مؤسسة حمد الطبية
Hamad Medical Corporation
HEALTH • EDUCATION • RESEARCH صحة • تعليم • بحوث

In Collaboration with
Institute *for* Healthcare Improvement

Middle East Forum on Quality & Safety in Healthcare **2023**

16-19 March, Doha

An Introduction to Planned Experimentation

Saturday, 14th March (15:30-16:30)

Healthcare Resilience in Extraordinary Times

Brought to you by:
Hamad Healthcare Quality Institute

IHI Faculty

Robert Lloyd, PhD
Vice President Improvement Science
Sr. Improvement Advisor

Conflict of Interest

The speaker for this session, Dr. Robert Lloyd, has no conflict of interest or disclosure in relation to this presentation.

Learning Objectives

At the end of this session, participants will be able to:

1. **Explain why Planned Experimentations is critical to Quality Improvement efforts**
2. **Describe the history of Planned Experimentation**
3. **Describe various types of studies and experiments**
4. **Explain the Planned Experimentation Terminology**
5. **Describe the principles and tools of Planned Experimentation**

Why Planned Experimentation?

Is life this simple?

X → Y

Patient encounter with physician

A healthy and satisfied patient

(If only it was this simple!)

The Messiness of Life!

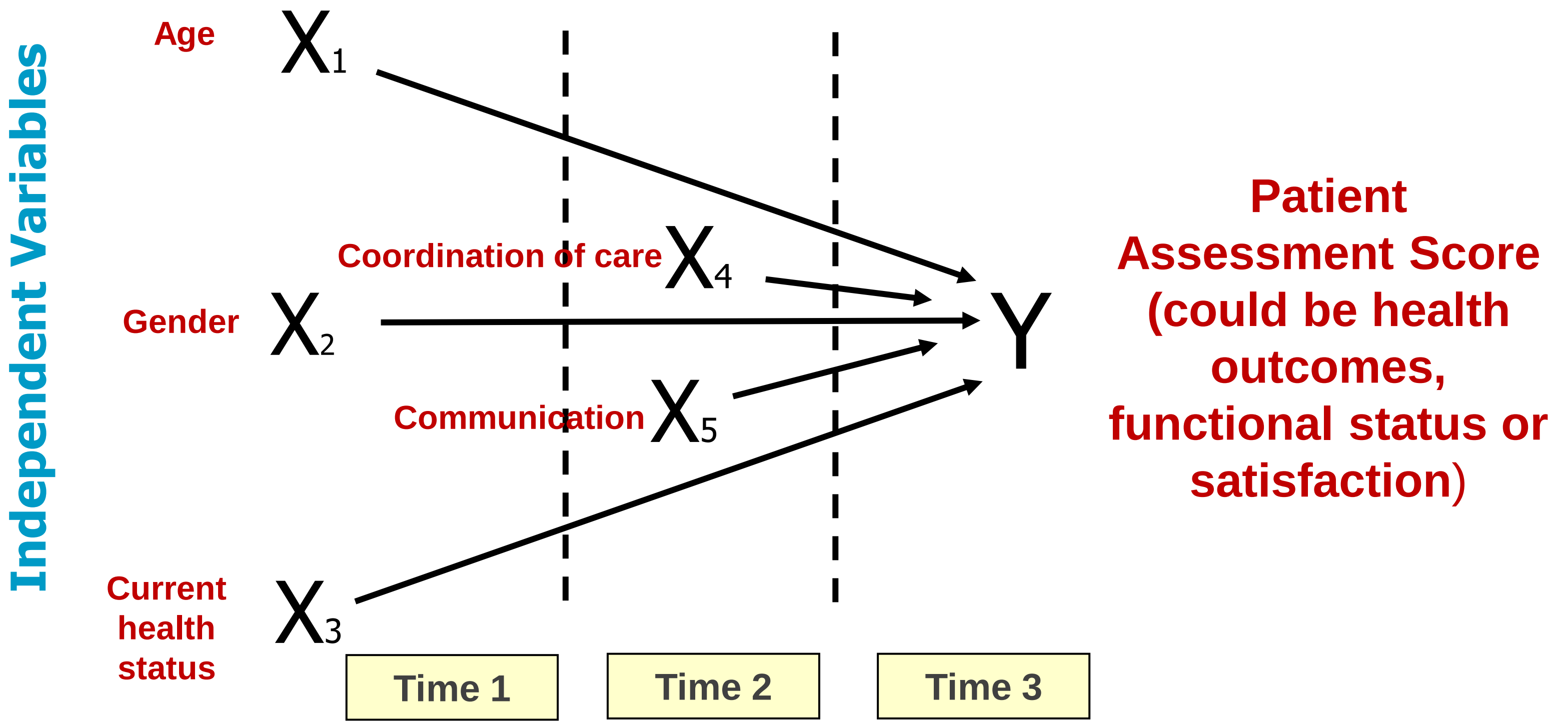
"Some problems are so complex that you have to be highly intelligent and well informed just to be undecided about them."

--Laurence J. Peter

A good reference on this topic is "Wicked Problems and Social Complexity"
by Jeff Conklin, Ph.D., Chapter 1 in *Dialogue Mapping: Defragmenting Projects through Shared Understanding*. For more information see the CogNexus Institute website at http://cognexus.org, 2004.

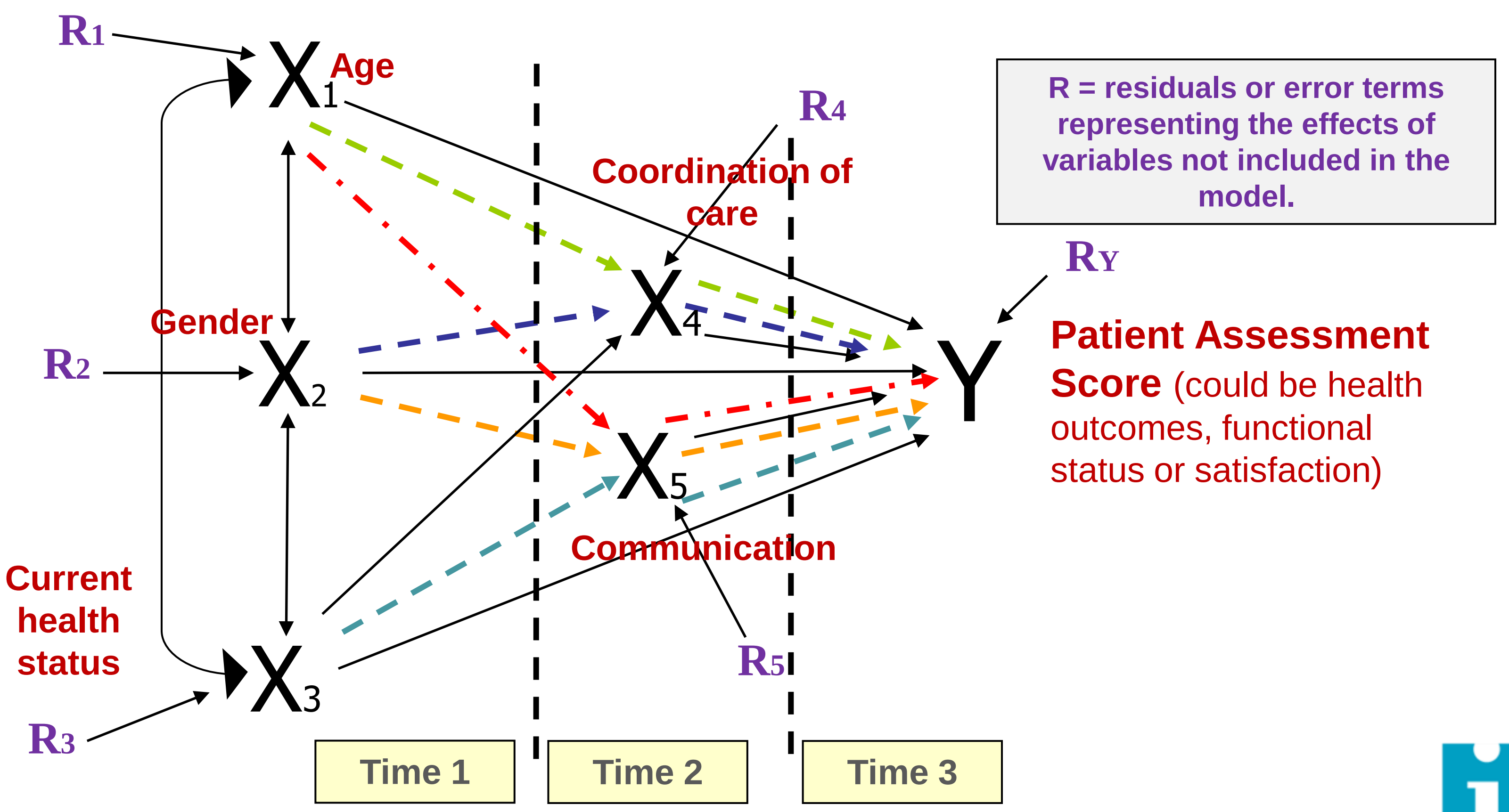
Life looks more like this…

There are numerous <u>direct effects</u> between the independent variables (the Xs) and the dependent variable (Y).



Actually life looks like this…

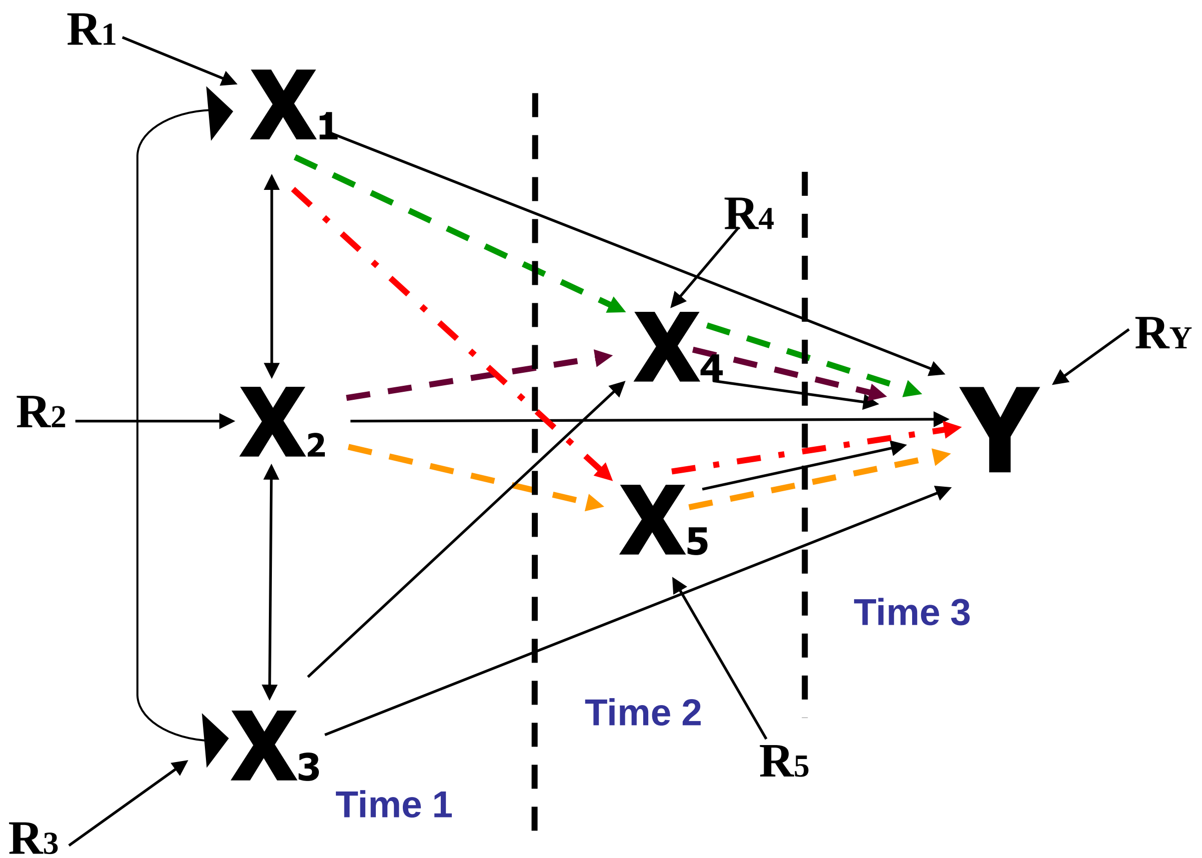
In this case, there are numerous direct and indirect effects between the independent variables and the dependent variable. For example, X1 and X4 both have direct effects on Y plus there is an indirect effect due to the interaction of X1 and X4 conjointly on Y.



Dr. Walter Shewhart on Applied Science

*"Both pure and applied science have gradually pushed further and further the requirements for accuracy and precision. However, **applied science**, is even more exacting than pure science in certain matters of accuracy and precision."*

Planned Experiments can help you understand the messiness of life!



I <u>REALLY</u>
do enjoy
the
messiness
of life!
Don't you?

Dialogue
Assessing the Messiness of Life!

- Do people within your organization regularly view issues as being rather messy and complex or do they see them as simple problems that should be resolved quickly and easily?

- List a few of these messy problems and why they are this way.

- On a scale of 1-10, how messy are each of these problems? (1 = a simple problem that is not very messy to 10 = a very messy problem)

- Do you have measures for these messy problems that allow you to determine just how complex and challenging each problem is?

- If you are measuring, do you feel that the measures you have are valid, reliable and appropriate given the complexity of the issues you face each day?

OK, enough of this messy talk. Let's start untangling this stuff!

Improvement requires two types of Knowledge

Subject Matter Knowledge

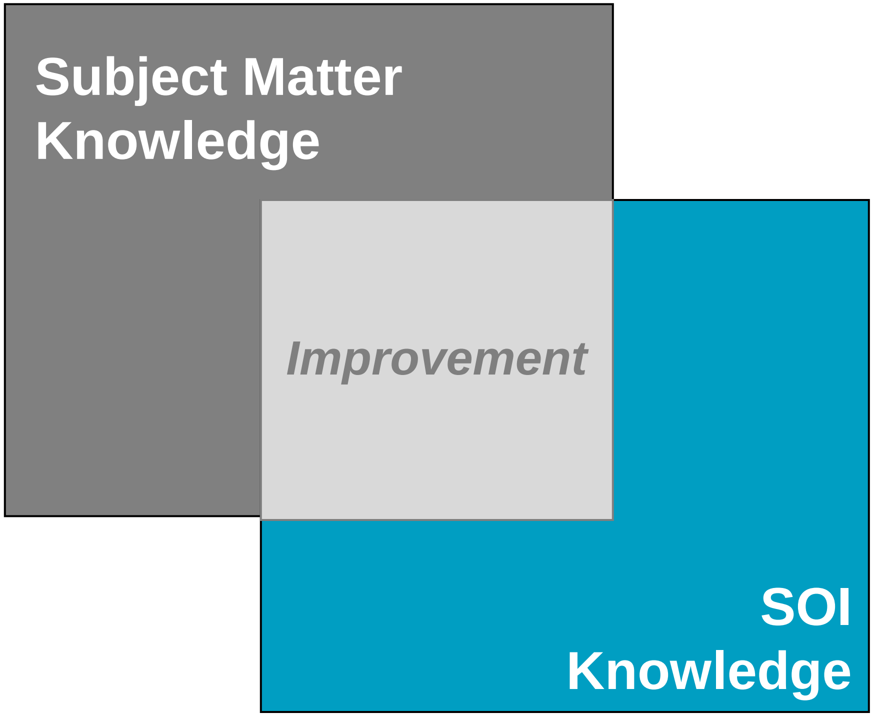
Subject Matter Knowledge: Knowledge basic to the things we do in life. Professional knowledge. Knowledge of work processes.

