Design of Experiments
Part 1

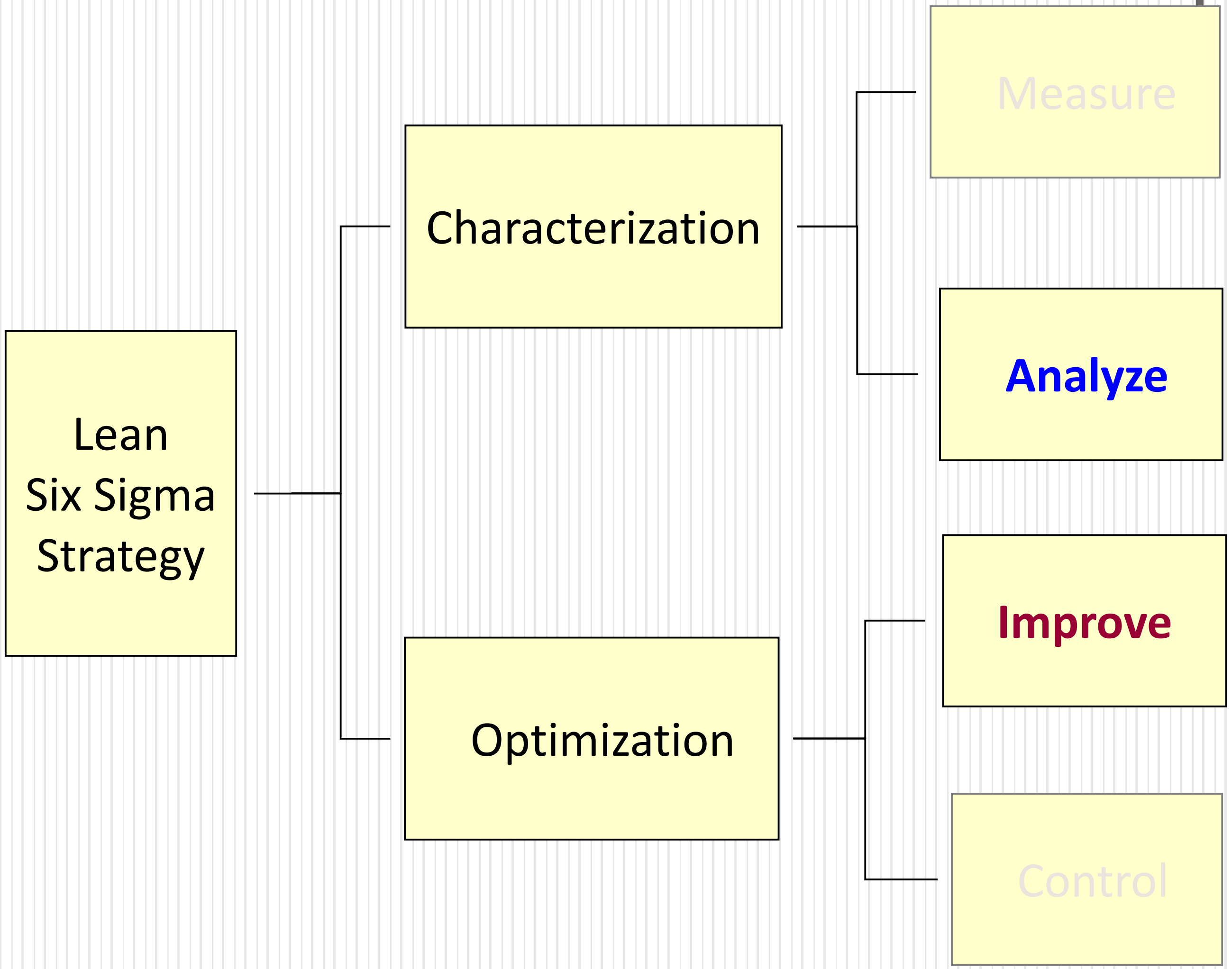
Dr. Bob Gee
Dean Scott Bonney
Professor William G. Journigan
American Meridian University

Learning Objectives

Upon successful completion of this module, the student should be able to:

- What is a Designed Experiment (DOE)?
- Why use DOE?
- Components of a DOE
- Full Factorial Experiment
  - Factor and Interaction Effects
- Steps for Designing and Conducting a DOE
  - DOE Planning Worksheet
  - Types of Outputs
  - Selecting a Good Output
  - Repetition and Replication
  - Doe Strategies
- Types of Experiments*
  - Fractional DOE
  - DOE for Variation

DMAIC Roadmap:
What is DOE & Where Does it Apply?



