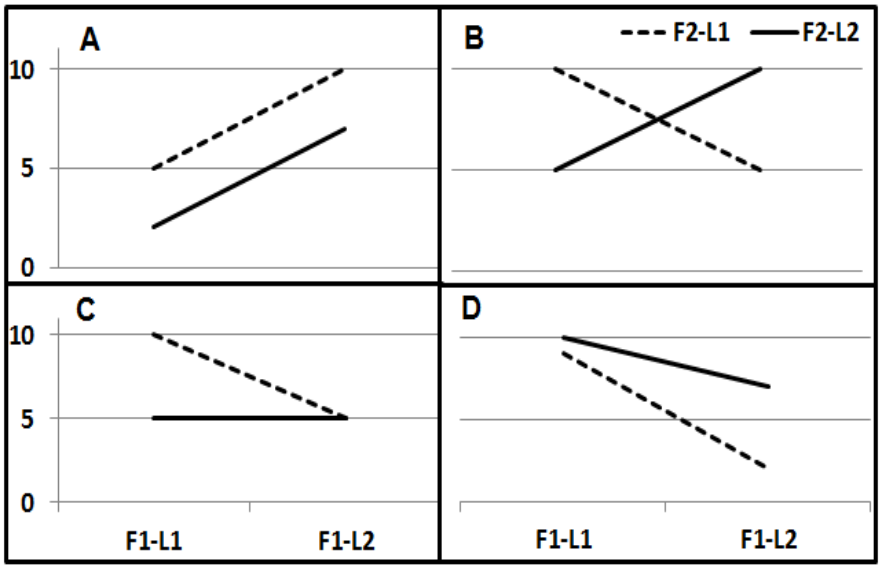
profitability, and customer satisfaction. Often single factor experiments are narrow in scope and fail to address the broader spectrum of issues that marketing managers faces. It is therefore beneficial to consider more than one factor when conducting a study. This is especially important when response to the factor of interest such as price points is expected to differ under different levels of the other factor such as promotion.

A factor is a discrete variable used to classify experimental units. Any of the marketing mixes such as price, promotion, etc. can be a factor. The treatments within the selected factors are termed as levels. In a pricing and promotion experiment for instance the price points (2%, 5%, 10% discounts) and promotion types (features, displays, newspaper inserts) are levels. The total number of treatments in a factorial experiment is the product of the levels in each factor. As the number of factors and levels increased the total number of treatments increases exponentially. In 2 by 2 and 3 by 3 factorials, the total number of treatments is 4 and 9, respectively. For large number of treatments a fractional factorial arrangement provides an option to reduce the number to a manageable size. This is accomplished through some business rules, established procedures and a combination of both [12].

An experiment that consists of all possible combinations of selected levels of two or more factors is said to have a factorial arrangement of treatments. Any of the complete block experimental designs for single factor experiments or the split-plot designs for two or more factor experiments may accommodate factorial arrangement of treatments using similar procedure of randomization and layout.

In a two or more factors study one or more than one way interactions may exist. Interactions occur when the responses of one factor levels differ under different levels of the other factor (Figure 3).

Figure 3. Effects of Treatment on the Response of Two Variables*.



*F = Factors; L = Level.

If the responses of the two levels of the second factors are the same in direction of change for levels 1 and 2 of the first factor (case A of Figure 3) there is no interaction. Case B is the extreme case of high interaction where levels 1 and 2 of the second factor responded in completely opposite direction to levels 1 and 2 of the second factor. There may also be varying levels of interactions between levels of the two factors (Cases C and D) that range in effect between the extreme case A and case B. Unlike the classical designs, statistical designs decipher the extent of interactions among the factors included in the study using the ANOVA method. However, as the number of factors included in a study increases the number of interactions increases resulting in the difficulty of result interpretation.

Split-Plot Design: This is a suitable design that accommodates factorial experiments and is a popular design in agricultural research where the relative influence of factors such as fertilizers, cultivars, pesticide, irrigation schemes, planting days all have differing impact on crop yield or other matrix of interest. Also called a multi-level design [15], it may incorporate one or more of the completely random, randomized complete block or Latin square designs [16, 21]. The distinguishing feature of this design is that one of the factors of study is assigned to the main-plot and the other factor is assigned to the sub-plot. The main-plot is divided into the sub-plot to which the second factor (the sub-plot) is assigned (Table 9). This way the main-plot becomes the block for the sub-plot treatments. In other words the smaller experimental units the sub-plots are nested within the larger ones the main plots [12].

Table 9. Split Plot Design with two main-plots, three sub-plots and three replications - Layout, ANOVA and SAS Procedure.

Layout*

A1	A2	A2	A1	A1	A2
B1	B3	B2	B3	B3	B1
B2	B1	B1	B2	B2	B3
B3	B2	B3	B1	B1	B2
Replication 1		Replication 2		Replication 3	

ANOVA

Source of Variation	df**
Replicatioins	R - 1 = 2
Main-Plot Factor (A)	A - 1 = 1
Error (A)	(R - 1)(A - 1) = 2
Sub-Plot Factor (B)	B - 1 = 2
AB	(A - 1)(B - 1) = 2
Error (B)	A (R - 1)(B - 1) = 8
Total	R*A*B - 1 = 17

SAS Procedure

```
proc glm data=mydata ;
      class R A B
      model y = R A R*A B A*B ;
      test=A e=R*A ;
run ;
```

*Numbers within the layout box indicate main-plot (A) and sub-plot (B) treatments.
**df = Degree of freedom.
Adopted from [12, 21] with some modifications.