Science of Improvement (SOI) Knowledge: The interplay of the theories of systems, variation, knowledge, and psychology.

SOI Knowledge

Knowledge for Improvement

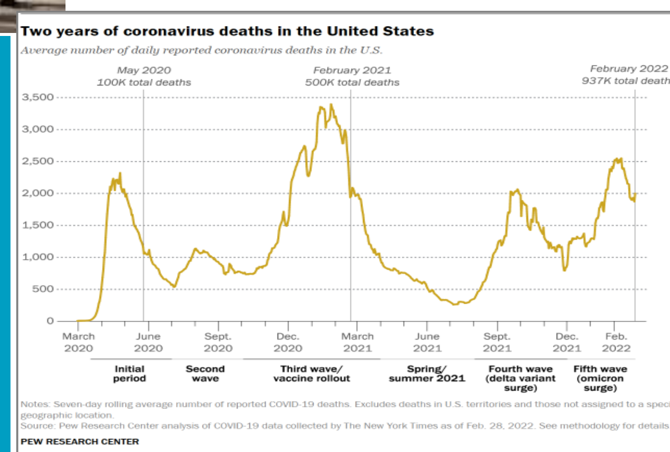
Improvement: Learn to combine subject matter knowledge and SOI knowledge in creative ways to develop effective changes for improvement.

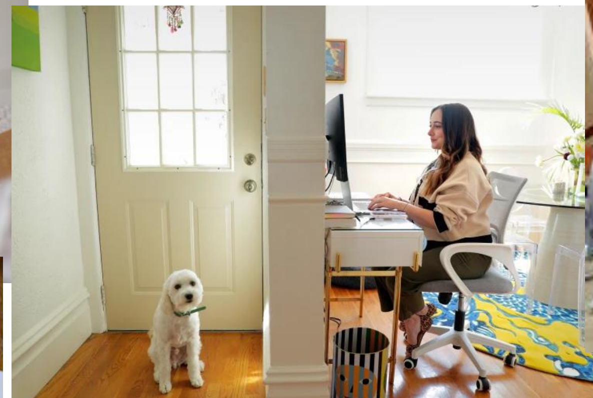


The PDSA Cycle for Learning and Improvement is central to successful PEs

Source: Improvement Guide

We are in a period of continuous planned experiments!





People are running experiments every day!

History of Planned Experimentation


PE Journey
Starts Here
THIRD EDITION
QUALITY
IMPROVEMENT
THROUGH
PLANNED
EXPERIMENTATION
RONALD D. MOEN
THOMAS W. NOLAN
LLOYD P. PROVOST

A Brief History of PE (aka DOE)

The agricultural origins (1908 – 1940s)

- W.S. Gossett and the t-test (1908)
- R.A. Fisher, his co-workers and his books
- Randomization, replication and blocking
- Factorial designs, ANOVA
- Profound impact on agricultural sciences

The industrial era (1951 – mid-to-late 1970s)

- Box & Wilson, response surfaces
- Application in the chemical and processing industries
- Important developments in designing various types of experiments
- Applications to mixtures

Source: Dr. Douglas Montgomery, Arizona State University, School of Engineering,
SAS On Demand Video, 17 January 2017

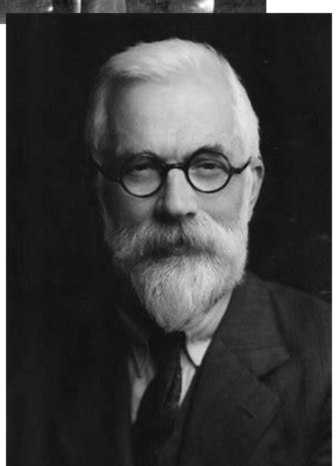
A Little More History on PE (DOE)

The second industrial era (mid-to-late 1970s to 1980s)

- QI initiatives introduced in many companies
- Many organizations discover “DOE”
- Applications of DOE flourished in many industrial sectors beyond the chemical and processing industries
- More methodological work (e.g., hybrid designs, optimal design tools, software enhancements)
- Taguchi and robust parameter designs and process robustness
- Some increase in controversy over methods and applications

Source: Dr. Douglas Montgomery, Arizona State University, School of Engineering,
SAS On Demand Video, 17 January 2017

Sir Ronald Aylmer Fisher

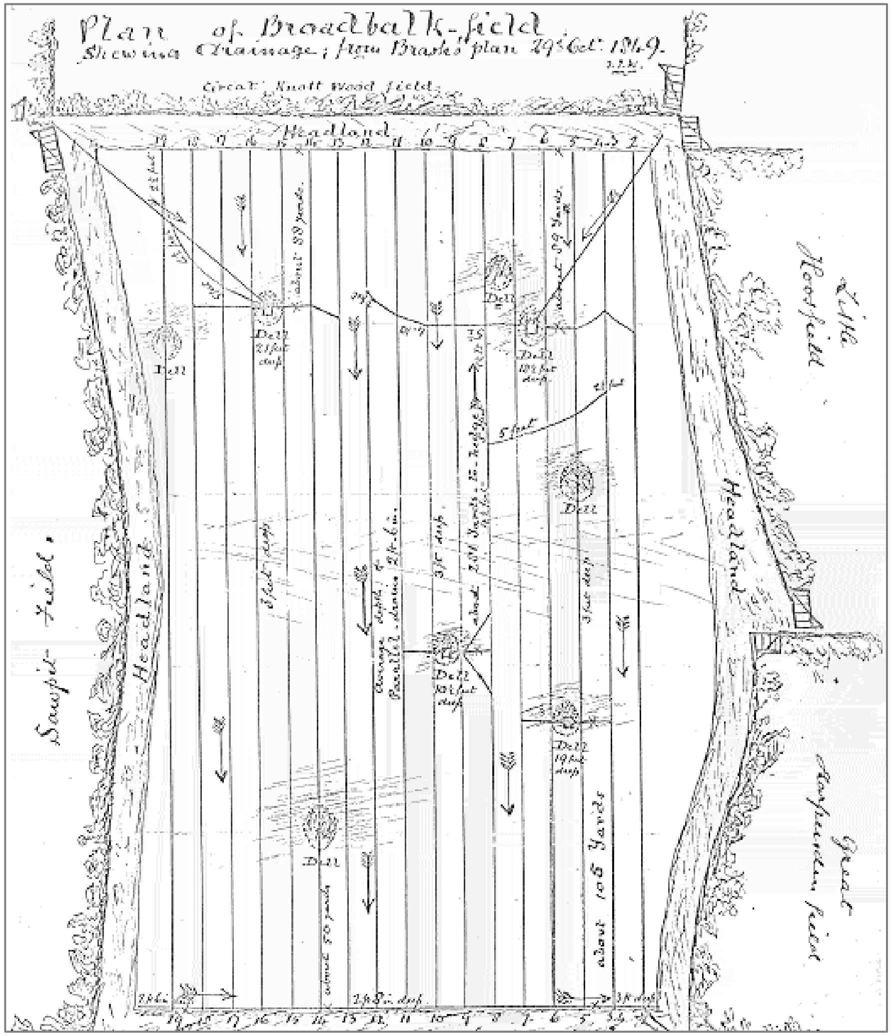


- Sir Ronald Aylmer Fisher (17 February 1890 – 29 July 1962) was a British statistician and geneticist.
- For his work in statistics, he has been described as "a genius who almost single-handedly created the foundations for modern statistical science“ and "the single most important figure in 20th century statistics".
- In genetics, his work used mathematics to combine Mendelian genetics and natural selection; this contributed to the revival of Darwinism in the early 20th-century revision of the theory of evolution known as the modern synthesis.
    - From 1919 onward, he worked at the Rothamsted Experimental Station (England) for 14 years.
    - While at Rothamsted, he analysed its immense data from crop experiments since the 1840s, and developed the analysis of variance (ANOVA).
    - He established his reputation there in the following years as a biostatistician.

Development of Planned Experimentation

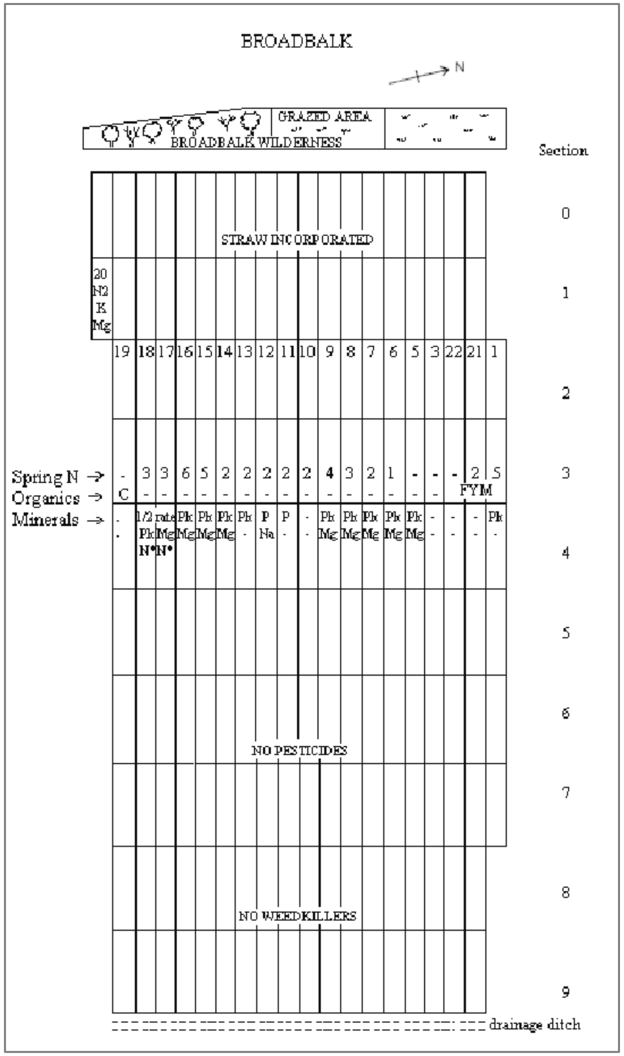
Sir Ronald Aylmer Fisher
(1890-1962)

The primary text for PE
***The Design of Experiments* (1935)**



Rothamsted Agricultural Research Center, 1844
Fisher was here from 1919 – 1933.

Rothamsted, England



1968

Rothamsted (2000)

PennState
College of Agricultural Sciences

1855

2022

A 168 year old experiment!

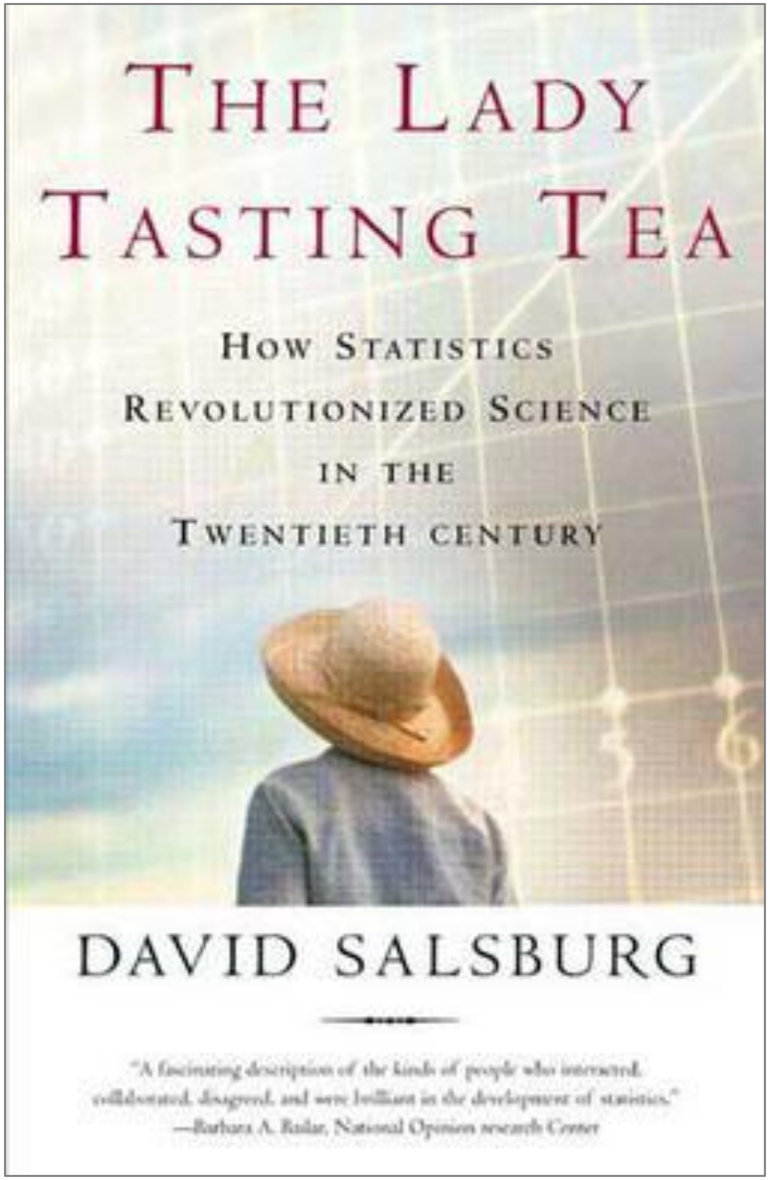
An Experiment
The Lady Tasting Tea

The lady in question (Lady Muriel Bristol) claimed to be able to tell ***whether the tea or the milk was added first to a cup***. Fisher proposed to give her eight cups, four of each variety, in random order. One could then ask what the probability was for her getting the specific number of cups she identified correct, but just by chance.

Fisher's description is less than 10 pages in length and is notable for its simplicity and completeness regarding terminology, calculations and design of the experiment.

Tea-Tasting Distribution Assuming the Null Hypothesis

Success count	Permutations of selection	Number of permutations
0	oooo	1 × 1 = 1
1	ooox, ooxo, oxoo, xooo	4 × 4 = 16
2	ooxx, oxox, oxxo, xoxo, xxoo, xoox	6 × 6 = 36
3	oxxx, xoxx, xxox, xxxo	4 × 4 = 16
4	xxxx	1 × 1 = 1
Total		70



The experiment provides a subject with 8 randomly ordered cups of tea – 4 prepared by first adding milk, 4 prepared by first adding the tea. The subject must select 4 cups prepared by one method. Judging cups by direct comparison is allowed. The method employed in the experiment is fully disclosed to the subject.

The null hypothesis is, that the subject has no ability to distinguish the teas. In Fisher's approach, there was no alternative hypothesis.

Types of Studies and Experiments

Planned Experimentation

An experiment is a ***study*** designed to provide a basis for action. It is structured around changing one or more measures of components of a system to determine the affect that these components have on a process or outcome measure.

A Plan, Do, Study, Act Cycle (PDSA) is a type of experiment.

Planned experimentation (PE) is a collection of approaches and methods to help increase the rate of learning about improvements to systems, processes or products.

Three principles of testing a change:

- Test on a small scale and build knowledge sequentially
- Collect data over time
- Include a wide range of conditions

API, 2015

Types of Planned Experiments

- An experiment consists of a **series of tests to a system** carried out by changing levels of factors and background variables and an observation of the effect of that change on one or more response variables.
- Experiments are a **natural part of life** and happen all the time. But usually, the choice of the specific tests is made haphazardly or conveniently, without planning.
- The PDSA cycle introduced in Chapter 1 provides a structure to formally plan the test that is conducted.
- The degree of up-front planning distinguishes observational studies (also called natural experiments or retrospective studies) from planned experimentation.