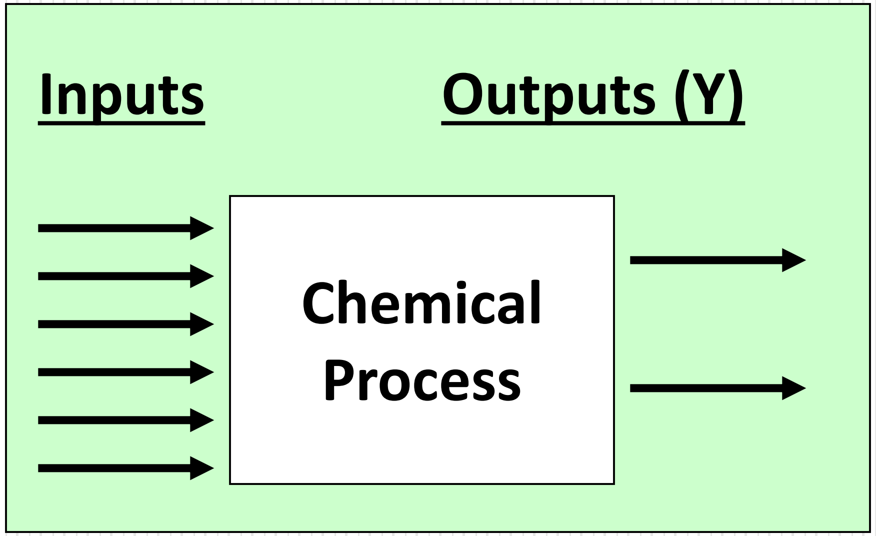
Fractional DOE used for <u>screening</u> a large number of variables for significance

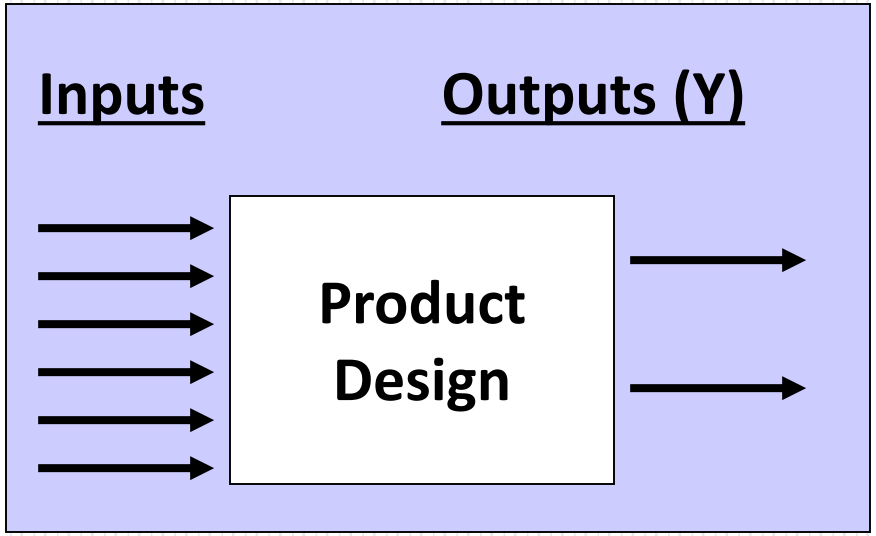
Fractional DOEs evaluate main factors and interactions & provide prediction equation <u>Characterizing</u> impact of vital input factors on output.
Full factorial DOE used to refine inputs for <u>Optimal</u> output

What is a Designed Experiment?

Deliberate and systematic changes of input variables (factors) to observe the corresponding change in the output (response)

Specifically, DOE is a way of organizing the way in which one changes one or more **input factors** (**X**) to see if one *or any combination* of those factors affect the **output variables** (**Y**) in a significant way.





Why Use Designed Experiments?

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

– William Diamond

IBM Retired Statistician

Benefits of Experiments

- Statistical experiments are rich in benefits:
  - Short time to get results
  - Relatively low cost
  - Excellent chance of detecting optimal variable levels
  - Very high confidence in results*
  - Able to identify both main effects and significant interaction effects

***Note: High confidence due to controlled data collection allows the experimenter to accept slightly higher risk. Industry standard is to accept α = 0.10 (10% alpha risk = 90% confidence) when evaluating DOE factors for significance.**

