EXPERIMENTATION SCIENCE: A PROCESS APPROACH FOR THE COMPLETE DESIGN OF AN EXPERIMENT

D. D. Kratzer

K. A. Ash

EXPERIMENTATION SCIENCE: A PROCESS APPROACH FOR THE COMPLETE
DESIGN OF AN EXPERIMENT.

D. D. Kratzer Ph.D., Pharmacia and Upjohn Inc., Kalamazoo MI,   and
K. A. Ash D.V.M., Ph.D., Town and Country Animal Hospital, Charlotte MI

ABSTRACT

Experimentation Science is introduced as a process through which the necessary steps of experimental design are all sufficiently addressed. Experimentation Science is defined as a nearly linear process of objective formulation, selection of experimentation unit and decision variable(s), deciding treatment, design and error structure, defining the randomization, statistical analyses and decision procedures, outlining quality control procedures for data collection, and finally analysis, presentation and interpretation of results. The protocol description form (PDF) is introduced as an instrument to guide the implementation and documentation of the Experimentation Science process.

INTRODUCTION

The objective of this paper is to document a process that has been evolving over the past 65 years. The process is part of the work agricultural statisticians do. We relate this process to the *so called* new paradigm of Total Quality Management (TQM) by giving a brief review of Total Quality Management, explaining Experimentation Science as a process of science, and a protocol description form (PDF) as an instrument to guide the documentation and implementation of the Experimentation Science process. An example of how Experimentation Science was applied in a real problem is planned for next year's (1997) conference.

TOTAL QUALITY MANAGEMENT REVIEW

The phrase: Total Quality Management is very much in vogue in management circles. This collection of words is not well understood and consequently often misused. Does it mean: "Management of Total-Quality"?, "Total Quality-of-management"?, "Quality of Total- Management" or "Total Management-of-Quality"? It is defined and explained by different "gurus" of management philosophies in different ways. The management philosophy we advocate is that of W. Edwards Deming. (1982).

Our reason for believing in the Deming philosophy is that it has foundations in the appropriate theories. Four theories are relevant. The theory of processes (or systems), the theory of psychology, the theory of knowledge and, most importantly to us, the theory of variation (or statistics). Deming recognized that, with regard to TQM, it is the word quality that is most misunderstood. To illustrate this, look in the business section of any phone book. For example, the Kansas City phone book shows 133 businesses that have quality as the first word in the name of their business. Many

more use the word some place in their names and even more use the word in their advertising. This misuse is world wide. The implication in using the word quality in this way is that the supplier is defining their product to be of highly accepted quality. Would they advertise otherwise? Deming understood and taught that it is only by the consumer, or customer, that quality is defined. Deming says it is the "voice" of the customer that establishes the quality of your product and it is only through scientifically sound surveys that the voice of the customer can be reliably determined.

From a statistical standpoint, the voice of the customer has a distribution with central tendency and variance. The center of the customer's distribution should become the target of the process. The processes (work) that produce the product also have a voice with central tendency and variance. A successful work process has a voice close, if not coincident, with the voice of the customer. The variation present in all things, will prevent most work outputs from being right on the target. For a graphical presentation of these concepts see W.W. Scherkenbach (1991).

Deming understood and taught that management has responsibility for determining the targets of the processes and management controls the sources of variation in the work processes. The goal of true Total-Quality-Management is to get the voice of the process coincident with the voice of the customer, and to continually manage all work processes to minimize process variation. Achieving this goal requires an understanding of process theory.

A process is simply a transformation of inputs into outputs, and the resources that go into the process are categorized in generic terms as: people, methods, materials, equipment, and environment. Each category involved in the process undergoes some transformation: people get trained, tired or burned out, methods get improved or worsened. Materials get used or wasted. Equipment is calibrated, wears out, fails or is replaced. The resources within these categories only appear to stay the same due to the coarseness of measurement. Every work process has some of these aspects, and within any organization there are many inter-linking work processes.

Processes get improved to produce higher quality product by moving the center of their distribution closer to the target and by reducing the variance around their center. This improvement is most effectively achieved by following the steps of the Deming Cycle. This aspect of process science has a sequence of steps that involves a feedback loop system to achieve continual improvement. The steps are Plan, Do, Study and Act (PDSA) . Deming's basis for these steps was his interaction with W. A. Shewhart (1939) who he invited to lecture at the Graduate School of the Department of Agriculture where Deming worked in 1938. You will recognize the similarity of the PDSA cycle with that of the scientific method's steps: hypothesize, experiment, analyze and decide (HEAD). It is probably not mere coincidence that Deming developed PDSA shortly after he returned from a years study with R. A. Fisher. Similarly the Scientific Method and Deming Cycle both have feed back loops that result in continual improvement. It is of interest to note that this feed back loop technique is also present in nature in the form of natural selection which continually improves species and learning which continually improves behavior.