Types of Experiments

Very Informal

1. *Trial-and-learning methods (PDSA tests of change)*

   Introduce a change and see what happens. One-shot case studies (Campbell & Stanley)
2. *Running special lots or batches*

   Produced under controlled conditions
3. *Pilot runs*

   Set up to produce a desired effect
4. *One-factor experiment*

   A single change with background variables
5. *Experiment planned with two to four factors*

   Study separate effects and interactions
6. *Experiment with 5 to 20 factors*

   Screening studies
7. *Comprehensive experimental plan with many phases*

   Modeling, multiple factor levels, optimization

Very Formal

PE book: 34

Key PE Articles*

On Probability As a Basis For Action*

W. EDWARDS DEMING**

Abstract

The aim of the author is improvement of statistical practice. The author distinguishes between enumerative studies and analytic studies. An enumerative study has for its aim an estimate of the number of units of a frame that belong to a specified class. An analytic study has for its aim a basis for action on the cause-system or the process, in order to improve product of the future. A fair price to pay for an inventory is an example of an enumerative study. Tests of varieties of wheat, insecticides, drugs, manufacturing processes, are examples of analytic studies: the choice of variety or treatment will affect the future out-turn of wheat, future patients, future product. Techniques and methods of inference that are applicable to enumerative studies lead to faulty design and faulty inference for analytic problems.

facturing? What is malpractice in medicine?
work in consumer research is in a sorry st
money being spent on it year by year,
worsening examples of practice and present
These problems can not be understood an
even be stated, nor can the effect of any allege
be evaluated, without the aid of statistical t
methods. One can not even define operation
tives like reliable, safe, polluted, unemploye
(arrivals), equal (in size), round, random,
green, or any other adjective, for use in busi
government, except in statistical terms. A
(as of safety, or of performance or capability
meaning for business or legal purposes, must
in statistical terms.
The label on a blanket reads "50 per ce

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 218 JUNE, 1942 *Volume 37*

ON A CLASSIFICATION OF THE PROBLEMS OF STATISTICAL INFERENCE

BY W. EDWARDS DEMING
Bureau of the Census

Rethinking methods of inference

Analytical studies: a framework for quality improvement design and analysis

Lloyd P Provost

Predictable

"Why has it taken so long to understand that processes need analytic methods, not enumerative ones?"

Quality Digest Published: Wednesday, June 13, 2018 - 12:03

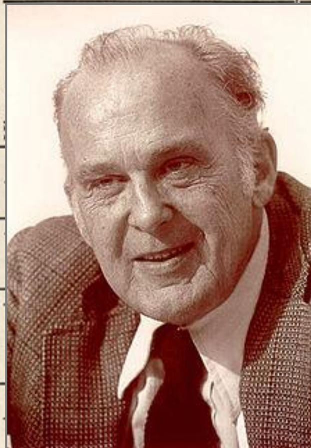
Quality is related to processes. A process is "a series of actions or steps taken in order to achieve a particular end." It doesn't matter whether the process is the handling of invoices, customers in a bank, the manufacture or assembly of parts, insurance claims, the sick passing through a hospital, or any one of thousands of other examples. A process involves movement and action in a sequential fashion.

***These have been posted for your reading enjoyment!**

Campbell, D. and Stanley, J. *Experimental and Quasi-Experimental; Designs for Research*. Houghton Mifflin Co., 1963.

TABLE 1

Sources of Invalidity for Designs 1 through 6



	Internal	External: Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Pre-Experimental Designs:					
1. One-Shot Case Study X O					
2. One-Group Pretest-Posttest Design O X O		–			
3. Static-Group Comparison X O / O					
True Experimental Designs:					
4. Pretest-Posttest Control Group Design R O X O / R O O		–			
5. Solomon Four-Group Design R O X O / R O O / R X O / R O	+ + + + + + + +	+			
6. Posttest-Only Control Group Design R X O / R O	+ + + + + + + +	+			

Note: In ... trolled, a ... is not rele... indicates ... a blank ind...
It is wi... "too help... in the text ... there. In ... or confide... nted becau... nplex and ... ler compreh... reate unco...

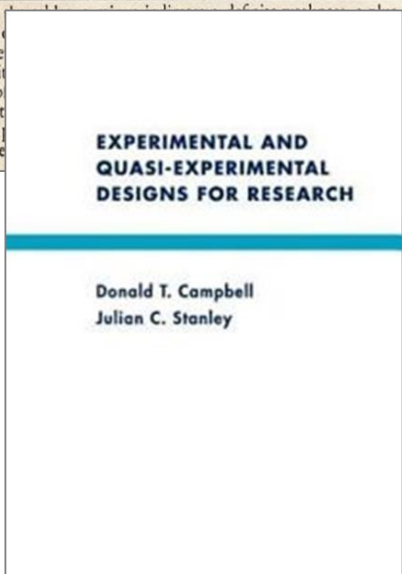


TABLE 2

Sources of Invalidity for Quasi-Experimental Designs 7 through 12

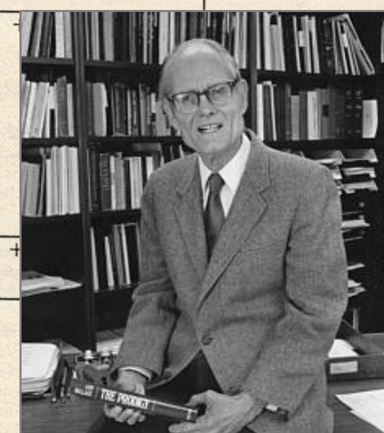
Quasi-Experimental Designs:	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive
7. Time Series O O O OXO O O O	–	+	+	?	+	+	+	+	–	?	
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O, etc.	+	+	+	+	+	+	+	+	–	?	
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O, etc.	+	+	+	+	+	+	+	+	–	?	
10. Nonequivalent Control Group Design O X O / O O	+	+	+	+	?	+	+	–	–	?	
11. Counterbalanced Designs X_1O X_2O X_3O X_4O / X_2O X_4O X_1O X_3O / X_3O X_1O X_4O X_2O / X_4O X_3O X_2O X_1O	+	+	+	+	+	+	+	?	?	?	
12. Separate-Sample Pretest-Posttest Design R O (X) / R X O	–	–	+	?	+	+	–	–	+	+	+
12a. R O (X) / R X O / R O (X) / R X O	+	–	+	?	+	+	–	+	+	+	+
12b. R O_1 (X) / R O_2 (X) / R X O_3	–	+	+	?	+	+	–	?	+	+	
12c. R O_1 X O_2 / R X O_3	–	–	+	?	+	+	+	–	+	+	+

TABLE 3

Sources of Invalidity for Quasi-Experimental Designs 13 through 16

Quasi-Experimental Designs Continued:	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
13. Separate-Sample Pretest-Posttest Control Group Design R O (X) / R X O / R O / R O	+	+	+	+	+	+	+	–	+	+	+	
13a. R O (X) / R X O / R O (X) / R X O / R O (X) / R X O / R O / R O / R O / R O / R O / R O	+	+	+									
14. Multiple Time-Series O O OXO O O / O O O O O O	+	+	+	+								
15. Institutional Cycle Design: Class A X O_1 / Class B_1 RO_2 X O_3 / Class B_2 R X O_4 / Class C O_5 X / [a]Gen. Pop. Con. Cl. B O_6 / [a]Gen. Pop. Con. Cl. C O_7												
$O_2 < O_1$, $O_5 < O_4$	+	–	+	+	?	–	?		+	?	+	
$O_2 < O_3$	–	–	–	?	?	+	+		–	?	+	
$O_2 < O_4$	–	–	+	?	?	+	?		+	?	?	
$O_6 = O_7$, $O_{2y} = O_{2o}$		+						–				
16. Regression Discontinuity	+	+	+	?	+	+	?	+	+	–	+	+

[a] General Population Controls for Class B, etc.



Q Manage Health Care
Vol. 13, No. 1, pp. 17–32

Study Designs for PDSA Quality Improvement Research

Theodore Speroff, PhD; Gerald T. O'Connor, PhD, DSc

"The purpose of this article is to advocate for the use of <u>quasi-experimental strategies</u> to improve the scientific foundation of PDSA quality improvement in health care.

PDSA quality improvement – data are collected to demonstrate that change by intervention resulted in improvement."

(2004)

Deming on Prediction!

"Why does anyone make comparisons of two methods, two treatments, two processes, or two materials? Why does anyone carry out a test or an experiment? The answer is ***to predict*** *– whether one of the methods or materials tested will in the future, under a specified range of conditions perform better than the other one.*

Prediction is the problem*, whether we are talking about applied science, research and development, engineering, or management in industry, education or government. The question is, what do the data tell us? How do they help us to predict?"*

From the Forward in *Quality Improvement Through Planned Experimentation,* page xiii.

Deming on Prediction! (continued)

"Unfortunately, the statistical methods in textbooks and in the classroom do not tell the student that the problem in the use of data is prediction. What the student learns is how to calculate a variety of tests (t-test, F-test, chi square, goodness of fit, etc.) in order to announce that the difference between the two methods or treatments is either significant or not significant. Unfortunately, such calculations are mere formality. ***<u>Significance or the lack of it provides no degree of belief - high, moderate or low – about prediction or performance in the future</u>****, which is the only reason to carry out the comparison, test, or experiment in the first place."*

From the Forward in *Quality Improvement Through Planned Experimentation,* page xiii.

Analytic Studies and Prediction

Deming (1942) emphasized that the primary reason to carry out an experiment is to provide a basis for **action on the system**. He also classified studies into two types (1975) depending on the type of action that will be taken:

- An **enumerative study** is one in which action will be taken on the universe that was studied (examples: conducting a census, or sampling materials for a decision on acceptance or pricing).

- An **analytic study** is one in which action will be taken on a causal system to improve performance of a product, process, or system in the future (examples: a study to select a future raw material supplier or a using a Shewhart control chart to learn and improve a process).

PE book: 30

Enumerative versus Analytic Studies and Related Statistical Techniques

"The teaching of pure statistical theory in universities, including the theory of probability and related subjects is almost everywhere excellent. Application to enumerative studies is mostly correct, but application to analytic problems is deceptive and misleading.

Analysis of variance, t-test, confidence intervals, and other statistical techniques taught in books, however interesting, are inappropriate because they provide no basis for prediction and because they bury the information contained in the order of production. Most if not all computer packages for analysis of data, as they are called, provide flagrant examples of inefficiency."

Dr. Deming, *Out of the Crisis*, page 132.

On Probability As a Basis For Action*

W. EDWARDS DEMING**

Abstract

The aim of the author is improvement of statistical practice. The author distinguishes between enumerative studies and analytic studies. An enumerative study has for its aim an estimate of the number of units of a frame that belong to a specified class. An analytic study has for its aim a basis for action on the cause-system or the process, in order to improve product of the future. A fair price to pay for an inventory is an example of an enumerative study. Tests of varieties of wheat, insecticides, drugs, manufacturing processes, are examples of analytic studies: the choice of variety or treatment will affect the future out-turn of wheat, future patients, future product. Techniques and methods of inference that are applicable to enumerative studies lead to faulty design and faulty inference for analytic problems.

It is possible, in an enumerative problem, to reduce errors of sampling to any specified level. In contrast, in an analytic problem, it is impossible to compute the risk of making a wrong decision. The author provides a number of examples, and pleads for greater care in the writing and teaching of statistical theory and inference.

* * * * *

Aim and scope of this paper. The aim here is to try to contribute something to the improvement of statistical practice. The basic supposition here is that any statistical investigation is carried out for purposes of action. New knowledge modifies existing knowledge.

Urgent needs for statistical work. Challenges face statisticians today as never before. The whole world is talking about safety in mechanical and electrical devices (in automobiles, for example), safety in drugs, reliability, due care, pollution, poverty, nutrition, improvement of medical practice, improvement of agricultural practice, improvement in quality of product, break-down of service, break-down of equipment, tardy busses, trains, and mail, need for greater output in industry and in agriculture, enrichment of jobs. The consumer requires month by month ever greater and greater safety, and he expects better and better performance of manufactured articles. The manufacturer has the same problems in his purchases of materials, assemblies, machines, and use of manpower. He must, in addition, know more and more about his own product. What is due care in manufacturing? What is malpractice in medicine? Statistical work in consumer research is in a sorry state, more money being spent on it year by year, with ever worsening examples of practice and presentation.

These problems can not be understood and can not even be stated, nor can the effect of any alleged solution be evaluated, without the aid of statistical theory and methods. One can not even define operationally adjectives like reliable, safe, polluted, unemployed, on time (arrivals), equal (in size), round, random, tired, red, green, or any other adjective, for use in business or in government, except in statistical terms. A standard (as of safety, or of performance or capability) to have meaning for business or legal purposes, must be defined in statistical terms.

The label on a blanket reads "50 per cent wool." What does this mean? Half wool, on the average, over this blanket, or half wool over a month's production? What is half wool? Half by weight? If so, at what humidity? By what method of chemical analysis? How many analyses? The bottom half of the blanket is wool and the top half is something else. Is it 50 per cent wool? Does 50 per cent wool mean that there must be some wool in any random cross-section the size of a half dollar? If so, how many cuts shall be tested? How select them? What criterion must the average satisfy? And how much variation between cuts is permissible? Obviously, the meaning of 50 per cent wool can only be stated in statistical terms. Mere words in English, French, or Japanese will not suffice. What means 80% butter fat in the butter that you buy?

Drastic changes in practice and in writing and in teaching are called for. As Shewhart said [18], the standards of knowledge and workmanship in industry and in public service are more severe than the requirements in pure science. He ought to have added that the requirements for statistical practice are also far more rigid than the requirements imposed on the teaching of statistics. It ought not to be that way, but it is. (More later on teaching.)

The frame, the universe, environmental conditions. A statistical study proceeds by investigation of the material in a frame [19]. The frame is an aggregate of identifiable tangible physical units of some kind, any or all of which may be selected and investigated. The frame may be lists of people, areas, establishments, materials, or of other identifiable units that would yield useful results if the whole content were investigated. It may be a lot of manufactured parts. Equally important in an analytic problem is a description of the environmental conditions that may affect the results (*vide infra*).

To facilitate exposition, we use a frame of N sampling units, numbered serially $1, 2, 3, \ldots, N$. However, there are circumstances in practice in which the size of

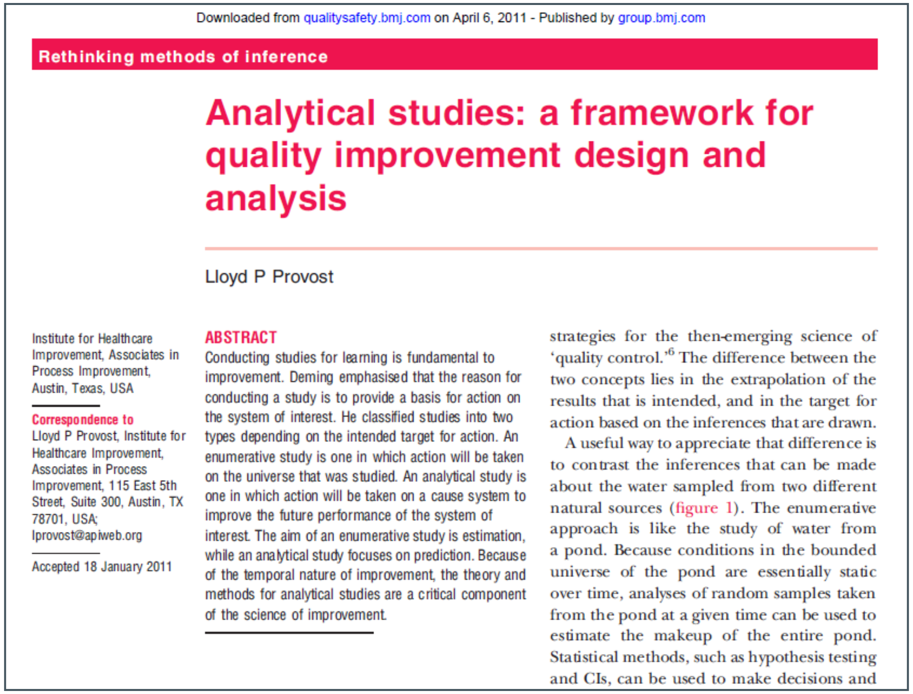
* I am indebted to many critics of earlier drafts of the manuscript for this paper; also to questions from the audience at lectures at a number of universities, including the Princeton meeting of the Biopharmaceutical Section of the American Statistical Association 4 Dec. 1974; the Universities of Mainz, Colorado, Wyoming, George Washington University, North Carolina, Inter-American Statistical Institute in Santiago de Chile.