This design is useful for two major reasons. First the experimenter is able to give differing emphasis and precision of measurements to the two factors according to perceived expectations. Differences among the main-plot factors are believed to be easier to discern whereas differences among sub-plot factor treatments require more

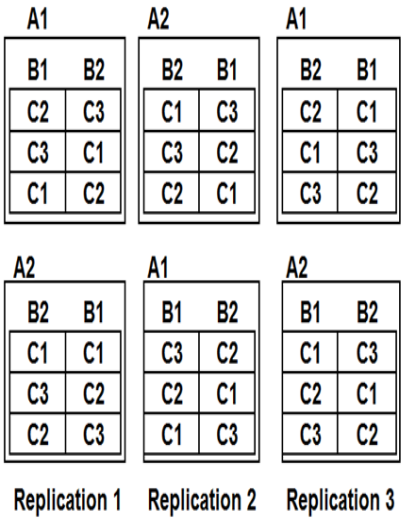
precision [12, 16, 21]. Second the design allow the experimenter to address size differences between the selected factors and the ease with which treatments of each factor are implemented and managed during the course of the study [12, 16]. The factor that is bigger in size is assigned to the main-plot while the smaller factor is allotted to the sub-plot. In an agricultural experiment where the interest lies in identifying the effect of tillage system and herbicides on crop yield, the investigator would assign tillage systems to the main-plot and herbicide rates to the sub-plot. In a hypothetical marketing research where the goal is to identity the best product price across markets, it would be appropriate to assign market to the main-plots and price points to the sub-plots. If the focus of the study is to measure the interactions of the two factors with better precision, the experimenter may choose to conduct the study using the strip-plot design instead of the split-plot [12]. The strip-plot design is less known outside of agricultural research and commonly viewed as a special case of split-plots. It requires special randomization and layout where the two factors under study are laid perpendicular to each other to create interaction plots.

Three or More Factor Experiments

Split-Split-Plot Design: This is an extension of the split-plot design to accommodate an additional factor (Table 10). It is most appropriate for a three-factor experiment where three different levels of precision are desired for the various effects.

Table 10. Split Plot Design with Two Main-Plots, Two Sub-Plots, Three Sub-Sub-Plots and Three Replications - Layout, ANOVA and SAS Procedure.

Layout*



ANOVA

Source of Variation	df**
Replicatioins	R - 1 = 2
Main-Plot Factor (A)	A - 1 = 1
Error (A)	(R - 1)(A - 1) = 2
Sub-Plot Factor (B)	B - 1 = 1
A*B	(A - 1)(B - 1) = 1
Error (B)	A(R - 1)(B - 1) = 4
Sub-Sub-Plot Factor (C)	C - 1 = 2
A*C	(A - 1)(C - 1) = 2
B*C	(B - 1)(C - 1) = 2
A*B*C	(A - 1)(B - 1)(C - 1) = 2
Error C	AB(R - 1)(C - 1) = 16
Total	RABC - 1 = 35

SAS Procedure

```
proc glm data=mydata ;
      class R A B C ;
      model y = R A R*A B
      A*B rep*A*B C A*C
      B*C A*B*C ;
      test=A e=R*A ;
      test=B e=R*A*B ;
      test=A*B e=R*A*B ;
run ;
```

*Numbers within the layout box indicate main-plot (A), sub-plot (B) and sub-sub-plot (C) treatments.
**df = Degree of freedom.
Adopted from [12, 21] with some modifications.

The main features of this design include that the three factors have differing sizes of experimental units. The largest, the medium and the lowest sizes are assigned to the

main-, sub- and sub-sub-plots, respectively [12]. Accordingly there are three levels of precision with the main-plot and the sub-sub-plot factors receiving the lowest and the highest degrees of precision, respectively. As more factors are accommodated analysis and interpretation of results are more complex especially if there are various levels of interactions among the three factors.

SUMMARY

The ultimate goals of market experimentation are to extract business insights that enhance profitability, market share, customer service, cost reduction, and device winning marketing strategy in the market. The marketing manager often manipulates one or more variables of interest to measure one or many responses of the marketing mix to establish a causal relationship. A given firm may conduct multitude of experiments on its marketing mix that include customers (perception, intent of purchase, satisfaction, loyalty, and attrition), products (launching, positioning, bundling), market (competition, segmentation, share, and expansion), promotion (messaging, outlets, types, duration) and pricing (strategies such as skimming and penetration, elasticity of demand, optimization, substitutes). Many choices of designs confer marketing managers to conduct diverse kinds of studies, learn their business better and draw actionable information for better decisions.

