

Experiment Planning an

Tab 6:

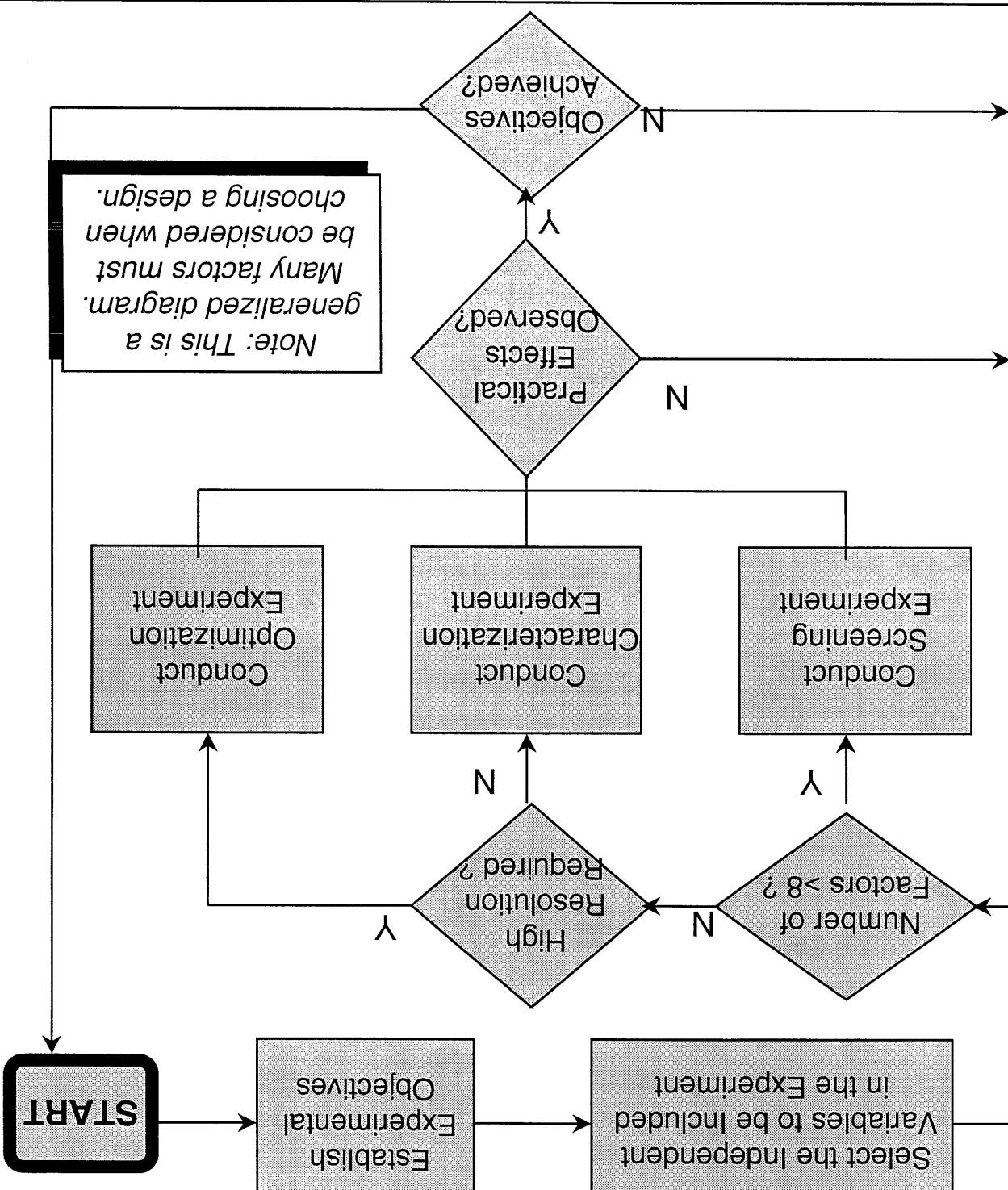
- 2) Recognize potential barriers to effective designs
- 1) Define steps in planning an experiment

Objectives:

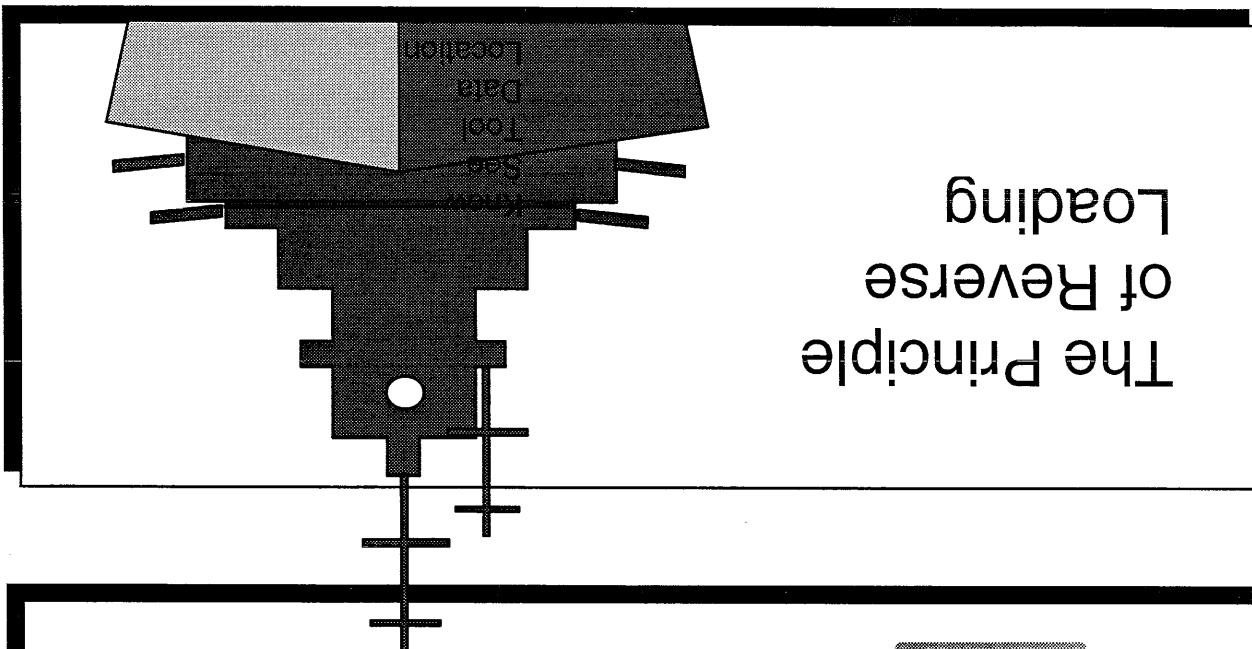
The purpose of this section is to give some practical guidelines in planning experiments. This section is intended to be a "checklist" of items to review before running an experiment.

Purpose:

Tab 6: Planning an Experiment

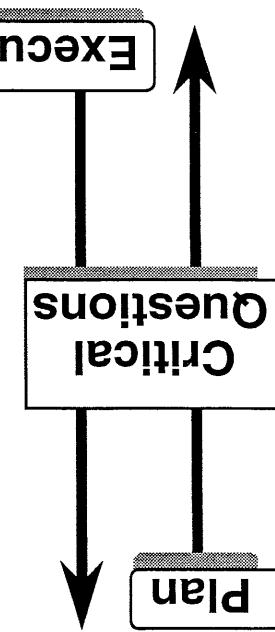


Selecting an Approach



The Principle
of Reverse
Loading

- 1) What do you want to know?
- 2) How do you want to look at what you need to know?
- 3) What type of tool will generate what you need to see?
- 4) What type of data is required of the selected tool?
- 5) Where can you get the required type of data?



The Planning Questions

What do you need at the Improve Phase dock?

Measure Phase Deliverables:

- Problem Statement, including Y response
- Process map of process
- Baseline of Y response
- 4-block for area of focus

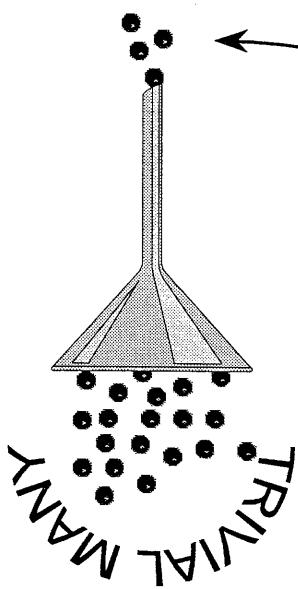
Analyze Phase Deliverables:

- Analysis of baseline data
- Graphs
- F, t and Chi² Tests
- ANOVA
- Suspected Vital X's

Management Team Buy-in:

- Time
- Cost
- Support of DOE Strategy
- Resources

DOE - Planning an Experiment



1. Screening experiments -- used to determine which variables are the most important (many variables, few levels - fractional factorial designs).

2. Detailed study of a few variables -- Use more levels of the independent variables (full factorial designs, response surface designs) to fit a more complex model via regression analysis.