** Consultant in Statistical Surveys, 4924 Butterworth Pl., Washington 20016.

"It is possible, in an enumerative problem, to reduce errors of sampling to any specified level. In contrast, in an analytic problem, it is impossible to compute the risk of making a wrong decision."

"On Probability as Basis for Action"
W. E. Deming, *The American Statistician*, November 1975, vol. 29, No. 4. Pages 146-152.

Analytic Studies in BMJ (2011)



Rethinking methods of inference

Analytical studies: a framework for quality improvement design and analysis

Lloyd P Provost

Institute for Healthcare Improvement, Associates in Process Improvement, Austin, Texas, USA

Correspondence to Lloyd P Provost, Institute for Healthcare Improvement, Associates in Process Improvement, 115 East 5th Street, Suite 300, Austin, TX 78701, USA; lprovost@apiweb.org

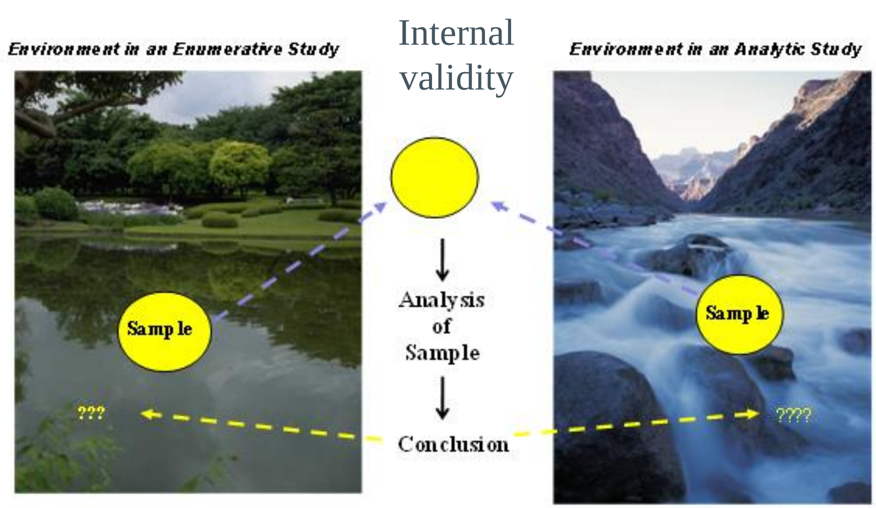
Accepted 18 January 2011

ABSTRACT
Conducting studies for learning is fundamental to improvement. Deming emphasised that the reason for conducting a study is to provide a basis for action on the system of interest. He classified studies into two types depending on the intended target for action. An enumerative study is one in which action will be taken on the universe that was studied. An analytical study is one in which action will be taken on a cause system to improve the future performance of the system of interest. The aim of an enumerative study is estimation, while an analytical study focuses on prediction. Because of the temporal nature of improvement, the theory and methods for analytical studies are a critical component of the science of improvement.

strategies for the then-emerging science of 'quality control.'[6] The difference between the two concepts lies in the extrapolation of the results that is intended, and in the target for action based on the inferences that are drawn.

A useful way to appreciate that difference is to contrast the inferences that can be made about the water sampled from two different natural sources (figure 1). The enumerative approach is like the study of water from a pond. Because conditions in the bounded universe of the pond are essentially static over time, analyses of random samples taken from the pond at a given time can be used to estimate the makeup of the entire pond. Statistical methods, such as hypothesis testing and CIs, can be used to make decisions and

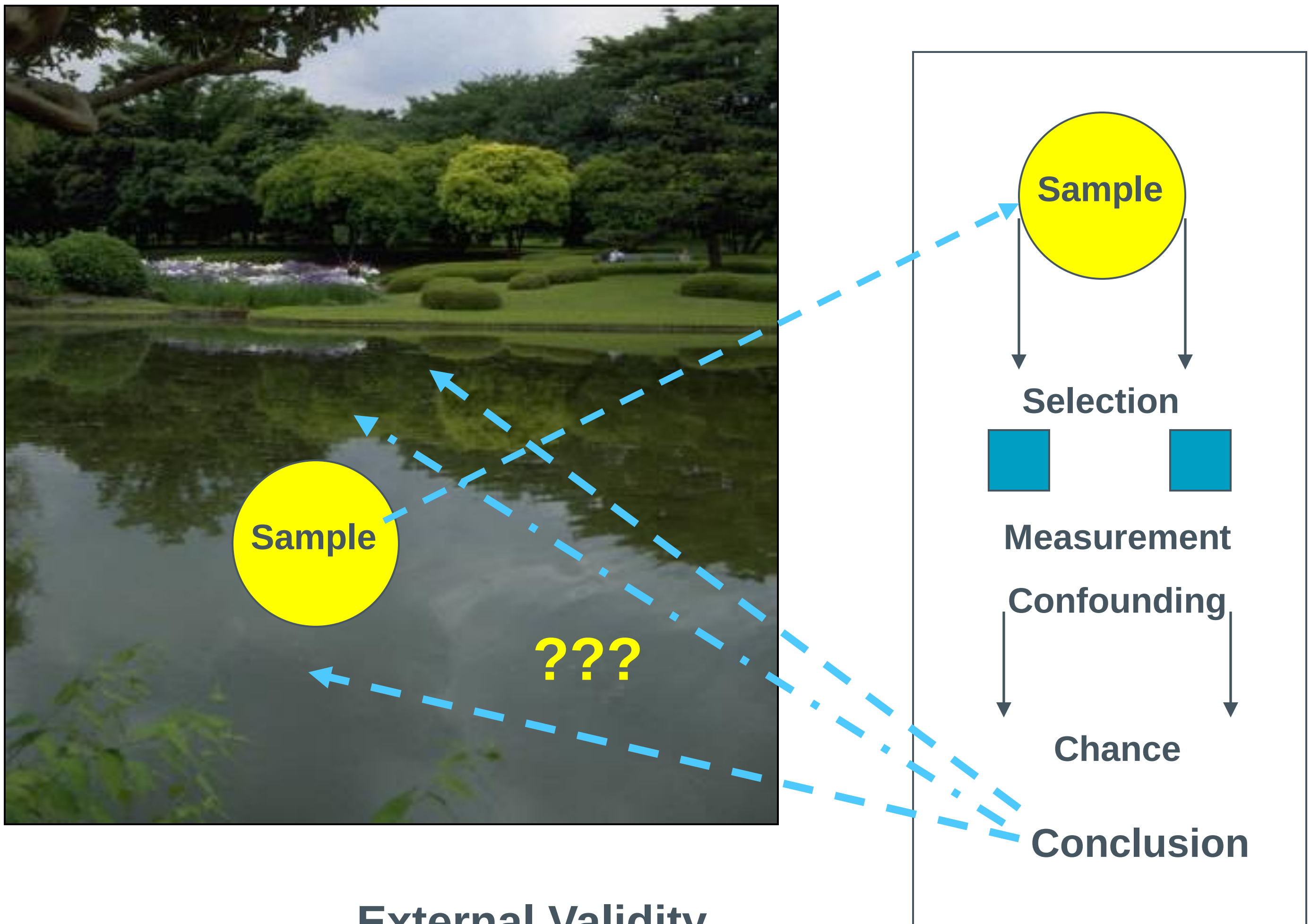
See Provost (2011) for further discussion of analytic studies. The classification of studies as enumerative or analytic ***depends on the intended target for action.***



p. 30

Environment in an Enumerative Study

Internal Validity



External Validity

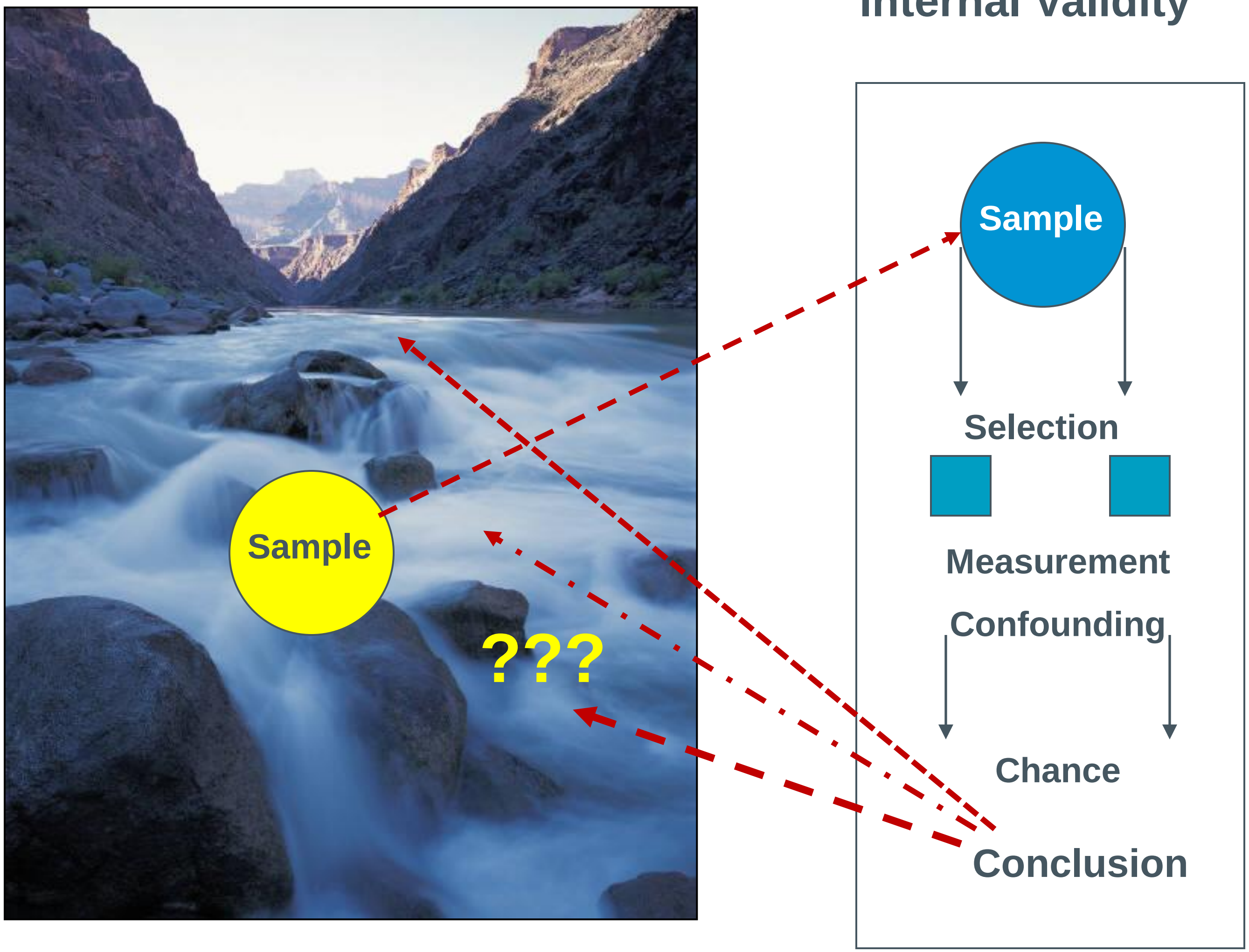
(generalizability)

Clinical epidemiology

Fletcher, Fletcher, Wagner

Environment in an Analytic Study

Internal Validity



External Validity

(generalizability)

Clinical epidemiology

Fletcher, Fletcher, Wagner

Analysis of Data from Analytic Studies

- In an **enumerative study**, the **existence of a distribution** for the characteristic of interest is ensured by the existence of a frame. Summary statistics such as a mean and standard deviation can be used to estimate parameters of the distribution. These estimates will have a quantifiable measure of uncertainty if the sample from the frame is chosen using a random number table. This type of analysis is usually an important step in accomplishing the aim of an enumerative study.
- In an **analytic study** **the aim is prediction**. A distribution useful for even short-term (days or weeks) prediction may not exist for any characteristics of a new product. The standard error of a statistic or the standard deviation does not address the most important source of uncertainty in an analytic study: **factors outside the conditions of the study that will change in the future**.

Estimation and Prediction in Different Types of Studies

Leverage for improvement	Application Examples	Theory to support use of the standard error	Role of subject matter expert
Low	**Estimation**		
	Acceptance sampling Valuing inventory Census Exit poll of voters	Probability distribution in combination with a frame and sampling by random numbers	Approval of the frame and definition of the complete coverage
High			

PE book: 32

Examples of Different Types of Studies

The approach to research and the statistical methods used should be based on the *question(s) being asked*.

A Descriptive Study Questions

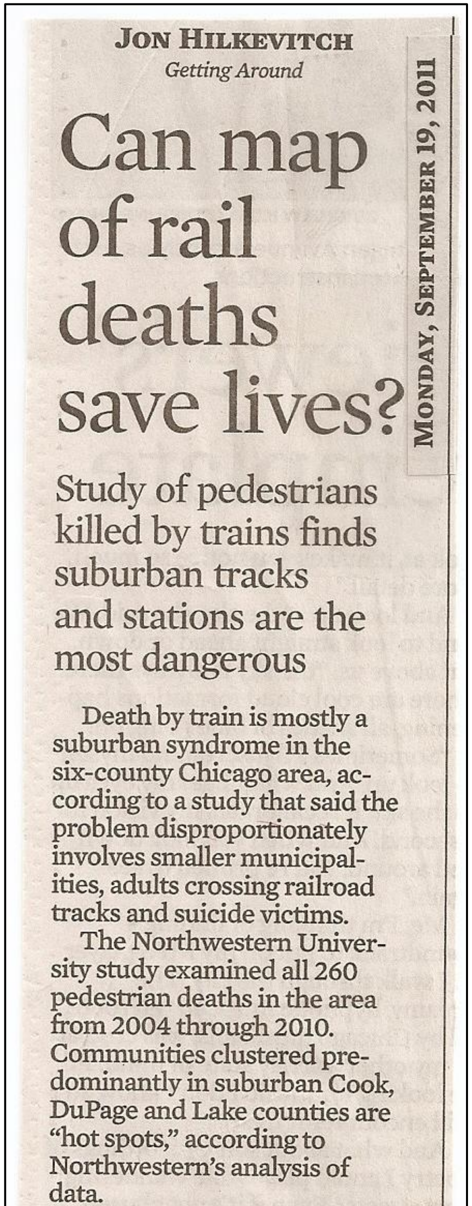
- How many barrels were filled today?
- How many pounds of fish were caught?
- What percent of fish today were cod?
- How many hours did it take to fish today?
- What is the average price we get for a barrel?

An Enumerative Study Questions

- How can we select one of the 50 barrels of fish on the deck? (randomly?)
- Let's use a test of significance to determine if today's catch is statistically different from last week's catch.
- Is the percent of cod in the selected barrel statistically different (using the .05 level of significance) from the percent of cod we took from a random barrel last week when we used a different type of bait?

An Analytic Study Questions

- What is the process by which certain types of fish end up in a barrel?
- How much variation is there in the quantity and types of fish we catch?
- Does our fishing method impact volume and type?
- What can we predict about the next catch?