Definitions & Examples

Components of DOE	Examples
Output Response The output that is measured as the result of the experiment and is used to judge the effects of factors. **NOTE: SHOULD BE VARIABLE DATA**	Part Dimension (in mm) Product Yield (in pounds) Prototype Cycle Time (in minutes) Product Quality (% Defective*)
Factors The process input variables that are set at different levels to see their effect on the output. **NOTE: MAY USE ATTRIBUTE OR VARIABLE DATA**	A. Cycle Time (Days) B. Mold Temp. (°C) C. Tank Pressure (PSI) D. Holding Time (Days) E. Material Type (Descriptive Attribute)
Levels The values at which factors are set.	<u>Factor</u> — <u>Levels</u> B. Mold Temp. — 200° — 400° D. Holding Time — 1 Day — 14 Days E. Mat'l. Type — Plastic — Aluminum
Interactions The degree to which factors depend on one another. Some experiments evaluate the effect of interactions.	Training Time x Experience: The best level for training time depends on previous operator experience.
Experimental Trials or Runs The specific test combinations of factors and levels that are run during the experiment.	Factors — A — B — C Runs 1 — -1 — -1 — +1 2 — +1 — +1 — -1 3 — +1 — -1 — +1

General Example -Setting the Factor Levels

Factor	Process Variable	Levels	Code
A	Lemon Flavor	Fresh Squeezed	+1
		Concentrate	-1
B	Sugar	Refined	+1
		Raw (Unrefined)	-1

General Example – Conducting the Experiment

INPUT DATA (FACTORS) MAY BE VARIABLE OR ATTRIBUTE

Sequence Order	Process Variables		Response
	A	B	
1	-1	-1	2.8
2	+1	-1	4.8
3	-1	+1	3.5
4	+1	+1	3.2

OUTPUT DATA (RESPONSE) MUST BE VARIABLE!

Full Factorial Experiment

- A full factorial experiment is any experiment in which all possible combinations of factor levels are tested.
- This chemical yield experiment is a simple example of a full factorial design:

	Factors	
Runs	A	B
1	-1	-1
2	-1	+1
3	+1	-1
4	+1	+1

Coded Data
(How Minitab sees it)

	Factors	
Runs	Pressure	Catalyst
1	1 ATM	A
2	1 ATM	B
3	4 ATM	A
4	4 ATM	B

Uncoded Data
(How we see it)

Would you rather see your DOE as Minitab sees it (left) or as your team members see it (right)? As a rule, we speak with Uncoded data. However, always remember that Minitab performs its calculations with Coded data.

Full Factorial Experiment

- How many factors are being evaluated?
- How many levels are there for each factor?
- Is this a full factorial DOE?
- Note that chemical yield is measured in grams of product

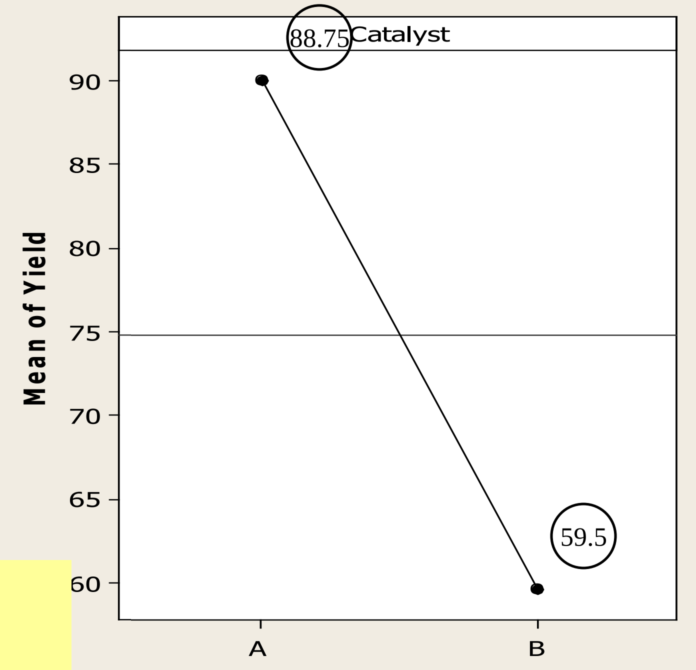
Pressure	Catalyst	Yield 1	Yield 2
1 ATM	A	65	85
1 ATM	B	55	63
4 ATM	A	98	107
4 ATM	B	68	52

Note

This simplified example is used to illustrate the concepts of factor and interaction effects. Scenario is from Minitab DOE Tutorial.
Notice that we *replicated* each experiment two times, for a total of eight data points. In the real world, we strongly encourage at least three replications so that variation may be better understood.