Total Quality Management, then, is deciding relevant targets for work processes and minimizing variation around those targets. Those institutions that seek to understand and use the theory of processes, the theory of knowledge, the theory of psychology and the theory of statistics are the most likely to succeed.

EXPERIMENTATION SCIENCE DEFINITION

The definition of Experimentation Science was created in the 1970's in an attempt to create a term and definition that best described the work biostatisticians were doing in the Agricultural Biostatistics unit at The Upjohn Company. We found the term most often used to describe our work, "doing the statistics", to be uninformative and inappropriate.

Experimentation Science (ES) is the collection of knowledge about how to most efficiently produce data for effective and reliable decision making. It is a process of deciding and documenting relevant objectives and creating appropriate experimentation designs and decision processes while making the most efficient use of available resources. The product is a relevant set of data of sufficient quality to make documentable, reliable decisions.

Figure 1 shows where ES fits into the Scientific Method (SM). The SM is commonly thought of as a process within science which produces knowledge. Our attempts to find documentation or detailed descriptions of the Scientific Method have been futile. Also our experience, observation, and queries find little if any evidence of formal attention given to the teaching of the Scientific Method as a standard methodology especially in Veterinary Science, Animal Science or Statistical Sciences. This also has been motivation for us to formalize ES as a step in the SM process.

Figure 2 shows a flow of the necessary steps to be completed in the ES process. The sequence is completed with the output of a written report which should include the decisions which have been made from the data generated by the ES process.

EXPERIMENTATION SCIENCE ORIGINS

We do not claim to have discovered Experimentation Science because it has been evolving during the last century and has involved many people of greater contributions and stature than ourselves. Our task here is to describe and formalize Experimentation Science as a science.

Most of the steps of ES have been advocated for many years but not identified as such. Experimentation Science has its origins in agricultural experimentation. The work and writings of R.A. Fisher, D. R. Cox, W. G. Cochran, G. M. Cox , W. T. Federer, O. Kempthorne and many others have espoused the details of Experimentation Science in varied forms. Unfortunately the detailed steps were never assembled into a complete process. Consequently, many of the important principles the early authors discussed and advocated are overlooked in today's teaching and texts.

Table 1 shows the formation and evolution of the steps as the authors expressed the steps they saw as necessary for the experimentation process. Leonard and Clark (1939) delineated essential steps for exercising the Scientific Method in agronomy, Kempthorne (1952) outlines 9 steps of a statistically designed experiment. D. R. Cox (1958) suggested the planning of an experiment is divisible into two distinct parts each with several sub-parts. LeClerg et. al. (1962) restated Leonard and Clark's (1939) essentials of effective method in the empirical sciences. Anderson and McLean (1974) conclude that their ordered list of requirements are needed for scientific experimentation. Most recently, Hinklemann and Kempthorne (1994) provide a logical decision tree approach.

There are similarities and differences among the list of steps suggested by the authors in Table 1. Each list has strong points and each list has some omissions. The collection of lists does show the evolving science for experimentation but to be completely defined as a science there needs to be detailed and systematic steps and clearly specified general principles. The property that a science has just a few general principles is an essential requirement and is a prime motivator of our description of Experimentation Science.

PRINCIPLES OF EXPERIMENTATION SCIENCE
Experimentation Science has three governing principles: relevance, replication and randomization.

RELEVANCE
The principle of relevance deals with the extent to which the results and decisions obtained from the sample can be reliably extended to the population of interest. To achieve relevance requires collaborative efforts amongst the members of the research team. That team must involve subject matter experts and persons with statistical expertise. Relevance will vary from broad to intermediate to narrow inference needs. It will depend on the specific needs of the decision which is begging to be made.

Objectivity is a key ingredient to maintaining relevance. To achieve and maintain objectivity one must record the intent of the experiment and the procedures for collecting and analyzing the data in a formal document. This document called a protocol will be discussed in detail later.

When considering relevance, beware of popularity, convenience, precedence and opinion.

REPLICATION
Replication is the second principle of Experimentation Science. This principle includes considerations of replicability and reliability. Replication in experimentation is the repeating of exposure of experimentation units under the conditions of the experiment. Replication is generally accepted as highly desirable because human experience has shown that: "repetition is the mother of proficiency". Without replication, there is only one outcome generated from each treatment and there is no information to help decide whether the results are due to chance or due to reality.