Jon Hilkevitch
Getting Around

Monday, September 19, 2011

Can map of rail deaths save lives?

Study of pedestrians killed by trains finds suburban tracks and stations are the most dangerous

Death by train is mostly a suburban syndrome in the six-county Chicago area, according to a study that said the problem disproportionately involves smaller municipalities, adults crossing railroad tracks and suicide victims.

The Northwestern University study examined all 260 pedestrian deaths in the area from 2004 through 2010. Communities clustered predominantly in suburban Cook, DuPage and Lake counties are "hot spots," according to Northwestern's analysis of data.

Case Study: The Chicago Tribune

Monday, September 19, 2011

"The purpose of the study, which represents the most comprehensive examination of railroad pedestrian fatalities in northeastern Illinois, was to determine the factors leading to the incidents (of death) and recommend solutions the researchers said."

Does this sound like an enumerative or analytic study?

The Chicago Tribune

Monday, September 19, 2011

Variables in the study

- Train type (Metra, Amtrak or Freight)
- Number of pedestrian deaths by age
- Number of pedestrian deaths by gender
- Pedestrian death rate by Metra route
- Pedestrian deaths (count) and rate by municipality
- Percentage of deaths by season

Does this sound like an enumerative or analytic study?

The Chicago Tribune

Monday, September 19, 2011

Data Analysis

Age disparity

Number of pedestrian deaths by age
In the Chicago area, 2004-2010

Under 19	26
19 to 40	90
41 to 65	88
65 and older	15

NOTE: Deaths where age was not reported not shown.

Men are more often victims

Males outnumber females by more than 2-to-1 in trespassing deaths and suicides, and account for 70 percent of all fatalities where gender is known.

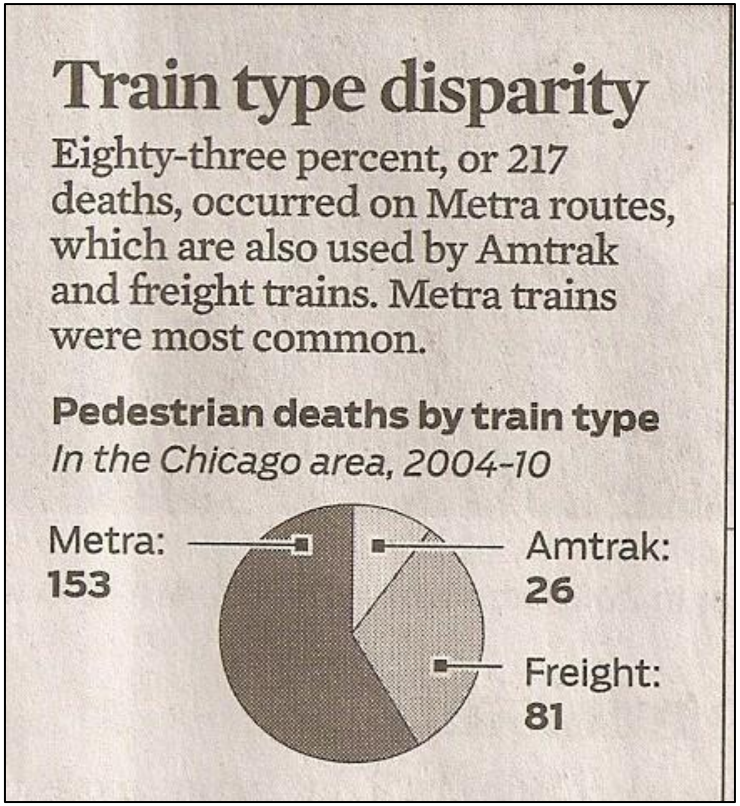
Number of pedestrian deaths by gender
In the Chicago area, from 2004-2010

	Male	Female	Not reported
Stations and crossings	31	23	5
Trespassers	49	14	18
Suicides	81	32	7

Fatalities higher in warmer months

More than 50 percent of all fatalities occurred in spring and summer, though trespasser deaths peaked in the fall months. Deaths at stations and crossings were highest in the summer.

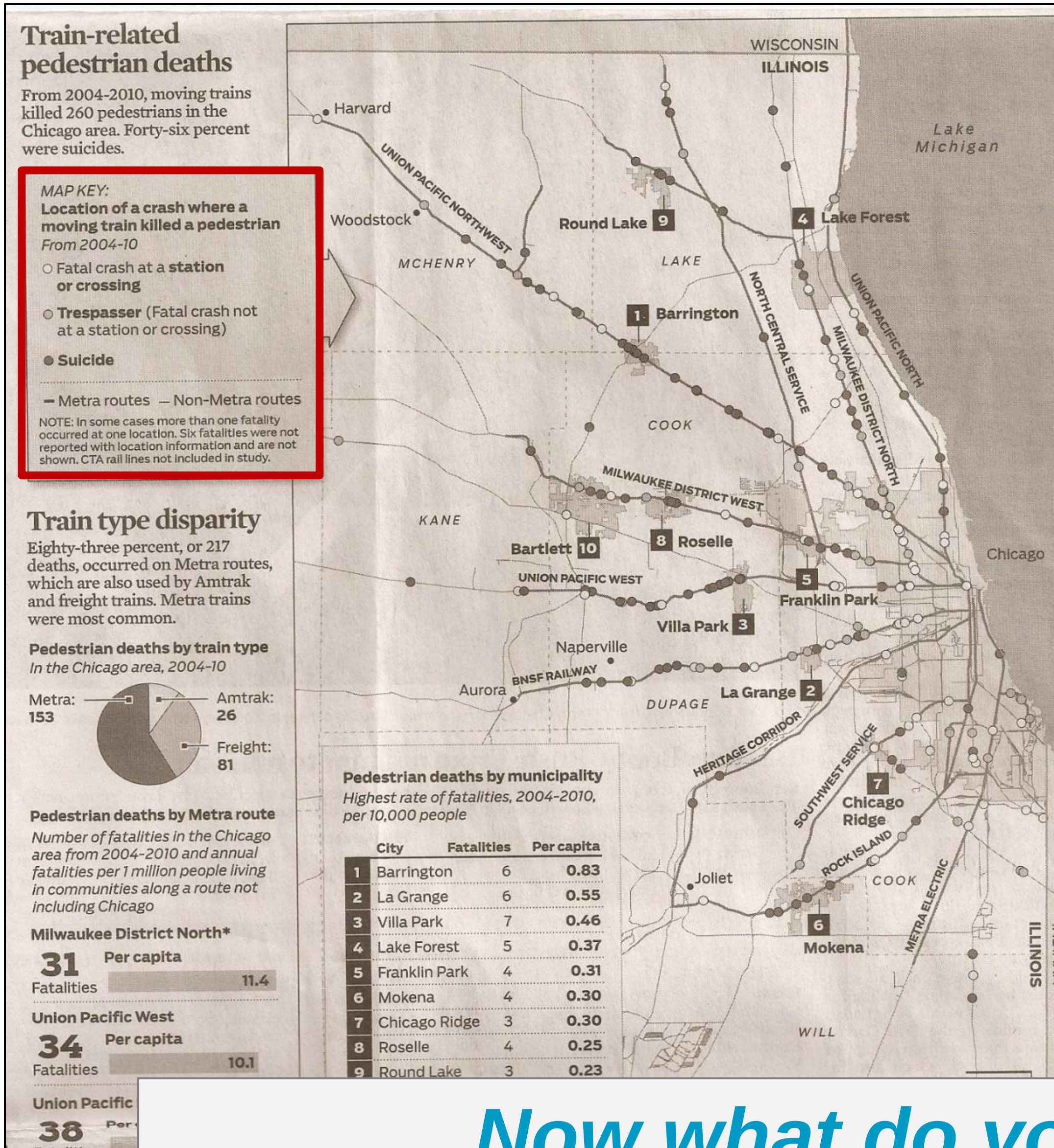
Percentage of pedestrian deaths by season
In the Chicago area, from 2004-2010



Train type disparity

Eighty-three percent, or 217 deaths, occurred on Metra routes, which are also used by Amtrak and freight trains. Metra trains were most common.

Pedestrian deaths by train type
In the Chicago area, 2004-10



The Chicago Tribune

Monday, September 19, 2011

Conclusion

"Fatal rail pedestrian incidents are occurring at an average of about one every 10 days in the Chicago area," the study said. "Last week, there were two, both on Thursday."

Now what do you think?

Is this an enumerative or analytic study?

Planned Experimentation Terminology

Terminology used in Planned Experimentation

- Variables in an Experiment:
  -
  -
  -
  -
- Level
- Experimental unit
- Blocks
- Effect

Who can tell us what these are?

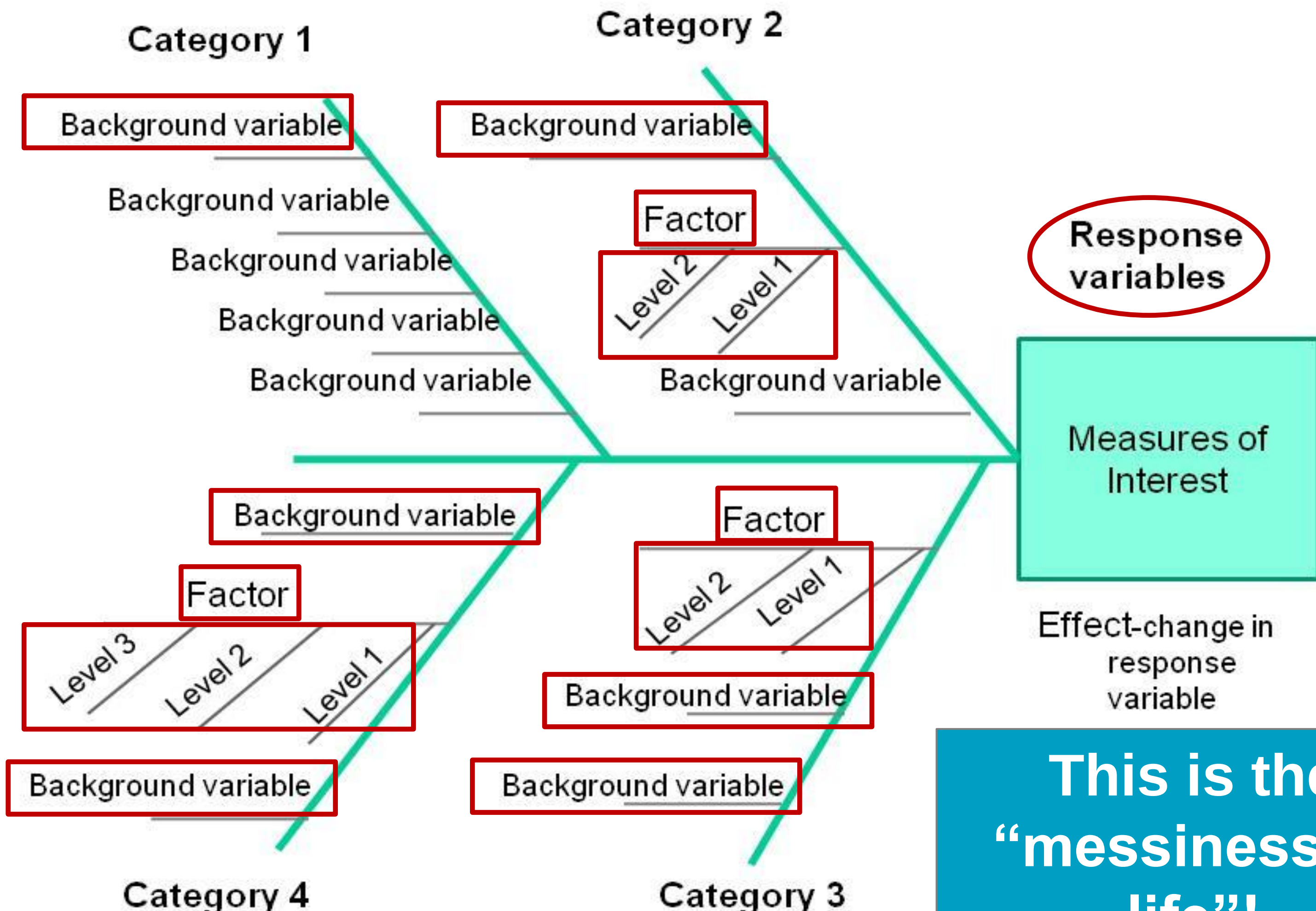
PE book: 26-29

Terminology in Planned Experimentation

- **Variables in an Experiment**:
  - **Response variable** (outcome measure)
  - **Factor** (treatment or intervention)
  - **Background variable** (baseline characteristics of those being studied)
  - **Nuisance variable** (noise)
- **Level** (a given value, a specific setting or a treatment option)
- **Experimental Unit** (e.g., student, school, patient or clinic)
- **Blocks** (groups of patients [EU's] with similar characteristics)
- **Effect** (change in the outcome measure; related to your aim. What happened to the response variable when we changed the levels of factors?)

PE book: 26-29

Applying the language of a Planned Experiment to a Cause-and-Effect Diagram



This is the "messiness of life"!

Response Variable – a variable observed or measured in an experiment; sometimes called a dependent variable (Y). The response variable is an output of an experiment and is often a quality characteristic or a measure of performance of the process. An experiment will have one or more response variables.

Factor – a variable which is deliberately varied or changed in a controlled manner in an experiment to observe its impact or effect on the response variable(s); sometimes called an independent or control variable or a causal variable.

Factor Level – A given value or specific setting of a quantitative factor or a specific option of a qualitative factor

Background Variable – a variable that can potentially affect a response variable in an experiment but will not be deliberately changed in the same way as a factor.

The PE Form*

Documentation of Planned Experiment

Objective:			
Background Information:			
1. **Experimental Variables:**			
	A.	Response variables	Measurement technique
		1.	
		2.	
		3.	
		4.	
	B.	Factors under study	Levels
		1.	
		2.	
		3.	
		4.	
		5.	
		6.	
		7.	
	C.	Background variables	Method of control
		1.	
		2.	
		3.	
		4.	
2. **Experimental Unit:**			
3. **Replication:**			
4. **Methods of randomization:**			
5. **Design matrix: (attach copy)**			
7. **Data collection forms: (attach copies)**			
8. **Planned methods of statistical analysis:**			
9. **Estimated cost, schedule, and other resource consideration:**			

***This form is included with the handouts for this session.**

Now it is your turn!

- Think of a QI project that you have worked on or one that is currently underway.
- **Can you identify the:**
  - Response Variable?
  - 1 or more Factors?
  - Levels for each factor?
  - Experimental Unit?
  - Background Variables?
  - Blocks?

Response Variable – a variable observed or measured in an experiment; sometimes called a dependent variable (Y). The response variable is an output of an experiment and is often a quality characteristic or a measure of performance of the process. An experiment will have one or more response variables.

Factor – a variable which is deliberately varied or changed in a controlled manner in an experiment to observe its impact or effect on the response variable(s); sometimes called an independent or control variable or a causal variable.