Evaluating Main Effects

Pressure	Catalyst	Yield 1	Yield 2
1 ATM	A	65	85
1 ATM	B	55	63
4 ATM	A	98	107
4 ATM	B	68	52



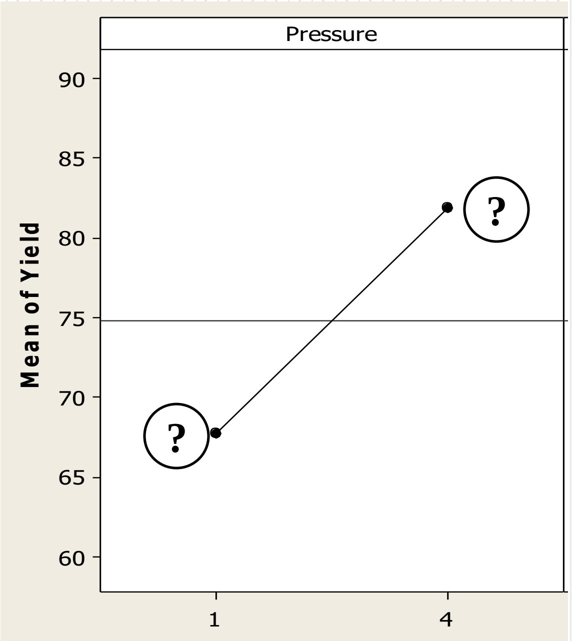
$$\bar{X}_{\text{Catalyst A}} = \frac{65 + 85 + 98 + 107}{4} = 88.75$$

$$\bar{X}_{\text{Catalyst B}} = \frac{55 + 63 + 68 + 52}{4} = 59.5$$

Does changing catalyst seem to change the average yield?

Evaluating Main Effects

Pressure	Catalyst	Yield 1	Yield 2
1 ATM	A	65	85
1 ATM	B	55	63
4 ATM	A	98	107
4 ATM	B	68	52



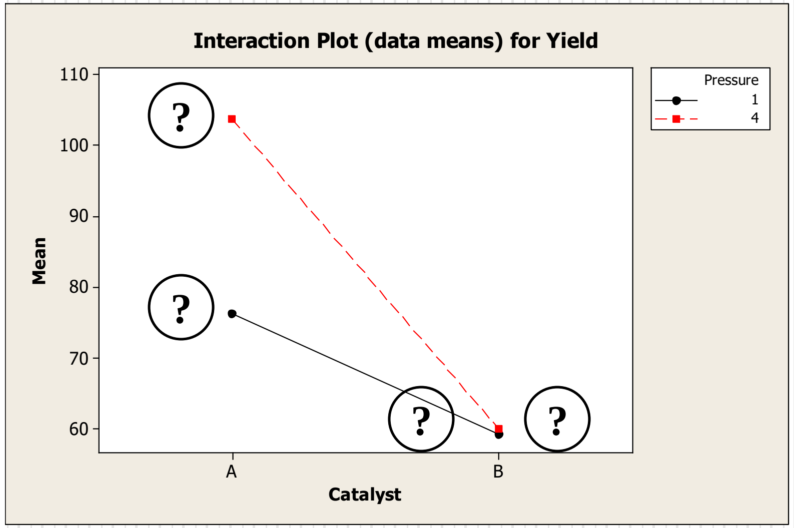
$$\bar{X}_{1ATM} = \frac{X_1 + X_2 + X_3 + X_4}{4} = \_\_\_\_\_$$

$$\bar{X}_{4ATM} = \frac{X_5 + X_6 + X_7 + X_8}{4} = \_\_\_\_\_$$

Does changing pressure seem to change the average yield ?

Evaluating Interactions

Pressure	Catalyst	Yield 1	Yield 2
1 ATM	A	65	85
1 ATM	B	55	63
4 ATM	A	98	107
4 ATM	B	68	52



$$\bar{X}_{1ATM - Cat\ A} = \frac{X_1 + X_2}{2} = \_\_\_\_\_$$

$$\bar{X}_{4ATM - Cat\ A} = \frac{X_5 + X_6}{2} = \_\_\_\_\_$$

$$\bar{X}_{1ATM - Cat\ B} = \frac{X_3 + X_4}{2} = \_\_\_\_\_$$

$$\bar{X}_{4ATM - Cat\ B} = \frac{X_7 + X_8}{2} = \_\_\_\_\_$$

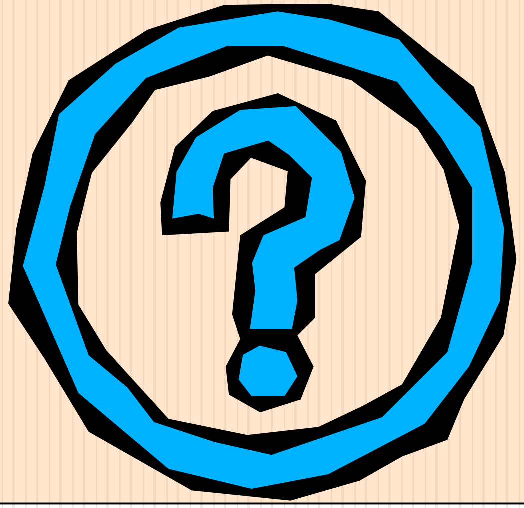
Does the pressure – catalyst interaction appear to be significant ?