Increased precision through the reduction of random error is the goal of a well-designed and executed experiment. Thus designs that identify the source and quantify the magnitude of error are superior research tools. Multitude of factors influence the outcome of an experiment as the result establishing a causal relationship between or among marketing variables requires proper planning and selection of the right experimental design. Experimental designs vary greatly in handling internal and external validity, ease and cost of execution, measurement of errors and interactions, sophistication of analysis and interpretation of results. The presence of diverse options to design an experiment is a blessing to the marketing manager as it offers flexibility of choice to tackle myriads of marketing problems.

REFERENCES

[1] Aaker D. A, V. Kumar, and G. S. Day. 2001. Market Research. Seventh Edition. John Wiley & Sons Inc. New York. 751 pp.

[2] Anderson M. J and S. L. Kraber. 1999. Keys to Successful Designed Experiments. American Society for Quality Conference. 11:1-6.

[3] Bilenas J. V. 2006. Application of Experimental Design in Consumer Direct-Mail Solicitations. SAS Users Group International. Paper 196-31. SAS Institute Inc. Carry, North Carolina.

[4] Box G.E.P, W.G. Hunter, and J. S. Hunter. 1978. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. John Wiley and Sons Inc. New York. 653 pp.

[5] Campbell D. T. and J. C. Stanley. 1963. Experimental and Quasi-Experimental

Designs for Research. Houghton Mifflin Company. Boston. 88 pp.
[6] Churchill Jr. G. A. 1995. Marketing Research Methodological Foundations. Sixth Edition. The Dryden Press. Orlando. 1117 pp.
[7] Dillon W. R, T. J. Madden, and N. H. Firtle. 1987. Marketing Research in a Marketing Environment. Third Edition. IRWIN Inc. Burr Ridge. 760 pp.
[8] DJ Research Ltd. The History of Market Research. 2015. http://www.marketresearchworld.net/content/view/3754/49/
[9] Feinberg F, T. Kinnear and J. Taylor. 2013. Modern Marketing Research: Concepts, Methods, and Cases. Second Edition. Cengage Learning. 678 pp.
[10] Figard S. 2009. The Basics of Experimental Design for Multivariate Analysis. SAS Global Forum 2009. Paper 284-2009. SAS Institute Inc. Carry, North Carolina.
[11] Fisher R. A. 1926. The Arrangement of Field Experiments. Journal of the Ministry of Agriculture of Great Britain. 33: 503-513.
[12] Gomez K. A. and A. A Gomez. 1984. Statistical Procedures for Agricultural Research. John Wiley and Sons Inc. New York. 680 pp.
[13] Kuehl R. D. 2000. Design of Experiments: Statistical Principles of Research Design and Analysis. 2nd Edition. Duxbury Thomson Learning. Pacific Grove. 666 pp.
[14] Mason R. L., R. F. Gunst, and J. L. Hess. 2003. Statistical Design and Analysis of Experiments with Applications to Engineering and Science. Wiley & Sons. Hoboken. 746 pp.
[15] Milliken G. A. 2003. Multilevel Designs and Their Analyses. SAS Users Group International. Paper 263-28. SAS Institute Inc. Carry, North Carolina.
[16] Oehlert G. W. 2010. A First Course in Design and Analysis of Experiments. University of Minnesota. 679 pp.
[17] Hamlin R. P. 2005. The Rise and Fall of the Latin Square in Marketing: a Cautionary Tale, European J. Marketing. 39:328-350.
[18] SAS. 2015. Knowledge Base Documentation. https://support.sas.com/documentation/onlinedoc/stat/indexproc.html.
[19] Smith M. S. and G. S. Albaum.2005. Fundamentals of Marketing Research. Sage Publications. Thousand Oaks, California. 863 pp.
[20] Smith S. M. and G. S. Albaum. 2012. Basic Marketing Research: Volume 1, Handbook for Research Professionals. Qualtrics.153 pp.
[21] Steel R. G. D and J. H. Torrie. 1980. Principles and Procedures of Statistics: A Biometrical Approach. 2nd Edition. McGraw Hill Publishing Company. New York. 633 pp.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please feel free to contact the authors at:

Dawit Mulugeta, Ph. D.

Director of Advanced Analytics, Pricing Analytics Team
Cardinal Health, Dublin, OH,
E-mail: dawit.mulugeta@cardinalhealth.com

Joseph Thurman
Manager of Advanced Analytics, Pricing Analytics Team
Cardinal Health, Dublin, OH,
E-mail: joseph.thurman@cardinalhealth.com

Yuhan Jia Ph. D.
Consultant of Advanced Analytics, Pricing Analytics Team
Cardinal Health, Dublin, OH,
E-mail: yuhan.jia@cardinalhealth.com

Maxx Friedauer
VP Strategic Pricing
Cardinal Health, Dublin, Oh,
E-mail: max.friedauer@cardinalhealth.com

Jason Greenfield
Senior Consultant of Advanced Analytics, Pricing Analytics Team
Cardinal Health, Dublin, OH,
E-mail: jason.greenfield@cardinalhealth.com

Lisa Conley
Director of Advanced Analytics, Pricing Analytics Team
Cardinal Health, Dublin, OH,
E-mail: lisa.conley@cardinalhealth.com