Factor Level – A given value or specific setting of a quantitative factor or a specific option of a qualitative factor

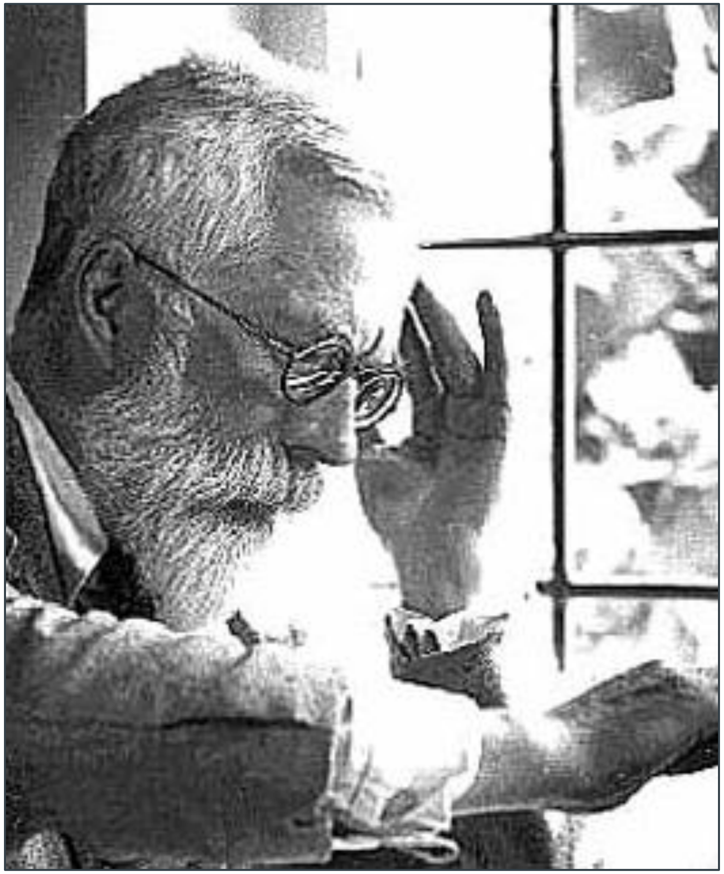
Background Variable – a variable that can potentially affect a response variable in an experiment but will not be deliberately changed in the same way as a factor.

Principles and Tools of Planned Experimentation

5 Principles for Designing Analytic Studies

1. Well defined objective
2. Sequential approach
3. Partitioning variation
4. Degree of belief
5. Simplicity of execution

See Appendix B for details on the 5 principles for designing analytic studies.



Sir Ronald A. Fisher 1890-1962

Now consider the 4 Key Tools for Analytic Studies

1. **Experimental pattern** - learn about each factor of interest
2. **Planned grouping** - managing background variables
3. **Randomization** - impact of nuisance variables
4. **Replication** - increase degree of belief

Now consider the 4 Key Tools for Analytic Studies

R. A. Fisher (1935) described the use of four tools that can be used to help ensure that an experiment follows these principles.

1. ***Experimental pattern***: The arrangement of factor levels and experimental units in the design.
2. ***Planned grouping***: Blocking of experimental units.
3. ***Randomization***: The objective assignment of specific combinations of factor and levels to specific experimental units.
4. ***Replication***: Repetition of experiments, experimental units, measurements, treatments, and other components as part of the planned experiment.

Using experimental tools to attain the principles of a good experiment:

Tool				
Property	**Experimental pattern**	**Planned grouping**	**Randomization**	**Replication**
Well-defined objective	XX			
Sequential approach	X	XX		X
Partitioning variation	X	X	X	X
Degree of belief	X	X	X	XX
Simplicity of execution	X	X	-X	
X - *direct effect* XX - *very strong impact* -X - *negative effect*				

PE, p. 40-51

Tool #1
Experimental Patterns for Factorial Designs

2 Factors, 2 levels each

F_2	- F_1	+
–		
+		

(a) 2^2 design

(Also as 2^2)

3 Factors, 2 levels each

F_2	- F_3		+	
	- F_1	+	- F_1	+
–				
+				

(b) 2^3 design

4 Factors, 2 levels each

F_4	F_2	- F_3		+	
		- F_1	+	- F_1	+
–	–				
	+				
+	–				
	+				

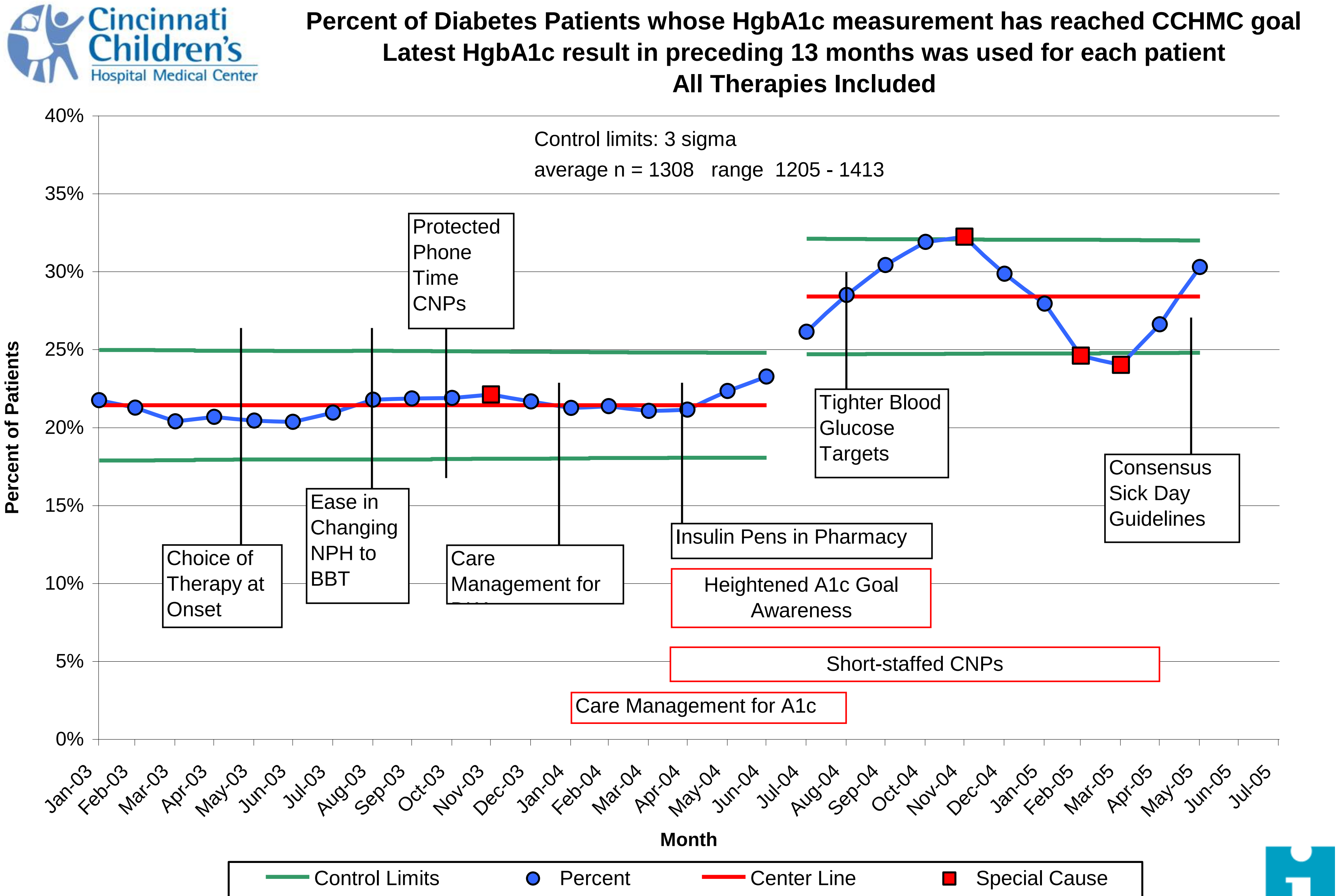
(c) 2^4 design

Factor 1 at 2 levels, Factor 2 at 3 levels

	F1 -	F1 +
F2 –l1		
F2 – l2		
F2 –l3		

(d) 2x3 design

Example: Using PDSA Tests to Improve Diabetic Management



Last update: 06-23-05 by H. Atherton, Data source: Disease Management Database

The Alternative: A Factorial Design

How many factors are being tested?

Figure 2.3: A factorial experimental pattern for improving care of diabetes patients

		No Care Management Program		Care Management Program	
		Therapy 1	Therapy 2	Therapy 1	Therapy 2
Manage to easy Glucose Targets	No Sick day guidelines	20 patients	20 patients	20 patients	20 patients
	Sick day guidelines	20 patients	20 patients	20 patients	20 patients
Manage to Tight Glucose Targets	No Sick day guidelines	20 patients	20 patients	20 patients	20 patients
	Sick day guidelines	20 patients	20 patients	20 patients	20 patients

Source: Cincinnati Children's Hospital Medical Center

Designs should incorporate response variables as a time series

Run charts or Shewhart charts can be used to assess whether improved levels of performance have been achieved and are being maintained. **Five time series designs** described in the literature are described in the table below.

Time Series Design	Description	Appropriate Experimental Pattern
Before-and-after time series	Testing a change where before (baseline) and after data are collected on an EU over time	One factor design
Time series with replication	Testing a change over time where the before and after data are repeated to see if the process returns to initial levels	One factor design with multiple experimental units
Time series with a control group	Testing a change over time where there is no baseline for comparison	One-factor design at two levels run in parallel
Time series with planned grouping	Testing a change over time under a wide range of conditions	One-factor design with blocks
Time series with two factors	Testing multiple combinations of factors and level on an EU over time	Factorial design

Langley, G, et al. The Improvement Guide, Jossey-Bass, 2nd Edition, 2009: 156-169.

And this is where Campbell & Stanley come back into the picture!

TABLE 1

SOURCES OF INVALIDITY FOR DESIGNS 1 THROUGH 6

Sources of Invalidity

	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Pre-Experimental Designs:												
1. One-Shot Case Study *X O*	−	−			−	−						
2. One-Group Pretest-Posttest Design *O X O*	−	−	−	−	?	+	+	−				
3. Static-Group Comparison *X O* ------ *O*	+	?	+	+	+	−	−	−				
True Experimental Designs:												
4. Pretest-Posttest Control Group Design *R O X O* *R O O*	+	+	+	+	+	+	+	+				
5. Solomon Four-Group Design *R O X O* *R O O* *R X O* *R O*	+	+	+	+	+	+	+	+				
6. Posttest-Only Control Group Design *R X O* *R O*	+	+	+	+	+	+	+	+				

Note: In the tables, a minus indicates a definite weakness, a plus indi
trolled, a question mark indicates a possible source of concern, and a b
is not relevant.

It is with extreme reluctance that these summary tables are presente
"too helpful," and to be depended upon in place of the more comple
in the text. No + or − indicator should be respected unless the reader c
there. In particular, it is against the spirit of this presentation to creat
or confidence in, specific designs.

TABLE 2

SOURCES OF INVALIDITY FOR QUASI-EXPERIMENTAL DESIGNS 7 THROUGH 12

Sources of Invalidity

	Internal								External		
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements
Quasi-Experimental Designs:											
7. Time Series *O O O OXO O O O*	−	+	+	?	+	+	+	+	−	?	?
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O, etc.	+	+	+	+	+	+	+	+	−	?	−
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O, etc.	+	+	+	+	+	+	+	+	−	?	?
10. Nonequivalent Control Group Design *O X O* ------ *O O*	+	+	+	+	?	+	+	−	−	?	?
11. Counterbalanced Designs X_1O X_2O X_3O X_4O X_2O X_4O X_1O X_3O X_3O X_1O X_4O X_2O X_4O X_3O X_2O X_1O	+	+	+	+	+	+	+	?	?	?	?
12. Separate-Sample Pretest-Posttest Design *R O (X)* *R X O*	−	−	+	?	+	+	−	−	+	+	+
12*a*. *R O (X)* *R X O* ------ *R O (X)* *R X O*	+	−	+	?	+	+	−	+	+	+	+
12*b*. *R* O_1 *(X)* *R* O_2 *(X)* *R X* O_3	−	+	+	?	+	+	−	?	+	+	+
12*c*. *R* O_1 *X* O_2 *R X* O_3	−	−	+	?	+	+	+	−	+	+	+

TABLE 3

SOURCES OF INVALIDITY FOR QUASI-EXPERIMENTAL DESIGNS 13 THROUGH 16

Sources of Invalidity

	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Quasi-Experimental Designs Continued:												
13. Separate-Sample Pretest-Posttest Control Group Design *R O (X)* *R X O* ------ *R O* *R O*	+	+	+	+	+	+	+	−	+	+	+	
13*a*. *R′* { *R O (X)* / *R X O* / *R O (X)* / *R X O* / *R O (X)* / *R X O* } *R′* { *R O* / *R O* / *R O* / *R O* / *R O* / *R O* }	+	+	+	+	+	+	+	+	+	+	+	
14. Multiple Time-Series *O O OXO O O* ------ *O O O O O O*	+	+	+	+	+	+	+	+	−	−	?	
15. Institutional Cycle Design Class *A X* O_1 Class B_1 *R*O_2 *X* O_3 Class B_2 *R X* O_4 Class *C* O_5 *X* [a]Gen. Pop. Con. Cl. *B* O_6 [a]Gen. Pop. Con. Cl. *C* O_7												
$O_2 < O_1$, $O_5 < O_4$	+	−	+	+	?	−	?		+	?	+	
$O_2 < O_3$	−	−	−	?	?	+	+		−	?	+	
$O_2 < O_4$	−	−	+	?	?	+	?		+	?	?	
$O_6 = O_7$, $O_{2y} = O_{2o}$		+						−				
16. Regression Discontinuity	+	+	+	?	+	+	?	+	+	−	+	+

[a] General Population Controls for Class B, etc.

Campbell, D. and Stanley, J. *Experimental and Quasi-Experimental; Designs for Research*. Houghton Mifflin Co., 1963.

Tool #2:
Planned Grouping to
Deal with Background Variables

Two Decisions to Make:

1. How to control the background variables so that the effects are not distorted by them.

2. How to use the background variables to establish a wide range of conditions:
   - to increase the degree of belief
   - to aid in designing a robust product or process.

PE book: 43-44

Example of Planned Grouping

Objective: Run an experiment to compare three material suppliers. Each of the three suppliers will submit four prototypes.

A. Identify background variables in the plant that could affect the response variables of interest:

Background variable	Level
Machine	#7, #4
Operator	Joe, Susan, George
Gage	G-102, G-322
Saw Blade	20 blades available
Time (day-to-day)	Many different days possible

Response Variables
- Nicks
- Burrs
- Surface (polished)
- Length

B. Create four blocks with widely varying conditions based on these background variables:

	Block 1	Block 2	Block 3	Block 4
Machine	#7	#4	#7	#7
Operator	Joe	Susan	George	Joe
Gage	G-102	G-322	G-102	G-322
Saw Blade	Blade 1	Blade 2	Blade 3	Blade 4
Time	Day 1	Day 2	Day 3	Day 4

The Aim is to minimize variation within a block and maximize variation between the blocks.

C. Evaluate one prototype from each supplier (A,B,C) in each block (random order within each block)

Test	Block 1	Block 2	Block 3	Block 4
1.	B	B	C	A
2.	A	C	A	C
3.	C	A	B	B

Figure 2.5, page 46

D. Analyze supplier difference within each block; evaluate consistency of differences across blocks

Tool #3: Randomization

- ***Randomization*** helps prevent the variation due to nuisance variables from being confused with the variation due to the factors or the background variables.
- ***Randomization*** is often compared to insurance: you only need it when a problem (i.e., a big special cause from a nuisance variable) occurs.

Randomization (cont.)

Situations in which randomization is particularly important include:

- Experiments conducted when the important response variables have not been brought into a state of statistical control (i.e., special causes are detected on the Shewhart chart).
- Experiments that will be conducted by many different individuals or research assistants.
- Experiments in which the variation due to nuisance variables is expected to be large relative to the magnitude of effects of important factors.
- Formal experiments in which results must be evaluated and used by others (such as customers or senior staff, or readers of a publication) for action to be taken.

How should Experimental Units (EUs) be selected in the study?

- In an **analytic study**, there is no universe from which to draw a sample. However, in designing the study there are decisions to make concerning the conditions under which the product, process, service or system will be run during the study and the outcomes that will be measured for each set of conditions.
- Deming (1975) makes the point that all analytic studies are conducted on ***judgment samples***. The judgment of the expert in the subject matter determines the conditions to be studied and the measurements to be taken for each set of conditions.
- It Is rare in an analytic study for a random selection of conditions or outcomes to be preferable to a judgment selection. The opposite is true of an enumerative study.