Asking Questions of Reality

When we conduct a designed experiment, we are asking questions of reality.

Therefore, we must first decide…

What questions do we want the experiment to answer?



Before we go much further,
let us go into more detail about how to conduct a designed experiment

Steps for Designing and Conducting DOE

1. **Look at historical data and/or collect data to determine current process capability.**
2. **Determine the objective of the experiment.**
3. **Determine what to measure as the output of the experiment.**
4. **Identify factors (control factors and noise factors) that could affect the output.**
5. **Determine the number of levels for each factor, and their actual values.**

Steps for Designing and Conducting DOE

6. **Select an experimental layout that will accommodate the selected factors and levels and decide number of repetitions or replications.**
7. **Verify all measurement systems (Remember: garbage in, garbage out!).**
8. **Plan and prepare the resources (people, materials, etc.) for conducting the experiment. Should we randomize the runs? Create a test plan.**
9. **Conduct the experiment. Where appropriate, assure that each unit is labeled according to the experimental condition by which it was produced.**
10. **Measure the experimental units.**
11. **Analyze the data and identify strong factors for mean and for variation.**

Steps for Designing and Conducting DOE

12. **Determine <u>combination of factor levels</u> that best achieves the objective(s).**
13. **Run a <u>confirmation experiment</u> at this "optimal" combination.**
14. **Assure these best levels for strong factors are <u>maintained over time</u> by implementing Standard Operating Procedures and visual controls.**
15. **Re-evaluate <u>process capability</u> (to quantify improvement(s)).**

Planning the DOE

Use the DOE Planning Worksheet to begin planning your designed experiment.

1. Collect data and determine process capability
2. Establish the objective of the experiment.
3. Determine what to measure (Y) as the output of the experiment.
4. Identify Factors (Xs) that could affect the output (Y).
   a. Identify the controllable factors that could affect Y.
   b. Identify the noise factors that could affect Y.
5. Determine the number of levels for each factor, and their actual values

Note

Any variables not deliberately changed as part of the experiment must be held constant during experimentation.

Questions for a Prototype Experiment

Objective: To *minimize* prototype *cycle time*.

Factors	Output
Time order received	Prototype Cycle Time
P.O. approach	
Who entered data	

- First, clearly state your objective, then ask the team:
  - Which factors affect the average cycle time?
  - Which factors affect the variation in cycle time?

Prototype Example for DOE Planning

DOE Planning Worksheet

Black Belt: Susie Queue

Dept./Process: 100PI.2

Project Title: Develop Prototypes: Minimize prototype time

Date: Octember eleventeenth, 06

DOE Objective: Minimize prototype cycle time

Relevant Background: Current average prototype cycle time exceeds 9 days (13,100 minutes).

Output Characteristics

CT Characteristic	What/How to measure	Specifications	Use as measure for experiment? (yes/no)
Prototype cycle time	From confirmation of receipt of Customer fax until confirmation of receipt of error-free supplier P.O. fax	Less than 5400 minutes	Yes
Accuracy of purchase orders	Number of errors Number of rework iterations	No errors No rework	No

Is there a single output measure that captures all (or several) of the CT characteristics? Yes

Explain: Yes. By using a measure of cycle time that includes time spent on rework, reducign cycle time also captures the objective of reducing errors and rework

Prototype Example for DOE Planning

Controllable Factors	Strength of impact on Y	Ease to change during experiment	Include as factor in experiment? (Y/N)	If Factor in DOE: Current Level	If Factor in DOE: Proposed -	If Factor in DOE: Proposed +	If Not factor in DOE: How to hold constant? At What level?
Purchase Order Approach	◎	●	Y	Manual	Manual	Automated	N/A
Who Enters Data	◎	◎	Y	Both the Launch Coordinator & Data Entry Clerk enter data	Launch Coordinator enters data	Data Entry Clerk enters data	
Method for creating a "Pick List"	●	▲	N	Manual			Use current method (Manual)

Legend:
- ◎ **:Strong impact, Easy to change (Win!)**
- ● **:Moderate impact, Moderately easy to change (Place)**
- ▲ **:Weak impact, Difficult to Change (Show)**

Prototype Example for DOE Planning

Noise Factors	Strength of impact on Y	Ease to change during experiment	Include as factor in experiment? (Y/N)	If factor in DOE: Robustness/Variation reduction	If strong, will counteract			Levels to set during experiment	If Not factor in DOE: How to hold constant? At What level?
Time of day order received from Customer			Y	☐	☑			Morning (8 - 11 AM) vs. Afternoon (1 - 4 PM)	
				☐	☐				
				☐	☐				
				☐	☐				
				☐	☐				
				☐	☐				

Legend:

:Strong impact, Easy to change (Win!)