3. Confirming experiments -- used to demonstrate that promising results can be repeated.

Types of Experiments

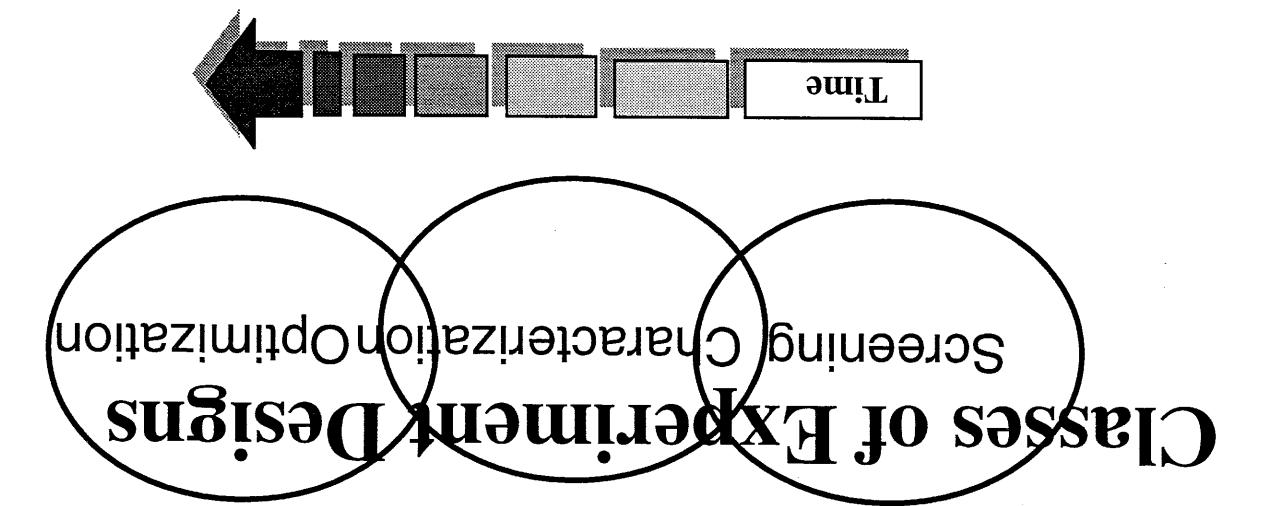
Note: The class descriptions are very general by nature. Such descriptions are intended for instructional purposes only and should not be used as the sole benchmark when designing an experiment.

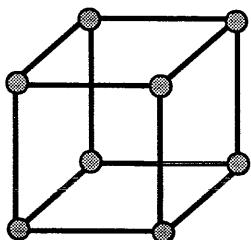
Optimization Designs: Resolution V and higher designs. Most often used to study highly complicated effects and relationships involving two or three variables. In general, such designs are employed to yield a reasonably precise mathematical model of the phenomenon under investigation.

Confounding of certain interactions does not pose a problem.

Characterization Designs: Resolution IV and V designs. Generally used to study relatively uncomplicated effects and interactions resulting from a relatively small number of factors. Generally speaking, such designs are based on the full factorial model using two or three levels. Sometimes, fractionals are employed for this purpose where there is a moderate number of factors to be studied and the confounding of certain interactions does not pose a problem.

Screening Designs: Resolution III designs. Used to isolate the "vital few" variables from the "trivial many"; i.e., those instances involving a very large number of factors. Such designs are most often limited to two levels and are principally used to for the study of main effects.





Example: Start with a 2^2 design (4 runs). If results show that you are not quite in the desired region, choose levels in the desired direction and run another small experiment.

Run an initial design and analyze results. Decide if more experimentation is warranted to test additional regions/variables, or to explore the current region/variable in more detail.

- **Sequential Experiments:**

2. How to conduct the experiments. This can be used to improve your experimentation technique in later experiments.

1. What variables are important. This may give insight into mechanisms that might control the process, and suggest variables to use in later experiments.

It is sometimes better to run a series of small or screening experiments first, rather than relying on one large experiment. In the early experiments you can learn:

- **Preliminary Experiments:**

Iterative Nature of the Experimental Process:

IBM - Retired Statistician
William Diamond

The purpose of an experiment is to better understand the real world, not to understand the experimental data



- # Steps in Planning an Experiment
- 1) Define the Objective
 - 2) Select the 'y'-Response (Dependent) Variables
 - 3) Select the 'x' (Independent) Variables
 - 4) Choose the 'x' Variable Levels
 - 5) Select the Experimental Design
 - 6) Run the Experiment & Collect the Data
 - 7) Analyze the Data
 - 8) Draw Conclusions
 - 9) Perform Confirmation Run

If the answer to # 3 is no, then re-design the experiment.

- 1) What decisions are to be made with the data?
- 2) How will the data be analyzed after it is collected?
- 3) Will the data and analyses allow the required decisions to be made?

the following questions:

While planning an experiment, it is useful to ask

- 2) Estimate the effects of region and % floor loading on sales volume.

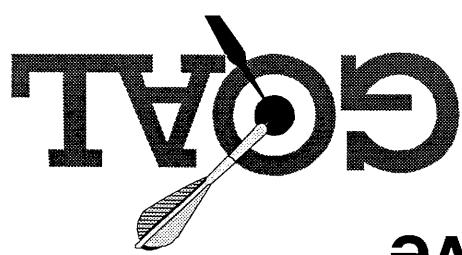
- 1) Estimate effects of time, temperature and cleaner concentration on the response of residue.

Examples:

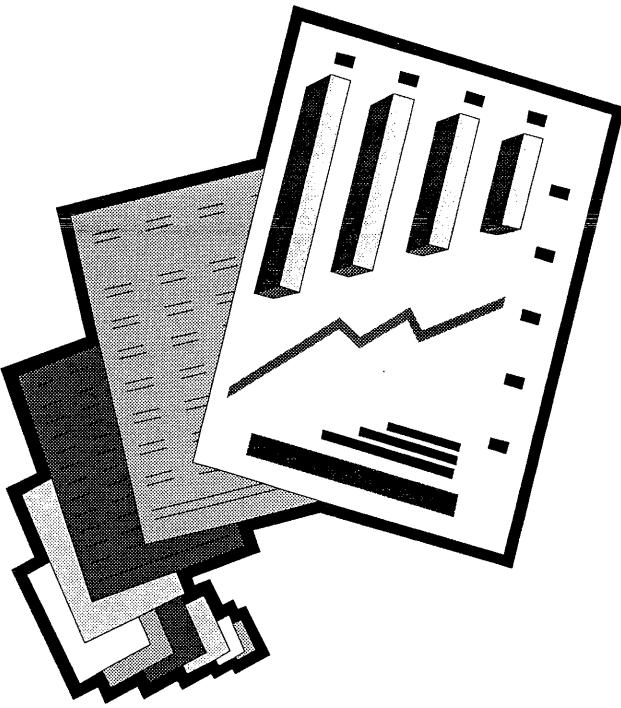
- Estimate the effects of: [list of independent variables]
- On the responses of: [list of dependent variables]

experiment in the form:

State the objective of an



1. Define the Objective



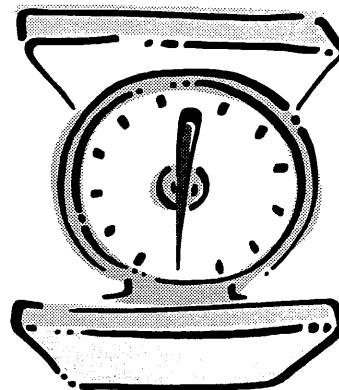
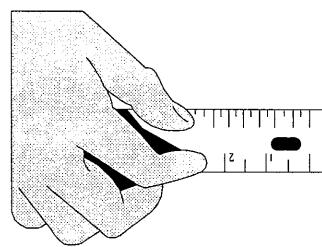
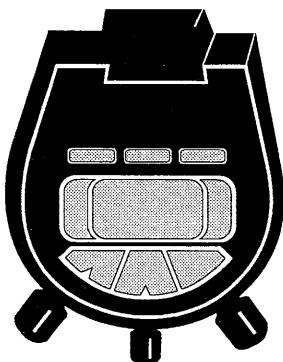
Valid conclusions. However, sometimes it is difficult to get quantitative measurements of the response of interest. It is often possible to measure comparisons or use rankings or standards.

Variable vs. Attribute data

It is possible to measure several dependent (response) variables and model each of these as a function of the independent variables at the same time.

2. Select the 'y', Response (Dependent) Variables

Remember - Gauge R&R is important for both Y's and X's!



Regardless of the type of data you have you must validate your measurement system before running your experiment. If it is unacceptable, it must be improved (reduced) to <20%.

Don't forget Gauge R&R...

- Cause & Effect diagrams
- Baseline data
- Flow charts
- Brainstorming
- Supplier input
- Walk the Process
- Rolled Throughput Yield
- Competitive analysis
- Expert opinion

Select the "X" variables to study through:

- Lurking variables
- Noise variables
- Intentionally changed
- Variables that are monitored, but not
- Variables that are held constant
- Blocking variables
- Design variables, intentionally varied in the experiment
- Design variables, intentionally varied in the

There are many types of "X" variables:

3. Select "X" (Independent) Variables



- There are often many candidate variables.
- Experiments give team members a chance to input and test their variables of interest.
- You can follow-up later with a confirming experiment or a more detailed experiment with fewer variables.

Rationale for studying many variables at once...

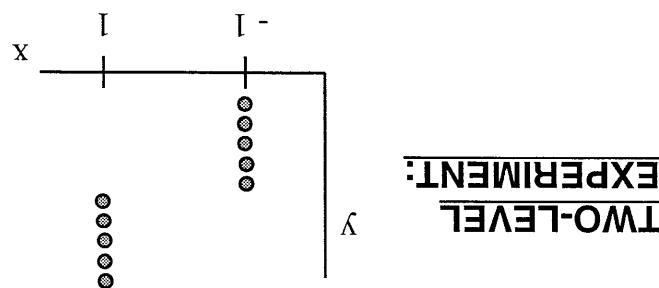
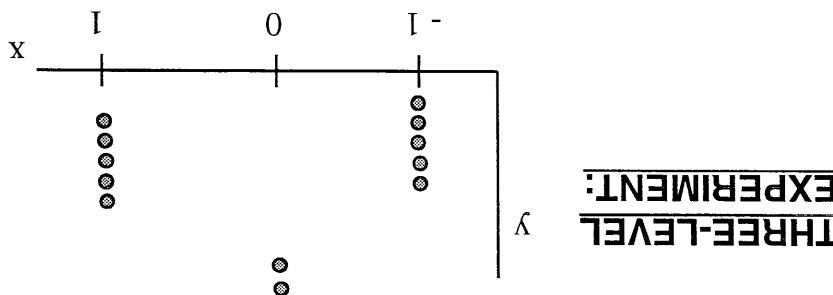
(cont'd)

3. Select 'X' (Independent) Variables



Sometimes you
could be wrong;
team ---
Your advocacy
made by you and
a judgment call
experience is
• Remember,

- Test the most *likely* candidates **first**
 - The challenge is to find the most influential "X"s and the *test range* that makes them visible.
 - Experience shows that **only 2 to 6 variables will end up being the "Vital Few"**
- Selecting the Right "X" Variables:**
- **Keep the design *simple* unless little is vital Few** (screening experiment) tests many variables to narrow in on the known about the process. In that case,
 - **Remember,**
experience is
made by you and
a judgment call
your advocacy
team ---
could be wrong;

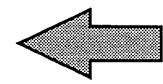


- If you know the variable is Vital and want to study its effect in detail, typically more than two levels are used.
- If you can't fit a line through the data (a potentially non-linear response), don't model with a line.
- If you are searching for the important variables with a screening experiment, typically two levels are used.