	Enumerative Study	Analytic Study
Random Selection	Required	Acceptable
Judgment Sample	No	Preferred

PE p 49

Tool #4: Replication

- **Replication refers to repeating particular aspects of an experiment.** It is the primary tool for studying stability of effects and for increasing the **degree of belief** in the results.
- There are many different types of potential replications in an experiment, including:
    - Repeated measurements of experimental units
    - Multiple experimental units for each combination of factors
    - Partial replication of the experimental pattern
    - Complete replication of the experimental pattern

Remember the 3 basic principles underlie the methods of analysis for analytic studies

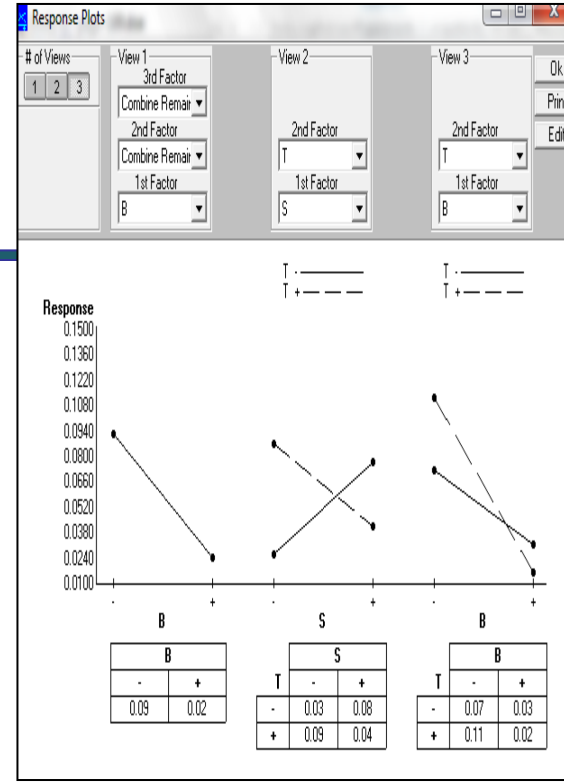
1. The analysis of data, the interpretation of the results, and the actions that are taken as a result of the study will be closely connected to the **current knowledge of experts** in the relevant subject matter.

2. **The conditions of the study will be different from the conditions under which the results will be used.** An assessment of the magnitude of this difference and its impact by experts in the subject matter should be an integral part of the interpretation of the results of the experiment.

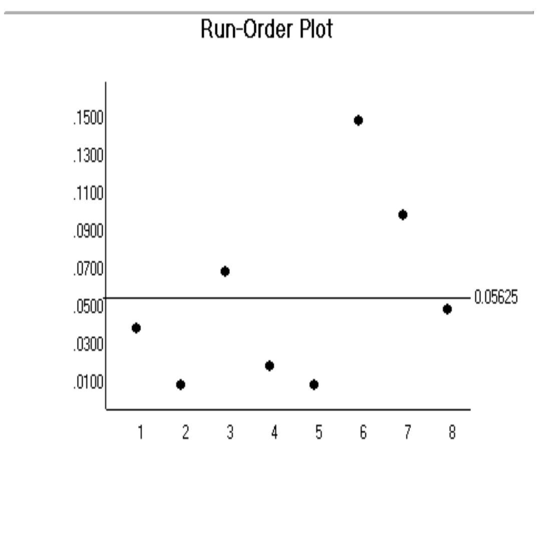
3. **Methods for the analysis of data will be almost exclusively graphical, with minimum aggregation of the data before graphical display.** The aim of the graphical display will be to visually partition the data among the sources of variation present in the study.

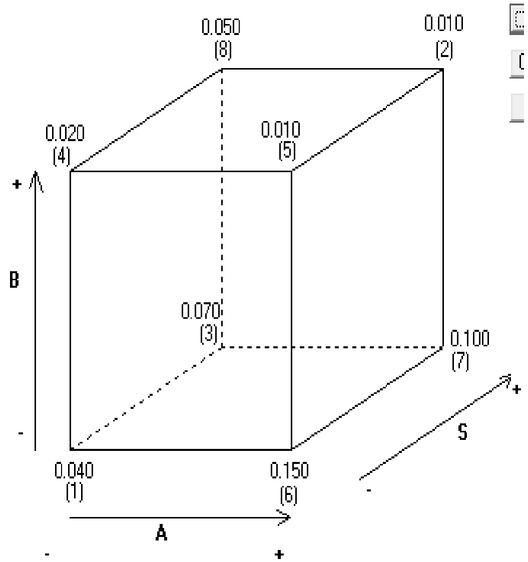
e Half of a 2^4 Factorial Design

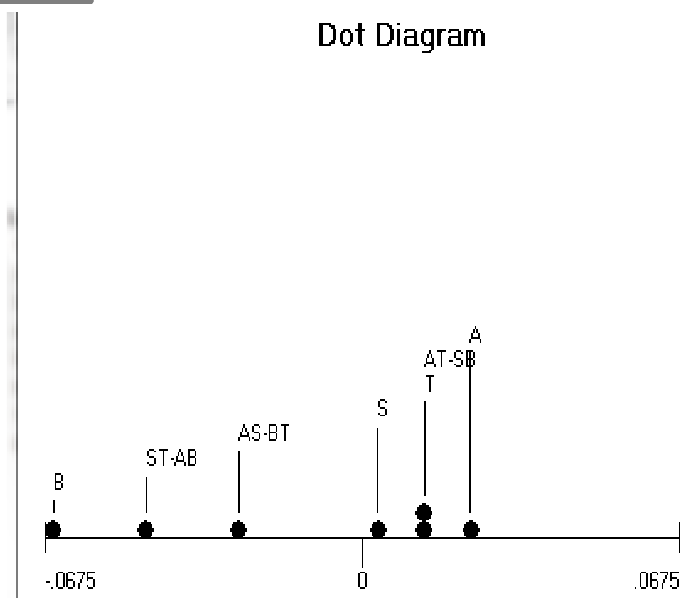
		-	-	-	-				
		+	+	+	+	14	24	34	
Test	RunOrder	A	S	B	T	AT-SB	ST-AB	AS-BT	Response
1	1	-	-	-	-	+	+	+	0.04
2	6	+	-	-	+	+	-	-	0.15
3	3	-	+	-	+	-	+	-	0.07
4	7	+	+	-	-	-	-	+	0.1
5	4	-	-	+	+	-	-	+	0.02
6	5	+	-	+	-	-	+	-	0.01
7	8	-	+	+	-	+	-	-	0.05
8	2	+	+	+	+	+	+	+	0.01
	Effect	0.0225	0.0025	-0.0675	0.0125	0.0125	-0.0475	-0.0275	



PE analysis is done primarily with graphic displays of the data.



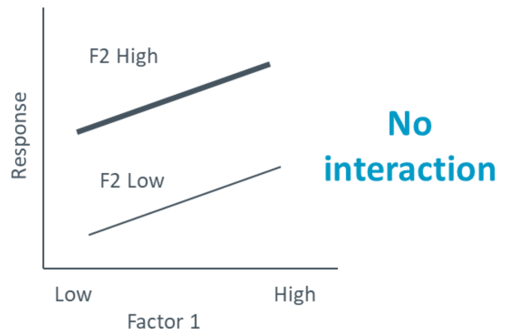


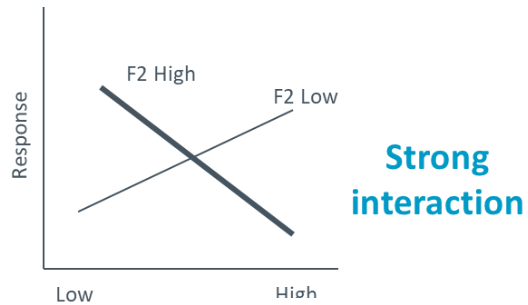


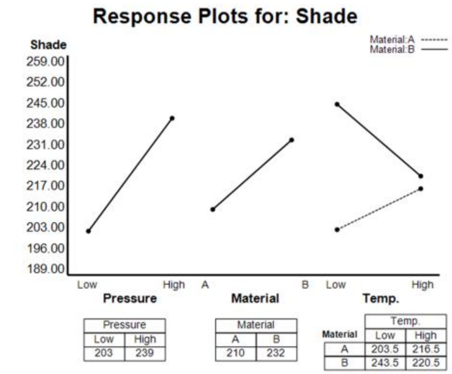
In Summary
Elements Common in the Analysis of Experiments

- **Show all the data before aggregation.**
- **Plot the data in the run order** in which the tests were conducted. This is an important means of identifying the presence of special causes of variation in the data.
- **Rearrange this plot to study other sources of variation (e.g., background variables)** that were included in the study design but are not directly related to the aim of the study. Examples of such variables might be volumes of patients, different protocols, measurement, different shifts or days of the week and environmental conditions.
- **Use graphical displays** to assess how much of the variation in the data can be explained by factors that were deliberately changed. These displays will differ depending on the type of experiment run.
- **Summarize the results of the study with appropriate graphical displays.**

Graphical Displays for Planned Experiments



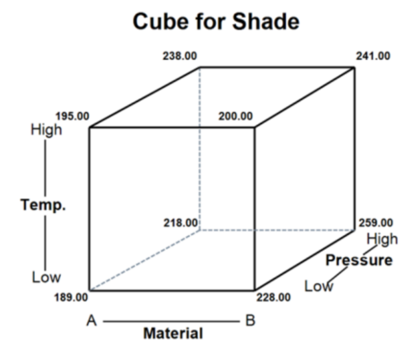


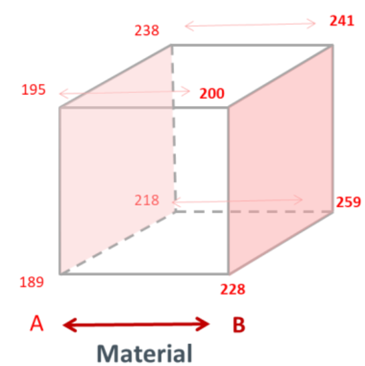


2 Factors, 2 levels each

F2	F1 -	F1 +
-		
+		

2^2 design, 4 runs





TEST	Factor 1	Factor 2	Factor 3
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

2^3 design, 8 tests

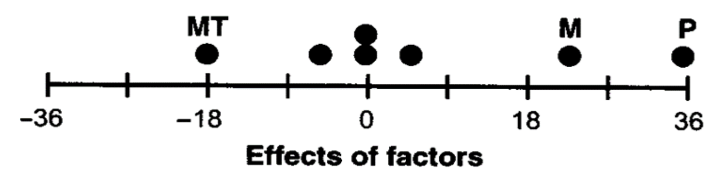
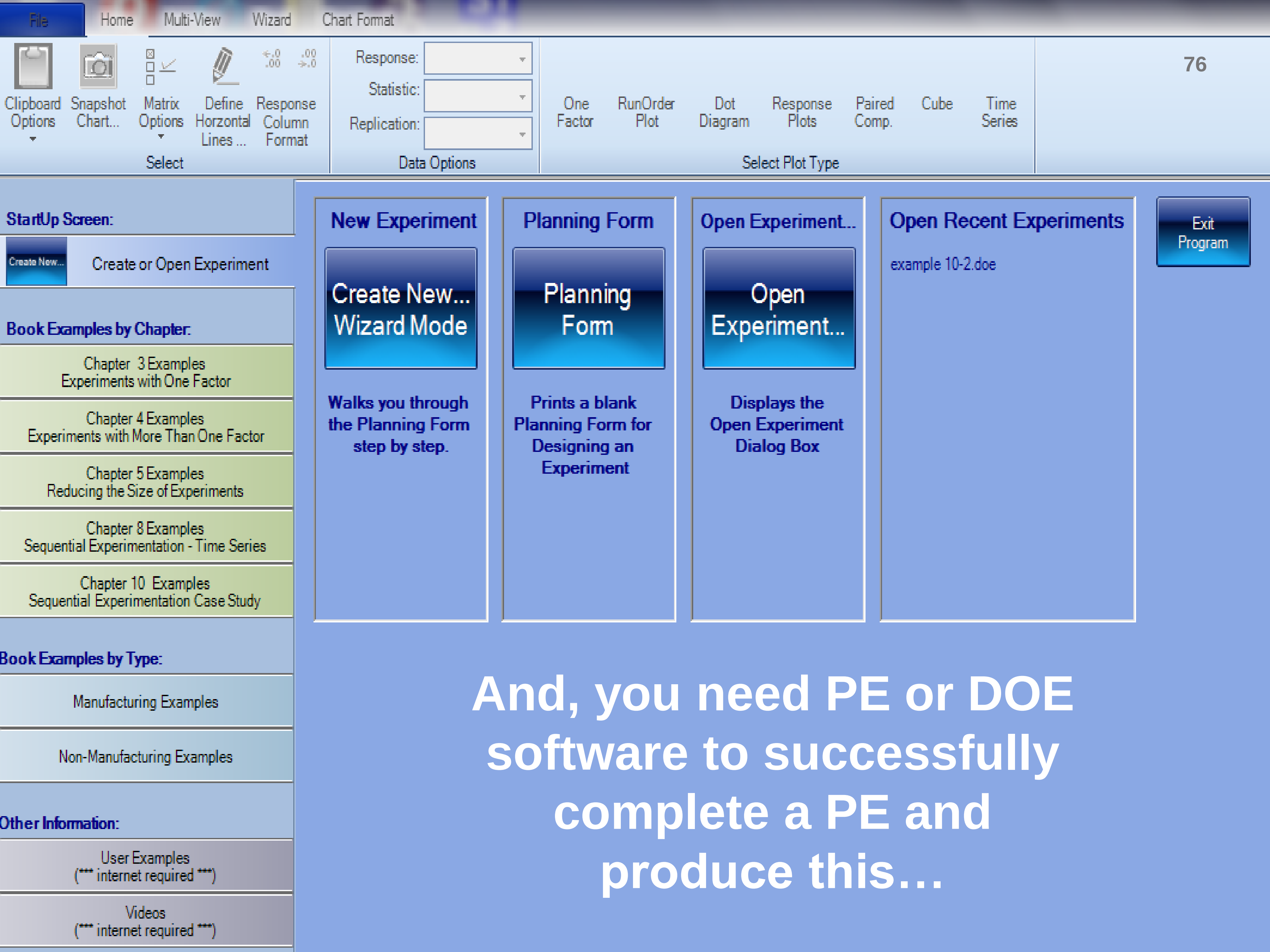
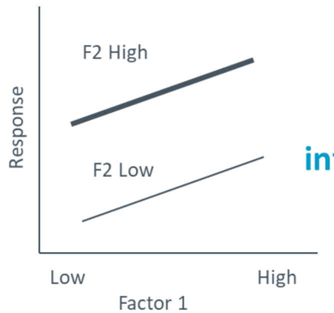


Figure 4.13 Dot Diagram of Effects in Dye Process Experiment

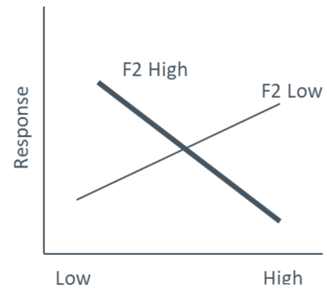


And, you need PE or DOE software to successfully complete a PE and produce this…

Graphical Displays for Planned Experiments



No interaction



Strong interaction

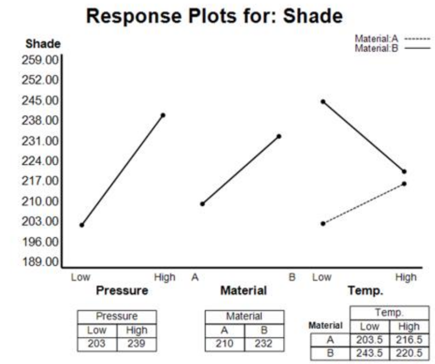
2 Factors, 2 levels each

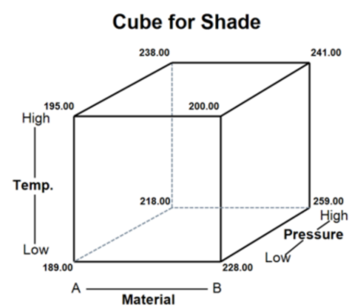
F2	F1 -	F1 +
-		
+		

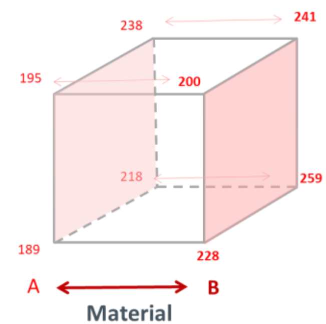
2^2 design, 4 runs

TEST	Factor 1	Factor 2	Factor 3
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

2^3 design, 8 tests







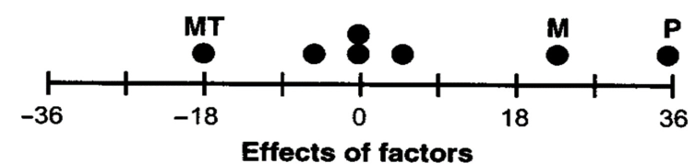


Figure 4.13 Dot Diagram of Effects in Dye Process Experiment

Thank you for joining me today to learn about Planned Experimentation!