:Moderate impact, Moderately easy to change (Place)

:Weak impact, Difficult to Change (Show)

Prototype Example for DOE Planning

Experimental Layout

1 How many factors do we wish to include in the experiment? 0

2 At how many levels do we wish to test each factor? 0

3 What experimental layout should we use given our number of factors, number of levels, experimental objective and resource constraints?

Measurement System

1 What measurement system will be used?

2 Should we conduct a Gage R & R?

Data Collection

1 How many data Values should we collect for each test combination? 0

Repetitions or replications? Repetitions

2 Should we randomize the order of the runs? No

3 How will we identify each experimental unit (or data value) in terms of the test combination under which it was produced?

4 Should we create a data collection sheet? Yes

The Objective

2. Establish the <u>objective</u> of the experiment.

-Specifically, what needs to be improved?

- What are all the “Critical to” (CTx) characteristics?

The Output Measure

3. Determine what to measure (Y) as the output of the experiment.

- Is there a single output measure that captures all the CT characteristics? Should we measure more than one output?

Examples

1. In transactional processes, when minimizing "time" is the objective, make sure the measure of time includes time spent on rework. Thus, minimizing "time" also captures the objective of minimizing errors and rework.
2. In manufacturing processes, reducing variation in a measure of function reduces the chances for problems to occur.

Use Variable Data

Attribute data such as counting defects is an ineffective output for a DOE because it….

- Provides insufficient detail for evaluating factor effects.
- Often leads to conclusions that are difficult to reproduce.
- Requires large sample sizes.

Variable Data overcomes these problems.

Selecting a "Good" Output

To obtain the most useful information from a DOE conducted to improve a process….

1. Use Variable Data!
2. Measure something that relates to the basic function of the process.

Discussion: Classify Outputs

CTx

Process	Metric	Unit of Measure	Output Type
Bond paneling to the basement wall	Adhesive strength	Newtons (Lbs-Force)	Maximize
Process purchase orders	Accuracy	# of errors/P.O.	Minimize
	Cycle Time	Hours	Minimize
Mold parts for cheap plastic toys	Diameter, length, etc.	mm	Optimize
Other examples?			

Note: While the spoken "goal" for every metric is clear (maximize, minimize, optimize), what is the unspoken desire for every metric, every time?

MINIMIZE VARIATION!

As a BB, we must never forget the unspoken goal!

Measure Outputs that Relate to Function

By using a measure of *function*, rather than a problem-related measure, we can….

Minimize the chance for all problems to occur!!

Consider the molding example. The function of the process is to create parts of a specific dimension. By achieving minimum variation in part dimension, we have discovered a way to uniformly distribute particles and stresses throughout the part. When this happens, we minimize the occurrence of voids, flash and other problems.

NOTE

We do not have the same power to reduce problems when the problems themselves are measured as the experimental output!!

Measures of Function

- In addition, in order to see the factor effects, we need data that varies from run to run.
- Often, when defect-based measures are used, many test combinations produce zero defects (at least with small sample sizes). In such a case, the experiment is not doing its job.

Data (Results Not Useful):

	Data		
Run 1	0	0	0
Run 2	1	0	0
Run 3	0	0	0
Run 4	1	0	0

Results Not Useful

Data (Useful Results):

	Data		
Run 1	2.5	2.8	2.6
Run 2	3.1	2.9	3.4
Run 3	2.0	2.6	2.1
Run 4	2.9	3.8	3.4

Useful Results

Sample Size

- Sample size (the number of data values at each test combination) impacts experimental error, power, and reliability of results.
- Generally, the more data (the more degrees of freedom), the better the estimate.
- However, practical considerations (time, cost, etc.) must be weighed against statistical considerations.
- While exceptions may exist, a good rule of thumb is to collect a minimum of 3 data values for each test combination (unlike the preceding example!).

Note: Degrees of Freedom and Experimental Error are important statistical concepts that exceed the scope of our practitioner-based training.

Tendency for Defect to Occur (Leading Indicators)

- Sometimes metrics of function simply are not available. But with careful thought, we may identify something to measure on the product that shows a tendency toward defect before the actual defect occurs.

Manufacturing Example:

In assembling baseballs, even though no tears occur on the seam, stress at the seam (elongation of the thread holes) can sometimes be seen.

By using both the elongation of the thread holes (leading indicator) and tear length (variable data), we increase our power to discriminate among test units, thereby increasing the power of the experiment.

Tendency for Defect to Occur (Leading Indicators)

Transactional Example:

The Organizations's key metric for safety is *number of OSHA Recordable* injuries each month. We want to minimize this metric. Problems: Attribute data, low rate of occurrence, reactive metric.

By measuring *number of sprains/strains* reported to the dispensary each month, we introduce much greater sensitivity, significantly increase sample size, and have a leading indicator *prior* to OSHA Recordables. Switch to *time between reported injuries*, and we now have variable data.

Lesson Learned: Data is everywhere.
Be thoughtful in what you collect and how you collect it and *maximize* your results!

Practice: Identify the “Most Useful” Output Metric

Process:
Molding
Storage
Containers

→

Outputs:

- Dimensions (width, depth, length)
- Warpage
- Number of Cracks

1. What is the *function* of the molding process?
2. Which of these outputs most closely relates to the function of the process?
3. What type of output is it?

Controllable Factors

4a. Identify the controllable factors that could affect Y.

-Which of these should we change as factors during the experiment?

-Which should we hold constant?

Include the factor in the experiment when…..

1. The team believes it could have a strong impact on Y, and…
2. It is reasonably easy to change during the experiment.

All other "factors" must be held constant.

About Controllable Factors

Definition:

- Controllable (Control) factors are factors for which we can set and maintain a desired level during a process' normal operation.

Molding Example:

- On an injection molding machine, if we find that setting the temperature at 800^o results in less part variation than setting it at 900^o, then we can decide to run the machine at 800^o.

Purchase Order Example:

- If we test two approaches for creating purchase orders and find one approach works better than the other (shorter time, fewer errors), then we can decide to use the better approach.

Note

The more control factors that affect Y, the greater the potential for improvement.

Noise Factors

4b. Identify the noise factors that could affect Y.

- Which of these should we change as factors during the experiment?
- Which should we hold constant?

Include the factor in the experiment when…..

1. The team believes it could have a strong impact on Y, and…
2. It is reasonably easy to change during the experiment. All other "factors" must be held constant.

Note

In addition, the team should determine the strategy regarding each noise factor included in the experiment.

Strategies Regarding Noise Factors

Definition: Noise factors are factors that vary during normal operation, and we cannot control them. Or, we prefer not to control them because doing so would be very expensive.

Strategies for including a noise factor in an experiment:

1. To quantify its impact on Y. If it has a strong effect, then the team would identify a counteraction (either *directly* control the noise or reduce the *effect* of the noise).

2. To deliberately create variability in Y for the purpose of determining levels of controllable factors that will reduce the variation. When enough strong control factors are used, we may even be able to achieve robustness (insensitivity to noise).

Question

Which of these two strategies leads to lower cost improvements?

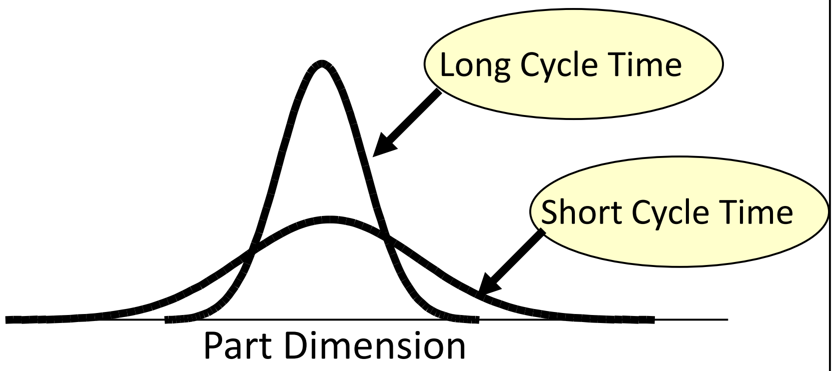
Setting appropriate levels

5. Determine the number of levels for each factor, and their actual values.

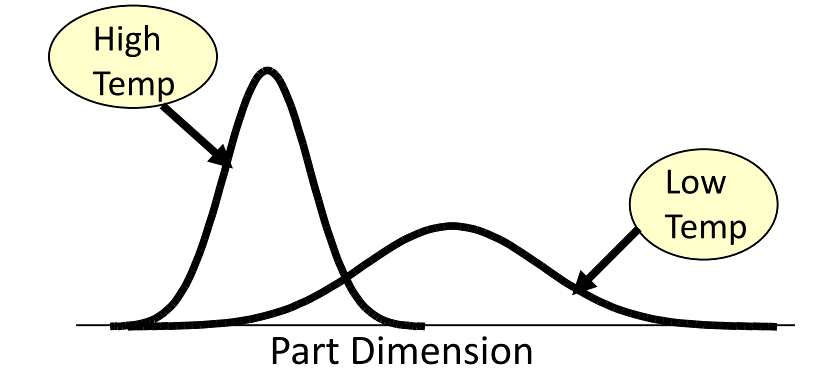
- Must be able to see the variation
- Realistic extremes so that you can learn something from the experiment

Factors Can Affect….

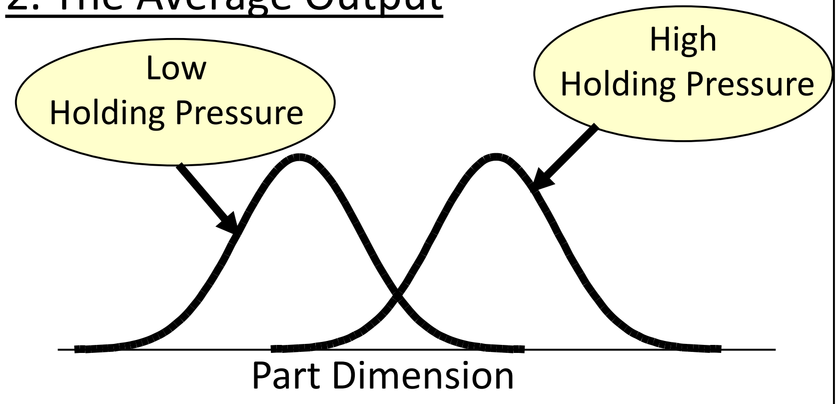
1. Variation in the output



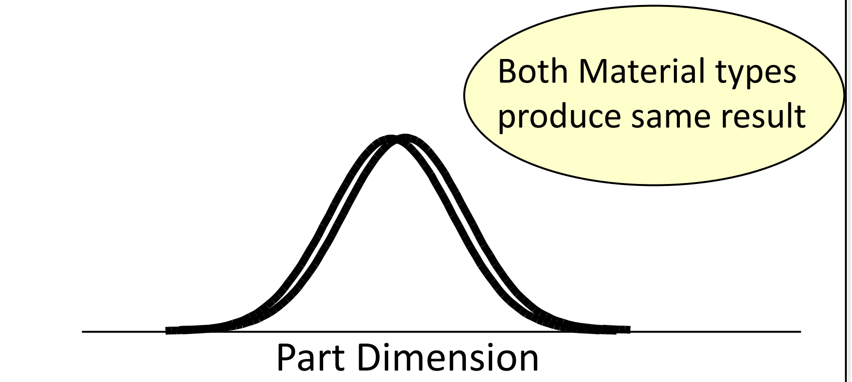
3. Both Variation and the Average



2. The Average Output



4. Neither the Variation nor the Average



For a given factor, which level would you pick for each example?

Let's take a stop here…

And see how we can apply what we have learned so far in the bounds of the Training Project Scenario

New Training Project Scenario

Our external customer has asked if we launch Satellites into space at a particular orbit such as 1,875 miles, 3,750 miles, 5,625 miles, etc… Since this relates to your current project you are brought in as part of a super team of 5 black belts to solve this issue. Since this is a competitive bid other Satellite Deployment Companies that are developing similar strategies. Your task is to develop a model that will simulate a robust Satellite Launching process that has the ability to Launch Satellites that range from 1,875 miles to 22,500 miles above the Earth. Understand that labor costs and quality costs are important concerns to our customers. An automatic award of two million dollars and sole source two year contract will be awarded by the contractor for the winning installation.

Step 1

- Set up a statapult configuration to minimize variation (10 minutes)
- New target XX"+/- 2.25" (every inch represents 187.5 miles). Make whatever adjustments you need to make in order to hit the target. Read....start the clock!!
- When your team can hit the target twice in succession, stop the clock.
- Collect (at least) 30 data points (5 shots/person minimum) and calculate process capability on current state process.
  - Remember to check for normality and stability (SPC) before calculating capability
  - While we are not performing a formal Gage R& R, now is the time to ensure minimal measurement system error (MSA)
- Instructor will review your Baseline process performance in 45 minutes

To do, Steps 1 - 5

- Using statapults, build a designed experiment using the DOE project worksheet.
- Answer the following questions:
  - What Input factors does your team plan to analyze to better understand their relationship with your metrics?
  - What is your data collection plan to analyze these factors?)
- Instructor will review/approve your Project Plan in 30 minutes.