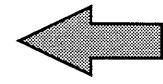
The number of levels you choose depends on the objective of the experiment and the graph of the response.

4. Choose Levels for the "X" Variables

Realize that some combinations in the test will produce unacceptable responses, and that these results are expected and desirable.

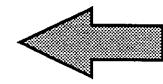


Understand that levels must not be beyond the range of feasibility (but they may be beyond current process range).



Example: Changing temperature by one degree will likely have little effect on the outcome, and you might incorrectly conclude that temperature is not an important variable.

The range must be wide enough to show a difference, IF there is a difference.



When choosing the Variable Levels...

- a) Orthogonality
- b) Randomization
- c) Replication
- d) Repetition
- e) Controls
- f) Lurking Variables
- g) Noise Variables
- h) Blocking
- i) Sample Size
- j) Confounding

*When designing the experiment,
consider 10 key elements...*

5. Select the Experimental Design

- Reduces the variability of an estimate (shorter same levels)
 - Gives an estimate of variation and confidence confidence intervals)
 - in the results.
- c) Replication** (completely re-setting the experiment and obtaining additional results at the same levels.

- Randomize:**
- Run order
 - Assignment of experimental units
 - Measurement order
- statistical tests of significance are valid.
- b) Randomization** to reduce the effects of extraneous variables, and ensure that the fractional factorial test plans) are used to separate the effects of the variables.

- a) Orthogonal arrangements** (fractional and fractional factorial test plans) are used to separate the effects of the variables.

10 Key Elements of Experimental Designs

- e) **Controls** and reference distributions. Most experiments are comparative. Including a control or baseline can be extremely useful.
- f) **Lurking Variables** are variables that can affect the results but are unknown, uncontrollable, uncontrolable or un-measurable.
- The influence of lurking variables can often be reduced by blocking and randomization.

(cont'd)

10 Key Elements of Experimental Designs

	Time	Gallons	Cotton	Cotton	Blend	Blend	Fabric Type	Load Size
(Test Variables)	6	10	6	10				
	X	X	X	X	X	X		
	X	X	X	X	X	X		
	X	X	X	X	X	X		
	X	X	X	X	X	X		

We could set up the experiment like this:

We want a washer design that works well for all fabrics and all loads because we cannot predict or control this in our consumer's homes. 'Time' and 'Gallons' are test variables. 'Load Size' (6 or 10 lbs) and 'Fabric Type' (Cotton or Blend) are noise variables.

Example:

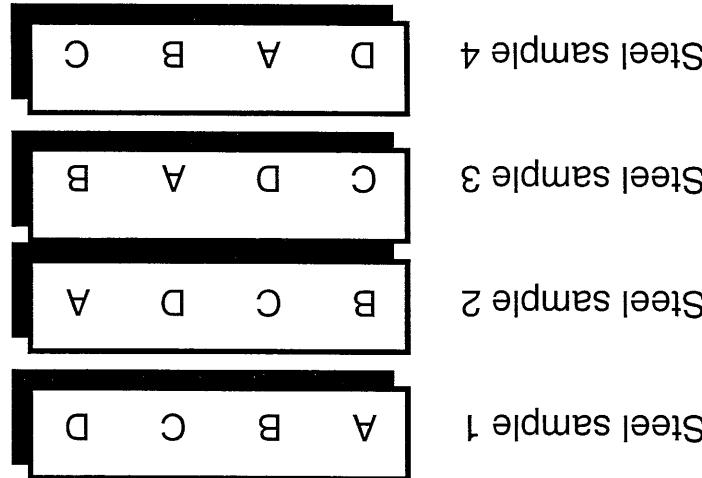
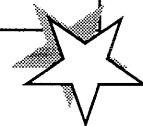
To reduce the effect of noise variables, select test variables that can be monitored across all levels in the experiment.

g) **Noise Variables** are known (or thought) to affect the results, but we either cannot or choose not to control them. Examples: humidity, day of the week, competitor incentives.

(cont'd)

10 Key Elements of Experimental Designs

A blocking variable is generally a variable that "gets in the way." You could consider it just another independent variable . . .



of steel.

Blocking example 1: Test 4 types of coatings (A, B, C, D) on 4 samples

(it is often good to block on time or run order)

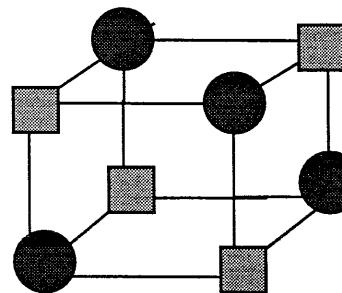
- Reduces variation and gives higher precision of the estimates
- Gives some protection against confounding and lurking variables
- Gives all independent variables an equal chance (a fair test)

Benefits of blocking:

Compare independent variables within blocks, not between blocks.
A 'block' is group of homogeneous units.

h) Blocking (cont'd)

10 Key Elements of Experimental Designs



Each block is a 2^3-1 fractional factorial experiment.