Dr Bob

Appendix A
Dr. Robert Lloyd Bio

Robert Lloyd, PhD, Vice President, Institute for Healthcare Improvement provides leadership in the areas of performance improvement strategies, statistical process control methods, development of strategic dashboards and capacity and capability building for quality improvement. He serves as primary faculty for the IHI Improvement Advisor (IA) Professional Development Program, the Improvement Science in Action (ISIA) Program, the Improvement Coach Program and various other IHI initiatives and demonstration projects. Dr. Lloyd works throughout the US, Canada, the UK, Sweden, Denmark, Africa, the Middle East, India, Malaysia, Australia and New Zealand. He is an internationally recognized speaker on quality improvement concepts, methods and tools.

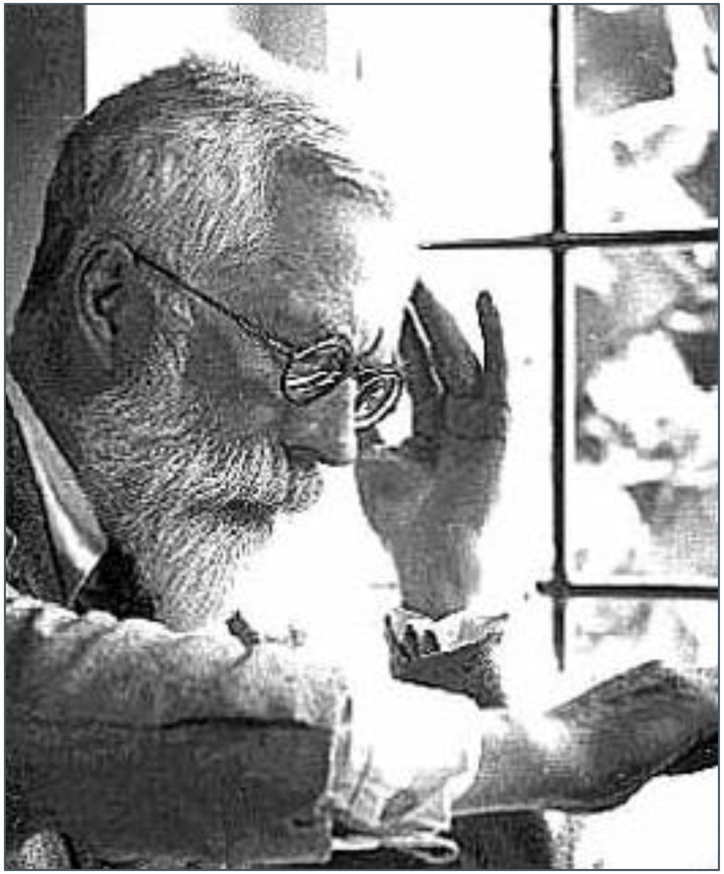
@rlloyd66

rlloyd@ihi.org

He also advises senior leadership teams and boards on how to create the structures, processes and cultures that will make quality thinking and behaviors part of daily work. He is the author of three leading books on measuring quality improvement in healthcare settings and numerous articles and book chapters on quality measurement and improvement.

Appendix B
Principles for Designing Analytic Studies

1. A well defined objective
2. Take a sequential approach
3. Partition the variation
4. Explore the degree of belief
5. Simplicity of execution



Sir Ronald A. Fisher 1890-1962

1. Objective of the study

- The objective of the cycle should be stated in such a way that it provides guidance to those designing the experiment.

- The objective should clarify whether the experiment involves screening a large number of variables to find the most important ones, studying in depth a few variables, or confirming the results of past studies under new conditions.

2. Sequential Approach

- The sequential nature of learning is fundamental to quality improvement. Planned experiments should be designed to learn in a sequential manner.
- The Model for Improvement (Chapter 1) stresses the iterative nature of development of knowledge of the product or process through multiple PDSA cycles.
- Many iterations of the PDSA cycle to develop and test changes will include the design of an experiment.

Strategy for Experimentation:

Current Knowledge	Types of Experiments
Low knowledge	Fractional factorials (screening studies) Nested designs (sources of variation)
Moderate knowledge	Fractional factorials (new levels, new factors) Factorial studies (study interactions)
High knowledge	Confirmatory studies (one-factor with blocks)

p. 35-36

3. Partitioning of variation

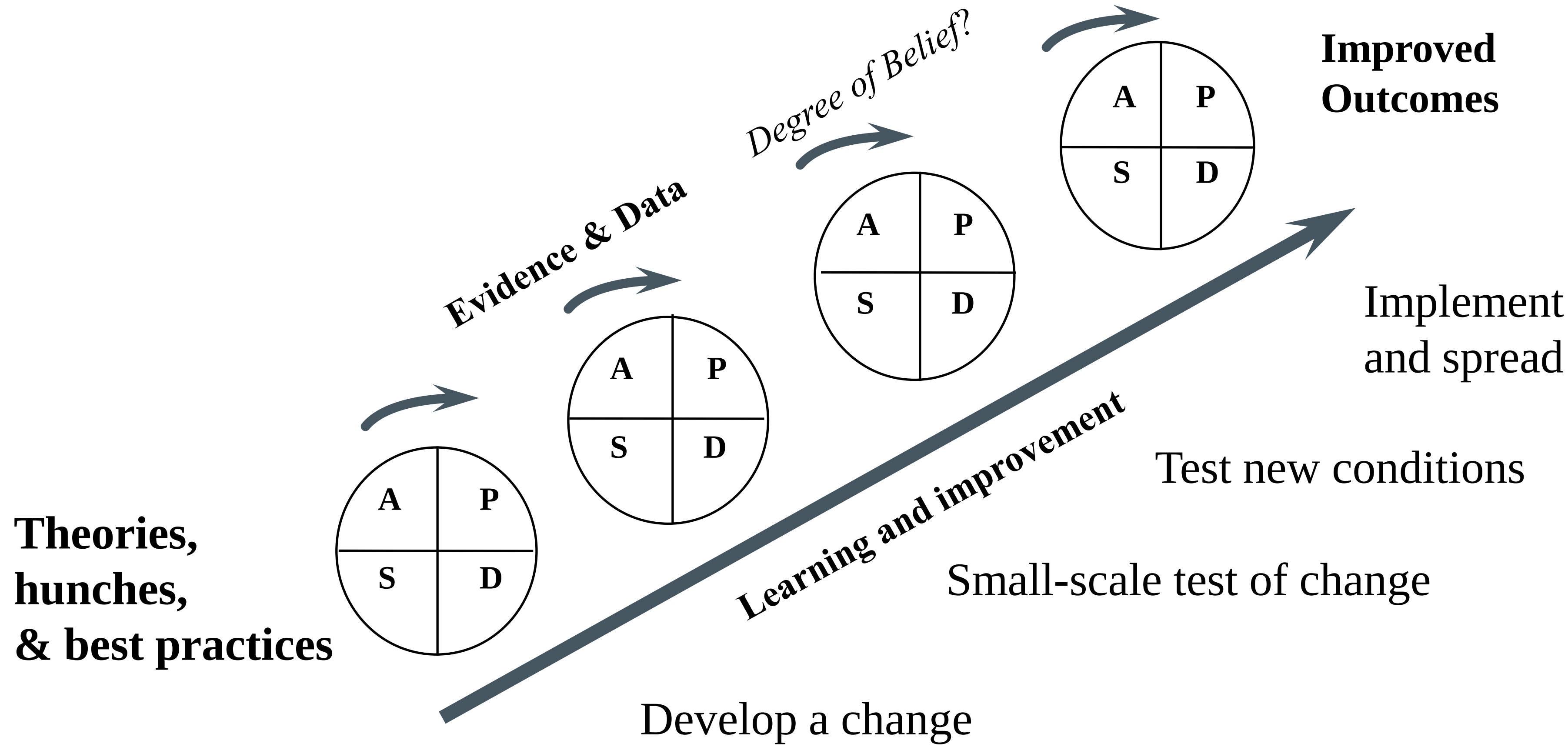
- Determination of important factors and estimation of the effects of these factors on the response variables is usually the objective in any experiment.
- These decisions and estimates should not be confounded by background or nuisance variables. The factors chosen for the experiment are usually those that the experimenter believes will have the greatest effect on the response variable.
- In many experiments, the variation due to background or nuisance variables will be as great as or greater than the variation due to the factors chosen. To help determine if the most important factors have been studied, the experimental design must allow the variation in the response variable to be partitioned into components due to factors, due to background variables, and due to nuisance variables.

p. 37

Improvement: Building Knowledge Sequentially

Sequential building of knowledge

Include a wide range of conditions in the sequence of tests



3. Partitioning of Variation (cont.)

- The ability to partition the variation begins with the selection of the experimental units for the study.
- The experimental unit (EU) for a study is the smallest division of units (structures, subjects, social groups, material, or time) in an experiment such that any two units may receive different combinations of factors and levels.
- The experimental pattern (discussed next) includes deciding which combination of factors and levels will be assigned to each EU.
- Selection of a useful EU is a key role of the subject matter experts associated with the factors of interest and the levels of these factors that are of interest.
- In health care, the EU is often the patient, the provider, the clinic, the hospital, or the region. In education, students, classrooms, teachers, or schools are often EU s.

4. Degree of Belief

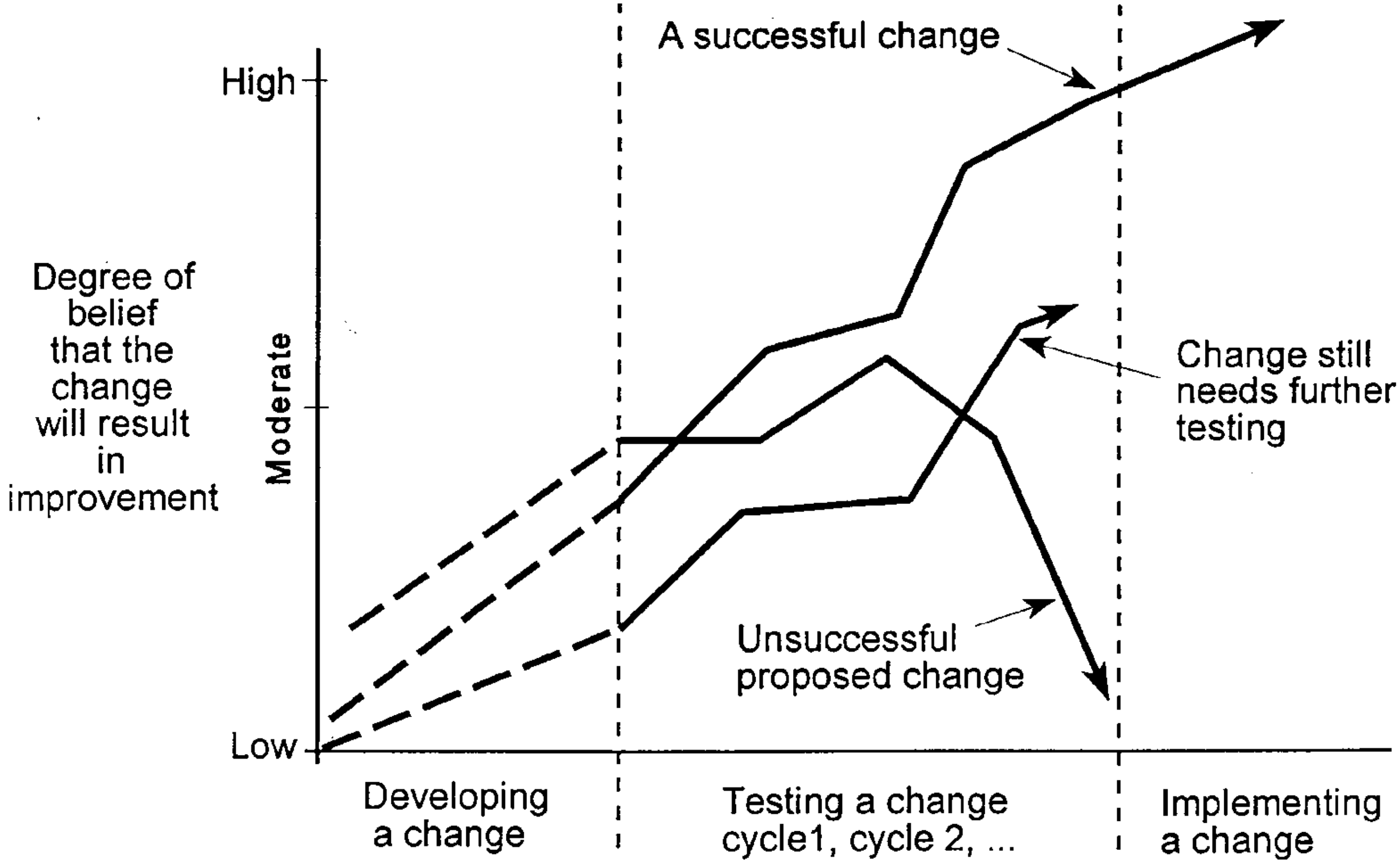
- Most experiments are carried out to determine if a change will result in better performance in the future. The wider the range of conditions included in the experiment, the more generally applicable will be the conclusions from the experiment.

- The degree of belief in the validity of the conclusions is increased by running the experiment using different clinics, different operators, different days, different times of the year, different batches of raw materials, and so forth.

- The range of conditions selected for the study will ultimately determine the degree of belief in the actions taken as a result of the experiment.

- The subject matter expert(s) must determine what is an "adequate" degree of belief for taking action.

Degree of Belief When Making Changes



5. Simplicity of Execution

- The simplicity of the experiment should be considered one of the most important properties of a planned study to improve a system.
- Deming in particular emphasized this aspect of study design. Simplicity is important in the design, the conduct, and the analysis of a planned experiment.
- An experimental design should be as simple as possible while still satisfying the other properties of a well planned experiment.
- Simplicity requires that all the practical aspects of conducting an experiment be considered.
- Some important aspects include the determination of the experimental unit, the difficulty in changing levels of a factor, the ability to control background variables, and the ability to measure important response variables.