More replication is always better, but there is a continuum of points of diminishing return. The challenge is to design an experiment with an efficient and effective amount of relevant replication. This challenge requires extensive collaborative activity between all scientists involved on the team.

The factors that affect the amount of necessary replication are contained in these questions: How much inherent variation exists among experimentation units?, How much difference or similarity needs to be detected among the treatments?, What is the cost of declaring a false positive?, and What is the cost of declaring a false negative? These questions do not have easy answers. The statistician is the primary resource on the research team for getting appropriate answers to these difficult questions.

Experimentation Science recognizes that the amount of replication is not the only consideration. It is also important to have relevant form and relevant frequency of replication. Since experimentation is for inference purposes, experimentation is in reality a sampling activity. Consequently the established principles of scientific sampling should be appropriately applied in creating experimentation designs.

Most texts discuss replication exclusively in terms of need to detect differences but, often the research question is whether two experimental groups are similar. For example; is a generic drug's efficacy comparable to that of the pioneer drug? This question comes under the category of equivalence testing. In general more replication is needed for reliable decisions in equivalence testing than in difference testing. Again it is the properly trained statistician that can sort out these issues.

RANDOMIZATION

Randomization is the third basic principle of Experimentation Science. It is a requirement that must be met before valid inferences can be made from experimental results.

Ostle (1954) gives appropriate emphasis to this critical point: "a statement **un**accompanied by a measure of reliability is of little value. If the devise of randomization is not employed, the measure of reliability will be biased and thus any tests of hypotheses can not be supported by an appropriate probability statement".

Cochran & Cox (1957) point out that randomization may be thought of as a form of insurance against systematic bias. This explanation is the common, and often exclusive, argument for randomization, but R.A. Fisher, the authoritative figure for randomization according to Kempthorne (1966), advocated over 50 years ago that one must also randomize to have a valid estimate of experimental error. This is the reason why any hypothesis test or interval estimate or variance estimate from a non-randomized trial can not be supported by a valid probability statement.

The mechanics of the randomization process are an initial extra effort and extra expense and consequently are often substituted with haphazard or systematic procedures which are subjectively judged to be unbiased. If the experiment is worth doing, it is worth doing so that it can be defended as scientifically valid even in a court of law if needed. This only occurs with use of a valid, **documentable** randomization.

Appropriate randomization requires preparation and planning but it is needed as insurance against unintentional bias **and** to have a valid estimate of error.

THE UNKNOWABLE CONSEQUENCES OF EXPERIMENTATION WITHOUT ES.

Research activities which do not employ all the principles of Experimentation Science can serve a role for creation of hypotheses, but such activities should not be substituted for appropriate experimentation because the result is data with un-measurable reliability. Such data generation results in an undeterminable plethora of false positive and false negatives, and in general, hampers reliable decision making, productivity and progress.

THE PROTOCOL DESCRIPTION FORM

When research is deemed necessary to make a decision, resources are allocated, inquiry and generation of data proceed. The quantity and quality of planning that precedes these steps vary tremendously. R.A. Fisher (1947), a founding father of statistics, said best what needs to be said about planning for experimentation: "experiments are carefully planned experiences". The plan for an experiment is the protocol. The major function of the protocol is to provide a formal descriptive document of the experimental design and procedures. The protocol also serves as a "contract for objectivity" for the decision making process. Due to the pre-data creation of the protocol, it controls the personal factor of human frailty that result in inevitable bias which was pointed out as a danger early on by Leonard and Clark (1939).

We advocate documenting the decisions made for the steps in the ES process in a written protocol. This documentation can be expedited using a Protocol Description Form (PDF). The sections of the sample PDF shown in figure 3 follow the flow of ES steps up to and including calculation of power and replication as shown in figure 2.

PROTOCOL DESCRIPTION FORM SECTION DEFINITIONS

The Experimentation Objectives are statements of the decisions to be made from the data.

The Population of Inference to which the decisions will apply is described. This is the scope of the experiment. The scope can vary from a population of "narrow" inference such as to a single location

or single group of animals, or to a population of "broad" inference such as the international market. Also, there can be many "intermediate" inference spaces. Defining the population of inference is critical for subsequent decisions about design and analysis. This terminology of broad, intermediate

and narrow is accredited to McLean et. al. (1991)

The <u>Experimentation Units</u> which will be subjected to the conditions of the experiment are described. Experimentation units are models for the entities to which treatments will eventually be applied in the population of inference. An example is a pen of 4 animals fed an experimental ration used to model a feedlot of 150 animals in the inference population.