Runs with $X_1 * X_2 * X_3 = 1$ become block 1, and are run on day 1
 Runs with $X_1 * X_2 * X_3 = -1$ become block 2, and are run on day 2

	\bar{X}_1	\bar{X}_2	\bar{X}_3	$\bar{X}_1 * \bar{X}_2 * \bar{X}_3$	Block	
1	1	1	1	1	1	2 - Wed
1	1	1	-1	-1	1	Mon
1	-1	1	-1	-1	1	Mon
1	-1	1	1	1	2	Wed
-1	1	-1	1	-1	1	Mon
-1	1	-1	1	1	2	Wed
-1	-1	1	1	1	2	Wed
-1	-1	1	-1	-1	1	Mon
-1	-1	-1	1	1	1	Mon
-1	-1	-1	-1	-1	1	Mon

2^3 experiment, 8 runs, 2 blocks (days) with 4 runs per block

Scenario: Take samples on Monday and Wednesday at a call-taking center, in order to capture the effects of day of the week.

Blocking example 2: Block on time or run order

	A	B	C	n / Cell
32 samples	8	+	+	+
	8	+	+	-
	8	+	-	+
	8	+	-	-
	8	-	+	+
	8	-	+	-
	8	-	-	+
	8	-	-	-

shown below:

will need to measure 8 samples at each level combination as see that $n = 33$. Since we have set up a 2^3 design (8 Runs), we shift of $\delta = 0.8$. From the sample size table ($\alpha = .05$, $\beta = .10$) we Example: You have a distribution with $G = 1$ and want to see a

in the Appendix of this tab)

(See the Bill Wunderlin paper on "Determining Sample Size",

sample size

- The level of confidence to determine the appropriate
- The size of difference that is important
- Variability

You must consider:

when deciding how many samples to take
There is usually a trade-off between cost and precision

i) Determining Sample Size

(cont'd)

10 Key Elements of Experimental Designs

A Basic Sample Size Table

Applies to continuous data only

$$\frac{Q_0}{Q} = \frac{20\%}{10\%} = \frac{5\%}{1\%} = \frac{10\%}{1\%} = \frac{5\%}{1\%} = \frac{10\%}{1\%} = \frac{5\%}{1\%} = \frac{10\%}{1\%} = \frac{5\%}{1\%} = \frac{10\%}{1\%} = \frac{5\%}{1\%}$$

DOE - Planning an Experiment Rev. 7 March 27/98

Choose a Fractional Factorial Design that minimizes confounding so that Main Effects and important interactions can be separated

Main Effects are usually more important than interactions. Interactions can be important, and must be investigated; however,

$X_1 * X_2$ interaction.

An apparent X_3 effect could also be due to an

Run	X_1	X_2	$X_1 * X_2$
4	1	1	1
3	1	-1	-1
2	-1	1	-1
1	-1	-1	1

interaction.

Example 2: A main effect confounded with an

Run	X_1	X_2	Did X_1 or X_2 cause the effect?
2	high	high	
1	low	low	

Example 1: Changing 2 variables together.

fractionals have some degree of confounding. interactions. Remember that all fractional effects of variables from one another, or from effects of variables from another, or from interactions. Remember that all fractional

(cont'd)

10 Key Elements of Experimental Designs

- A „center point“ is often included in a DOE. Typically the center point is the 2 level factorial design allows the inclusion of one or more center points in a normal operating condition.
- Use of a center point allows the „interior space“ (of the cube) to be investigated. The center point may be replicated without destroying the balance of the design.
- The number of center points will not affect the estimation of main effects or interactions.
- The number of center points will not affect the precision of effect estimates.

Use of a Design Centre Point

The Notion of Design Resolution

Resolution III:

No main effect is aliased with any other main effect
Main effects are aliased with second order interactions
Second order interactions are aliased with other second order interactions

No main effect is aliased with any other main effect
No main effects are aliased with any second order interactions
Second order interactions are aliased with other second order interactions

Resolution IV:

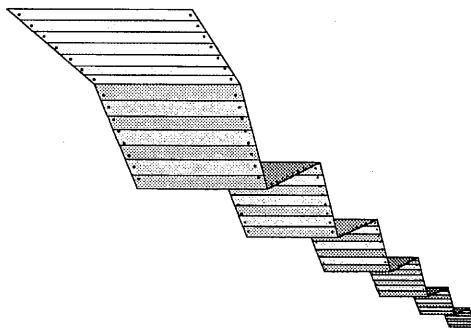
No main effect is aliased with any other main effect
No main effects are aliased with any second order interactions
Second order interactions are aliased with other second order interactions

Resolution V:

No main effect is aliased with any other main effect
No main effects are aliased with any second order interactions
Second order interactions are aliased with third order interactions

6.29

If you want your results to be generic and to apply over a wide range of conditions, then test over a wide range of conditions.



Demonstrating: To demonstrate success, you need repeated tests over time, with no failures.

Example: Sample 4 parts per day for 10 days, rather than 40 consecutive parts on one day.

- a lab or pilot plant (but not production)
- one product line
- one telephone operator
- one batch of raw materials

Your results must apply beyond:

When selecting an experimental design, remember...

- It is often useful to retain samples, which can be measured again later if there are questions about a particular measurement.
- Be sure to be there during the experiment - you never know what may happen!
- Be sure to set up your data sheet before running experiment to be sure all values are recorded in the desired form.

6) Run the Experiment & Collect Data

C) ANOVA tables (Session window) - look for low P-values that indicate significant factors.

B) Confidence intervals - valuable because they give a range of plausible values.

- Contour plots - for Response Surface designs

- Interaction plots - for both mean and standard deviation

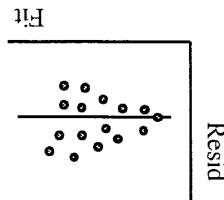
- Main effects plots - for both mean and standard deviation

- Cube plots

- Scatter plots

- Histograms

A) graph GRAPH GRAPH



7) Analyze the Data

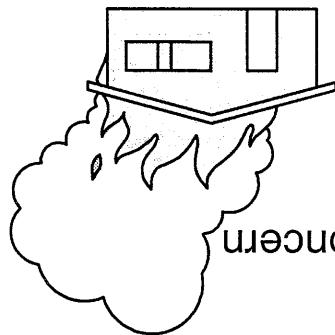
- Should we run additional experiments?
 - Does the experiment direct you towards how to fix the problem?
 - Are the results practically significant?
 - Is there evidence of change?
 - Are the results statistically significant?
- Questions to be answered:**

8) Draw Conclusions

- Be sure to allow the Vital Few X's to naturally vary within the levels set during your confirmation run...but again, do not tweak the process. You want to be sure to capture the natural variation of the process to verify that you made an improvement.
- The Confirmation Run is necessary to verify that you have truly made improvements. Confirmation Runs should be set up in rational subgroups, similar to process baselining. In effect, you are 're-baselining' your process at the new settings.

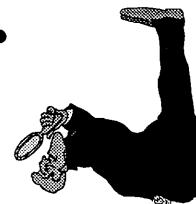
9) Perform Confirmation Run

- You cannot detect the effect of a 2 degree change if you can control only within +/- 1 degree



- Be sure the changes do not destroy the equipment, or create a safety concern
- variables are accurate
- Make sure that the levels of the independent

- Be present during the test run



WORDS OF WISDOM FROM PEOPLE IN PREVIOUS CLASSES

- High measurement variation -- cannot detect a change (Ex: washer, turnover) - the CU test is variable and must be repeated many times).
- Did not include important independent variables -- a series of changes were made, success was assumed.
- Tested over too narrow a range: favorable test results do not repeat. If you want your results to be generic and apply over a wide range of conditions, then test over a wide range of conditions.
- No controls, no way to determine if something else changed -- a "lurking" variable (Ex: compressor noise high voltage on test stands).
- Test that does not match usage conditions (Ex: Burcill defrost heater - the heater continues to degrade when it is off).
- Incorrect assumption about the "acceleration factor" in an accelerated life test (Ex: rotary compressor).

by Bill Wunderlin

WHY EXPERIMENTS FAIL

- Orthogonality
- Lurking Variables
- Randomization
- Noise Variables
- Blocking
- Replication
- Sample Size
- Confounding
- Control

designing an experiment:

- There are 10 key elements to consider when

- 1) Define the Objective
- 2) Select the "y"-Response (Dependent) Variables
- 3) Select the "x" (Independent) Variables
- 4) Choose the "x" Variable Levels
- 5) Select the Experimental Design
- 6) Run the Experiment & Collect the Data
- 7) Analyze the Data
- 8) Draw Conclusions
- 9) Perform Confirmation Run

- There are 9 steps to planning an experiment:

Key Concepts: Planning an Experiment

- Conditions or unplanned events.
- Be present during the trial; document match run sequence.
- Prepare logical, clear data sheets to handle changes of ambient conditions♦ Changes of test equipment♦ Effect of time trends♦ Difficulty of setup change
- Consider these factors:
 - Random or "Constrained Random",
the following:

■ Prepare run sequence considering

effects.

and adjust test plan to minimize the

■ Anticipate potential problems with trial

■ Preplan the trials

Trial Advice

- PHASE 2**
- The Missile**
- BM (Ballistic Missile) Range Optimization**
- Design of Experiments Exercise**
- Fractional Factorial Experiments**
- Rev. 6 August 1, 1997
1. Multiple teams, with approximately 7 people per team
 - A) Yellow missiles (known in military parlance as "soft & cushy", dependent upon our ability to optimize the travel distance of the southwestern desert of the United States. A lucrative government contract is dependent upon our ability to initial impact at a launch point to initial impact. This distance is measured by a complex, retractable metallic measurement instrument, known in layman's terms as a "tape measure". Your team's mission, should you decide to accept it, (which would be a REALLY good plan, in this case) is to find the centerline as a variable that will shoot the BM the required distance, which is 48 inches (measured from the leading edge of the catapult base), with a tolerance of +/- 6 inches.
 2. Objective: optimize the firing range of ballistic missiles (BMs).
 - A) Yellow missiles (known in military parlance as "soft & cushy", urgent delivery, or SCUDs). COST: \$1,000,000 per launch
 - B) Green missiles ("seek & point", urgent delivery, or SPUDs)
 - C) Blue missiles (known in military parlance as "hard & crusty", timely delivery, or CUDs). COST: \$1,000,000 per launch
 - D) Red missiles (known in military parlance as "hard & crusty", timely delivery, or RUDs). COST: \$1,000,000 per launch
 3. Situation: GE has recently bid on a contract with a top secret military installation somewhere in the vicinity of Area 51, somewhere in the southwestern desert of the United States. A lucrative government contract is dependent upon our ability to optimize the travel distance of the southwestern desert of the United States. A lucrative government contract is dependent upon our ability to initial impact at a launch point to initial impact. This distance is measured by a complex, retractable metallic measurement instrument, known in layman's terms as a "tape measure". Your team's mission, should you decide to accept it, (which would be a REALLY good plan, in this case) is to find the centerline as a variable that will shoot the BM the required distance, which is 48 inches (measured from the leading edge of the catapult base), with a tolerance of +/- 6 inches.

A. Brainstorm Potential Variables that could influence the launch distance - agree as a team to the top 5 (or fewer) on which to run the experiment. NOTE: You have worked with this missile launching system in the past, but you may want to baseline this process a second time. Since the government has agreed to pay for the baseline - regardless of sample size (translation: you don't have to count baselining missing missile shots). Remember, baseline is the foundation of our understanding of the process. Since the baseline is the foundation of our understanding of the process, the experiment - agree as a team to the top 5 (or fewer) on which to run the experiment. NOTE: You have worked with this missile launching system in the past, but you may want to baseline this process a second time. Since the government has agreed to pay for the baseline - regardless of sample size (translation: you don't have to count baselining missing missile shots). Remember, baselining means characterizing your process as it stands today, WITHOUT changing the process settings!

- J. Once the optimized launcher setting/missile configuration has present their findings to the class in a formal 10-minute presentation, using charts and whatever backup data is deemed appropriate by the team. A second confirmation run (consisting of 10 launches) will be done in front of the class, and will be part of the overall team rating (a Z-fit value will be calculated from these 10 shots). The winning team will be determined their performance during the second confirmation run. A standard rating form will be used.
- I. The data analysis and presentations should include plenty of graphs. Have all graphs available at the presentation, so you will be able to answer questions.
- H. Create a standard "data recording sheet" before running the experiment.
- G. Randomize the order of the runs.
- F. Describe the conounding pattern. Determine if it will be possible to estimate and interpret interactions.
- E. Choose one or more individuals on each team to fill the following roles:

1. Catapult operator
2. Distance judge #1
3. Distance judge #2
4. Distance judge #3
5. Data recorder

- E. Choose one or more individuals on each team to fill the following roles:
- D. Determine if blocking variables should be included. combined is 16.
- C. You have a total of 16 experimental runs that you are allowed to make. Design your 2-level experiment(s) so that the maximum total number of runs is 16. You may decide to perform multiple experiments (fractional and/or full factorial designs), as long as the total number of runs in all experiments combined is 16.
- B. Choose levels to test for the independent variables.
- A. Other creative designs) WILL COUNT towards your total dollars expended or responsibility (your conscience). Shots using pseudo-missiles (rolled-up tape or testing. One of the reasons we are being considered for this contract is the level of respect for our integrity policy . . . keeping track of all shots made is YOUR responsibility (your conscience).
- Every shot you make (after baselining) must be accounted for in order for GE to be reimbursed by the government for the cost of the preliminary testing. One of the reasons we are being considered for this contract is the level of respect for our integrity policy . . . keeping track of all shots made is YOUR responsibility (your conscience).

[s:\6_\sigmabottleindoeplns.doc]

and its cost will be deducted from your paycheck.
 their original design! If you do, you will destroy the missile launching station,
 multiple bands on the same launcher, or twist the bands in any way other than
 in creating the optimized settings, you may NOT under any circumstances use

answer questions.

Have back-up charts available with more detail, so you will be able to

5. The results of the confirming experiment.
 4. Recommended running conditions (levels of the independent variables).
 3. The one or two graphs of the results that are most enlightening (label the graphs well!).
 2. A listing of the test runs.
 1. The objective and the objective of this experiment. Indicate the independent and dependent variables.
- Be sure to show the following information in your presentation:

6.38

Appendix

1. Objective
2. Background information
3. Experimental variables:
 - A. Response Variables
 - B. Factors under study
4. Number of Replications:
5. Design matrices: (attach copy)
6. Data collection forms: (attach copy)
7. Planned method of statistical analysis:
8. Estimated cost, schedule and other resource considerations:

DOE Planning Worksheet

Designed experiments planned and executed well are a powerful tool to drive improvement yet done incorrectly are time consuming and expensive!

- Blocking
- Repetition and replication
- Sample size table
- Variables and variable levels
- 10 Step planning tool
- Characterization, optimization
- Design choice - screening,
- Attention to detail

DOE Planning Pointers