<u>Response Variables</u> are listed and subdivided based on the roles they will serve. <u>Decision Variables</u> are expected to reflect the effect of treatments and are used to make the decisions specified in the objectives. <u>Ancillary Variables</u> are secondary to the objectives and serve a monitoring, management or regulatory role.

The <u>Treatment Structure</u> specifies the category of design that will be used for the experimentation groups; for example, one-way dose response.

The <u>Experimentation Groups</u> section lists and describes the treatments which are to be assigned the experimentation units. This is a very important section because statistical comparisons among these groups will be the basis for decision making.

The <u>Design Structure</u> is a descriptor for the basic statistical design which will be used to assign experimentation units to experimentation groups. The designs are: Completely Randomized, Randomized Complete Blocks, and Latin Square, or some combinations of these three basic designs .

The <u>Randomization Procedure</u> specifies the formal process by which the treatments will be assigned to the experimentation units.

The <u>Statistical Analysis</u> outlines the methodology to be applied to the response variables.

The <u>Decision Process</u> outlines how the results of the statistical analysis will be used to make the desired decisions.

The <u>Power of Test</u> is a statistical theory based quantitative estimate of the sensitivity of the experiment. Usually power is controlled by the <u>Number of Replications</u>. True replication of a treatment effect occurs with each additional, autonomous experimentation unit. The statistician is the primary resource for appropriately completing this section of the protocol. The statistician serves by answering the often asked question of: "How Many Do I Need (HMDIN)"?

Summary

Agricultural statisticians have been evolving the process of Experimentation Science for the past 65 years. Here we have put together the details to document the process. Those that follow this

Experimentation Science process will be valuable assets to institutions truly practicing Total Quality Management (TQM). If we are collaborating with scientists who have good abilities in the four z's: ability to **summarize** former work, ability to **organize** existing knowledge, ability to **analyze** the problem and it's scope, and ability to **hypothesize** alternatives with details on what they predict, then relevant questions will be formulated. If these abilities are then used to construct a Protocol Description Form (PDF), the experimentation process will proceed objectively, efficiently, and effectively.

Epilogue

It is the well documented protocol that separates "an experiment" from "an experience".

REFERENCES

Anderson V. L. and McLean R. A. 1974. **Design of Experiments: a Realistic Approach.** Marcel Dekker, Inc. New York, NY.

Cochran, G. M. and D. R. Cox . 1957. Second Edition (1950) . **Experimental Designs** John Wiley & Sons, Inc. New York, NY.

Cox, D. R. 1958. **Planning of Experiments.** John Wiley & Sons , Inc. New York

Deming, W. Edwards.1982. **Out of the Crisis**. MIT Center for Advanced Engineering Study. Cambridge , Mass.

Fisher R .A. 1947. **The Design of Experiments.** 4th Ed. Oliver and Boyd, Edinburgh.

Hinklemann, K and O. Kempthorne. 1994. **Design and Analysis of Experiments . Volume I . Introduction to Experimental Design** . John Wiley & Sons, New York NY.

Kempthorne, Oscar. 1952. **The Design and Analysis of Experiments**. John Wiley & Sons, Inc. New York NY.

Kempthorne, Oscar. 1966. **Some Aspects of Experimental Inference**. Journal of the American Statistical Association 61: 11-34.

LeClerg, Erwin L., Warren H. Leonard and Andrew G. Clark. 1962. **Field Plot Technique.** Second Edition. Burgess Publishing Company. Minneapolis, MN.

Leonard, W. H. and A. G. Clark. 1939. **Field Plot Techniques.** Burgess Publishing Co. Minneapolis, MN.

McLean, R ., W. L. Sanders and W. W. Stroup. 1991. **A Unified Approach to Mixed Linear Models**. The American Statistician 45: 54-64.

Ostle, Bernard. 1954. **Statistics in Research. Basic Concepts and Techniques for Research Workers.** Iowa State College Press. Ames, Iowa.

Sherkenbach , William W. 1991. **Demings Road to Continual Improvement.** SPC Press, Knoxville, Tennessee.

Shewart , Walter A. 1939. **Statistical Method from the Viewpoint of Quality Control**. Dover Publications Inc., East 2nd Street, Mineola, 31 N.Y. 11501.

TABLE 1. Evolution of the Definition of the Steps in the Process of Experimentation.

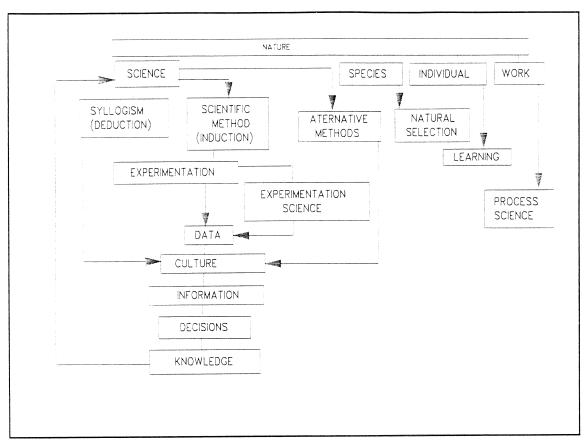| Author / Step	Leonard and Clark (1939):	Kempthorne (1952)	LeClerg et. al. (1960)	Anderson and Mclean (1974)	Hinkleman and Kempthorne (1994)
1	Formulate Hypothesis	State the problem	Clearly state the problem	Recognize the problem	Formulate questions and hypotheses
2	Logically Analyze the Problem	Formulate the hypotheses	Formulate the hypothesis	Formulate the problem	Formulate in subject matter terms
3	Use of Deductive Process to Design and Outline The Experiment and Quantitate	Devise the experimental technique and design	Logically analyze the hypothesis	Agree on factors and levels	Translate into statistical model
4	Control the personal equation (?)	Examine outcomes relative to the objectives	Outline the details of the experiment	Specify the response variables	Develop statistical design
5	Perform rigorous and exact experimental procedure	Consider outcomes relative to proposed statistical procedures to insure validity	Define the inference space	Define the inference space	Perform experiment
6	Sound and Logical Presentation	Perform experiment	Rigorously conduct experiment	Randomly select the experimental units	Collect data
7	Complete and Careful Report	Apply statistical techniques	Analyze data	Assign treatments to experimental units	Perform statistical analysis
8		Draw conclusions with measures of reliability relative to population of inference.	Precisely state the conclusions	Outline analysis	Interpret results
9		Evaluate the investigation relative to similar studies	Completely report data and methods	Collect data	Formulate new hypotheses
10				Analyze data	
11				Conclude	
12				Implement	

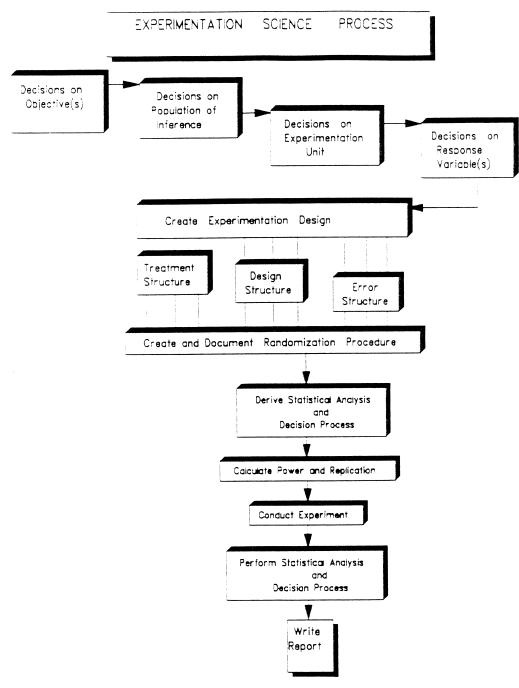Figure 1. Naturally Occuring Processes which have Feedback Loops for Achieving Continuous Improvement.

Figure 2. Experimentation Science Process Flow Chart.

EXPERIMENTATION SCIENCE PROCESS

Decisions on Objective(s) → Decisions on Population of inference → Decisions on Experimentation Unit → Decisions on Response Variable(s)

Create Experimentation Design

Treatment Structure — Design Structure — Error Structure

Create and Document Randomization Procedure

Derive Statistical Analysis and Decision Process

Calculate Power and Replication

Conduct Experiment

Perform Statistical Analysis and Decision Process

Write Report

Figure 3. Sample protocol description form.

Upjohn PROTOCOL DESCRIPTION FORM

Date: ___/____/___

Project No. _____   Experiment No. _____

EXPERIMENTATION OBJECTIVE(S):

POPULATION OF INFERENCE:

EXPERIMENTATION UNIT:

RESPONSE VARIABLES

Decision (In Order of Importance)

Ancillary

TREATMENT STRUCTURE

Classification

Experimentation Groups

DESIGN STRUCTURE:

RANDOMIZATION:

STATISTICAL MODEL AND DECISION PROCESS:

POWER AND REPLICATION: