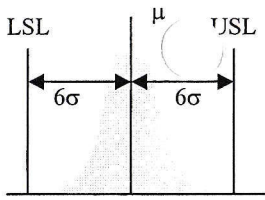


The Tech Group

Full Factorial Design Of Experiments

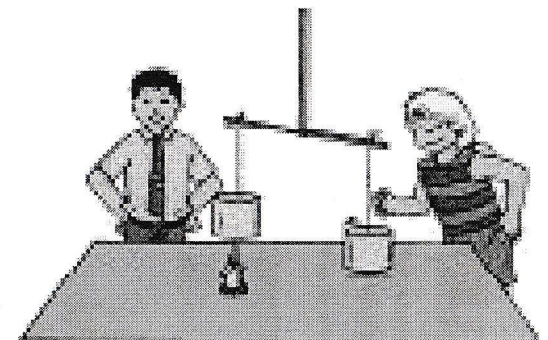
**Lean Sigma
Black Belt Training**

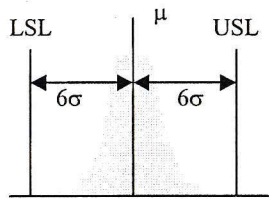


What is a Designed Experiment (DOE)?

A DOE is a carefully planned method of introducing variation into a process such that factors producing significant effects will be identified and quantified in an efficient manner.

In other words we will be able to effectively identify what factors make a difference and quantify the $Y=F(x)$ relationships within a process.

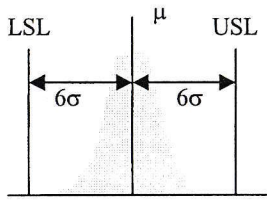




Designing an Experiment

All DOE's have the following in common:

- Factors will be selected that we want to study based on the knowledge gained about the $Y=f(x)$ relationships.
- Factor settings (a high and a low) will be selected for each factor.
- Response variable(s) will be chosen (what we measure).
- Runs (or treatments) will be created with specific factor settings defined (per DOE rules).

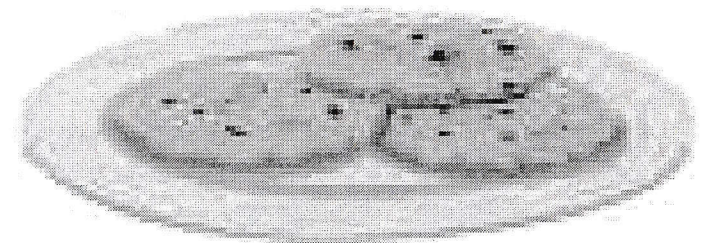


Example

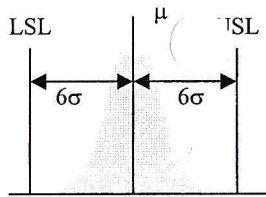
Full Factorial - Cookie DOE

AYTOCE Cookie Company produces packaged cookie dough. Their marketing department is interested in developing cookie instructions that will yield the best taste, texture and appearance of the cookie. They are interested in understanding the effects of two factors, **bake time** and **bake temperature**:

1. Bake Time (A).
2. Bake Temperature (B).



- What other factors are there?
- How would those factors be discovered?



Full Factorial - Cookie DOE (con't)

OBJECTIVE:

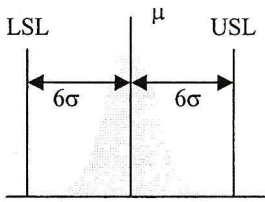
Improve taste, texture and appearance of cookies and to determine the range for the instructions.

TWO VARIABLES:

Bake Time (A).

Bake Temperature (B).

1. What questions should be asked?
2. What should the response be?
3. Construct a 2^2 factorial design.
4. How should the experiment be run?



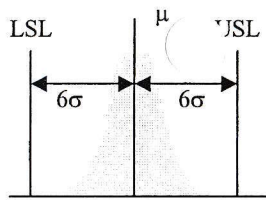
Full Factorial - Cookie DOE (con't)

DOE design and result:

A	B	Y
Bake Time (Min)	Bake Temp	Taste
6	375	41
10	375	50
6	450	47
10	450	35

Which factors are important?

How should the important factors be set?



Six Sigma Rules of Analysis:

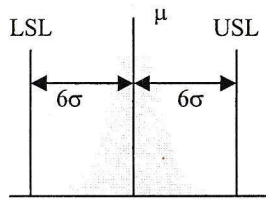
➤ **Practical**

➤ **Graphical**

➤ **Analytical**

➤ **Practical:**

- Once the experiment has been run, there are a host of methods to assist in analysis. The first consideration, before applying any statistical analysis techniques, is **whether the results are of any practical importance.**
 - ✓ **Did the response variable change?**
 - ✓ **Did it change the desired amount?**
- If the response variable did not change enough, it may be that:
 1. Factor levels were not set far enough apart.
 2. The selected factors do not affect the response variable.
 3. The measurement system is not adequate

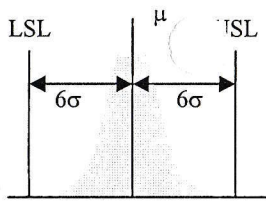


Practical View

2² Full Factorial Experiment

A	B	Y
Bake Time (Min)	Bake Temp	Taste
6	375	41
10	375	50
6	450	47
10	450	35

1. Practically speaking, what observations can be made about the data if:
 - a change of 10 is required?
 - a change of 50 is required?
2. What preliminary actions would you recommend right now?



Analytical View

To begin analyzing the importance of each factor, code the levels as (-) or (+) as shown to the right.

A	B	Y
Bake Time (Min)	Bake Temp	Taste
6 (-)	375 (-)	41
10 (+)	375 (-)	50
6 (-)	450 (+)	47
10 (+)	450 (+)	35

Create a coded column for the AB interaction. It is the product of the A and B signs.

Example: (-) x (-) = (+)

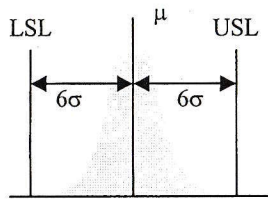
We have added the interaction column to the table on the right.

A	B	AB	Y
Bake Time (Min)	Bake Temp	Interaction	Taste
6 (-)	375 (-)	(+)	41
10 (+)	375 (-)	(-)	50
6 (-)	450 (+)	(-)	47
10 (+)	450 (+)	(+)	35

We will sort the response (Y) in ascending order and look for patterns of “level” separation. This is called an ANOG.

It appears our interaction is important

A	B	AB	Y
Bake Time (Min)	Bake Temp	Interaction	Taste
10 (+)	450 (+)	(+)	35
6 (-)	375 (-)	(+)	41
6 (-)	450 (+)	(-)	47
10 (+)	375 (-)	(-)	50



Analyzing the Main Effects

The importance of a factor is called the “effect”. It is determined by subtracting the averages of the high & low levels for a factor.

$$\text{Effect} = (\text{Avg. at “+” level}) - (\text{Avg. at “-” level})$$

Effect of Factor A :

We want $\bar{A}(+) - \bar{A}(-)$:

$$\bar{A}(+) = (50 + 35)/2 = 42.5$$

$$\bar{A}(-) = (41 + 47)/2 = 44.0$$

Therefore, the main effect of A = $42.5 - 44 = -1.5$

A	B	AB	Y
Bake Time (Min)	Bake Temp	Interaction	Taste
6 (-)	375 (-)	(+)	41
10 (+)	375 (-)	(-)	50
6 (-)	450 (+)	(-)	47
10 (+)	450 (+)	(+)	35

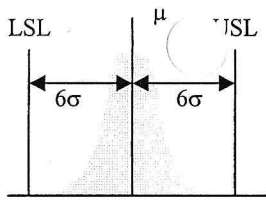
Effect of Factor B :

We want $\bar{B}(+) - \bar{B}(-)$:

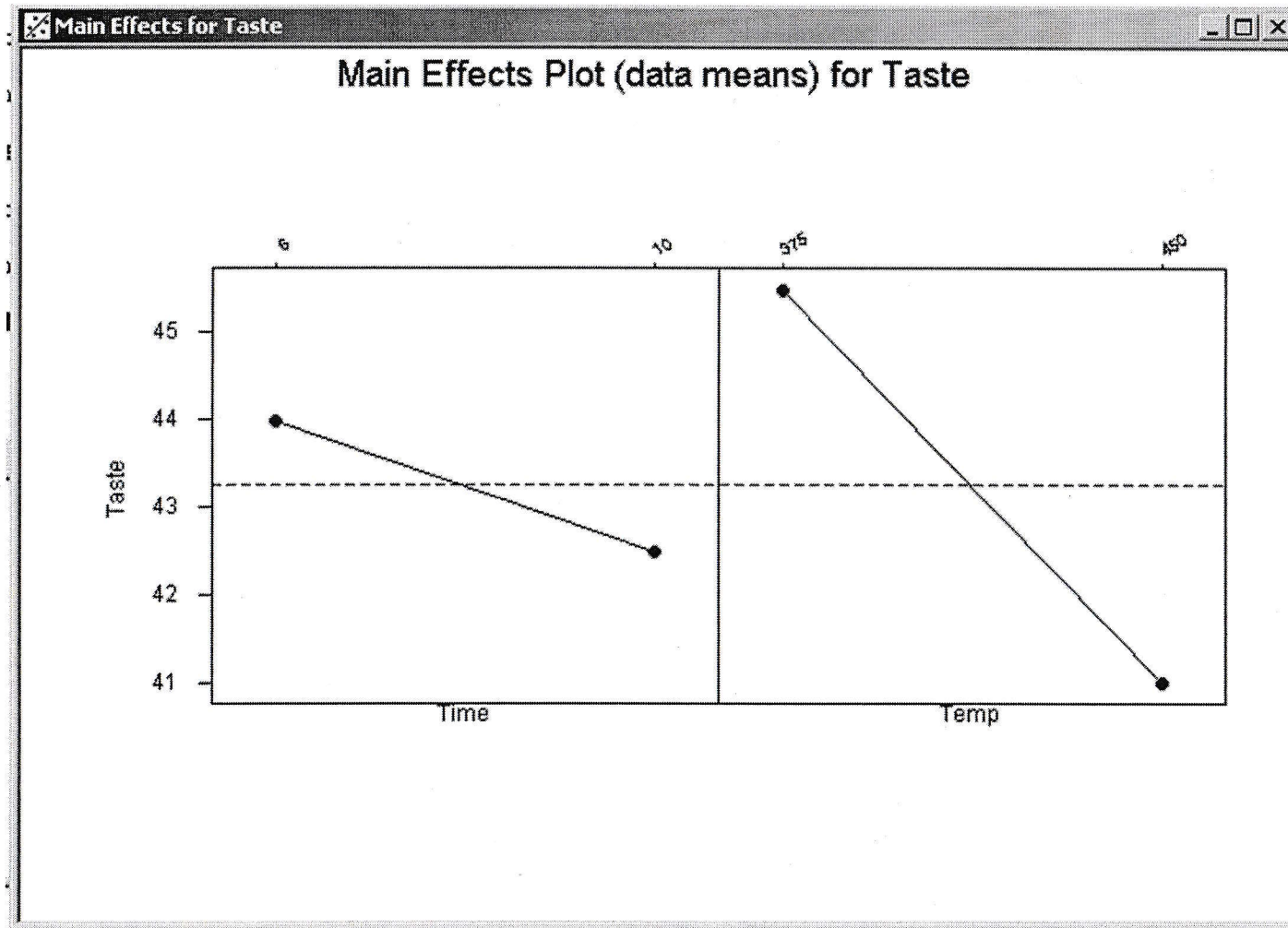
$$\bar{B}(+) = (47 + 35)/2 = 41$$

$$\bar{B}(-) = (41 + 50)/2 = 45.5$$

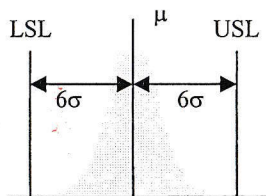
Therefore, the main effect of B = $41.0 - 45.5 = -4.5$



Graphical View



What does this tell us about the effects of time and temperature on taste?

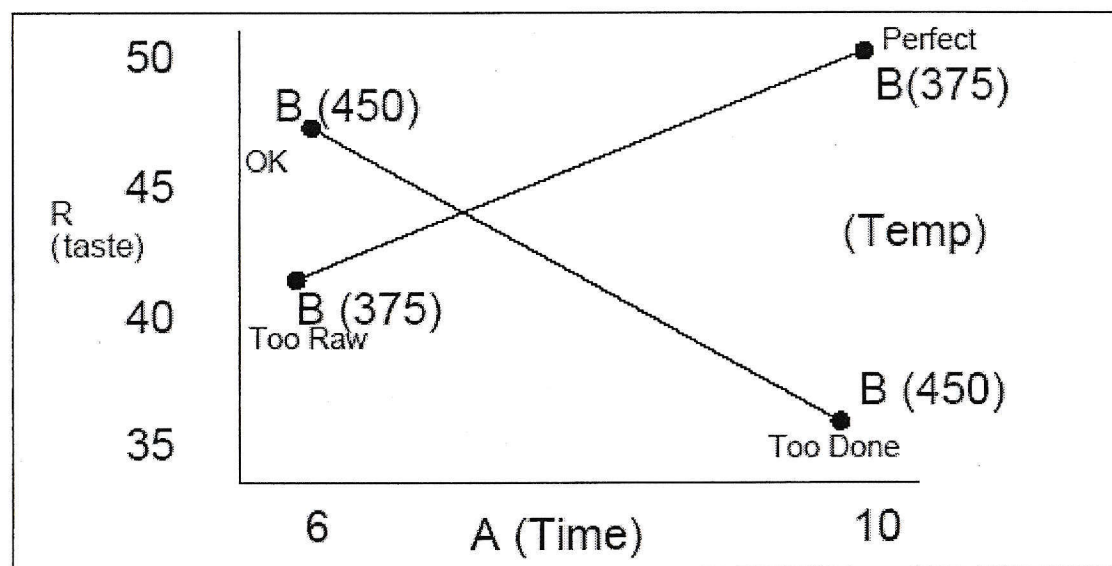


Analyzing the Interactions

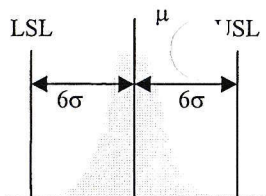
Interaction: 2 factors interact if one factor's effect on the response is dependent upon the level of the other factor.

The $\overline{AB}$ interaction = $(41 + 35)/2 - (50 + 47)/2 = -10.5$

Interaction Plot:



A	B	AB	Y
Bake Time (Min)	Bake Temp	Interaction	Taste
6 (-)	375 (-)	(+)	41
10 (+)	375 (-)	(-)	50
6 (-)	450 (+)	(-)	47
10 (+)	450 (+)	(+)	35



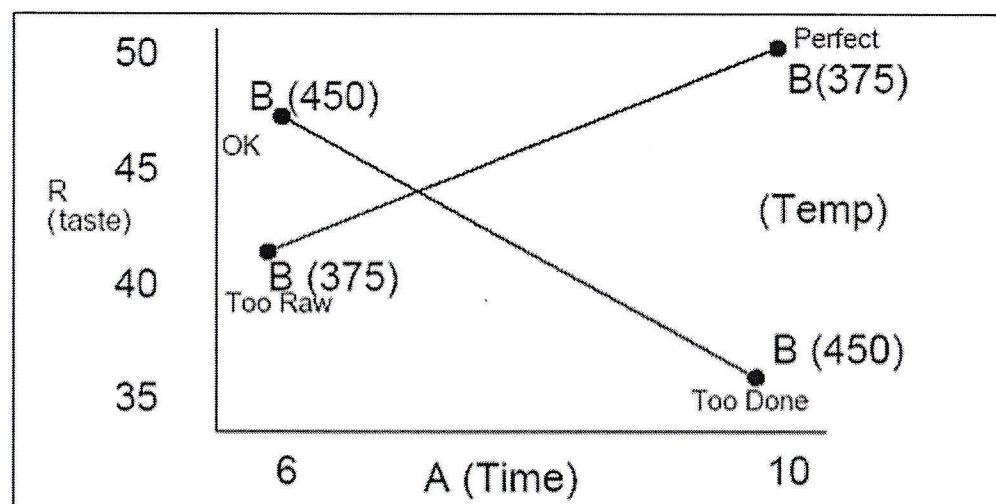
Analysis Summary

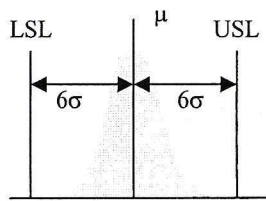
Now that we have calculated the effect for time, temperature, and the time/temperature interaction we can compare the relative importance of each:

<u>Factor</u>	<u>Effect</u>	In this experiment, the <u>Time x Temperature Interaction</u> has the largest effect and is, therefore, the most important.
Time	1.5	
Temperature	4.5	
Time x Temperature	10.5	This is consistent with our ANOG result!

Practical Significance:

Controlling Time or Temperature alone will not ensure good tasting cookies. Time and Temperature should be analyzed and set together to ensure the best tasting cookies.





ANalysis Of Good (ANOG)

In this example, we can see that the AB interaction is significant.

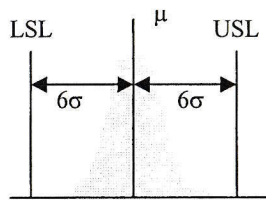
Run	Stop Pin	Launch Angle	Tension Pin	Interactions			Yavg
	A	B	C	AB	AC	ABC	
3	-1	+1	-1	-1	+1	+1	92
2	+1	-1	-1	-1	-1	+1	91
7	-1	+1	+1	-1	-1	-1	87
8	+1	+1	+1	+1	+1	+1	81
6	+1	-1	+1	-1	+1	-1	78
5	-1	-1	+1	+1	-1	+1	74
4	+1	+1	-1	+1	-1	-1	74
1	-1	-1	-1	+1	+1	-1	63

Sorted by Y_{avg}

(Example only)

Perform an ANOG on your data:

1. Color code all of one sign (yellow minus signs in this case)
2. Sort data by Y_{avg} (ascending or descending OK)
3. Look for patterns of separation in data. We want the signs to cluster together at opposite ends.



Coded vs. Uncoded Units

The predictive equation you created is based on the assumption that the high settings for each of your factors is “+1” and the low setting for each factor is “-1”. So the coefficients of the predictive equation are said to be in **“coded units”**.

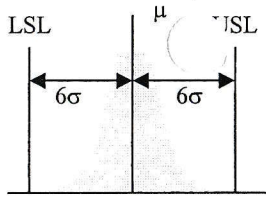
We must convert *coded units* to *uncoded units* if we wish to set our statapult based on the predictive equation.

For your convenience, a conversion chart has been provided on the right. Use this chart to complete the exercise on the next page.

Stop Pin (A)		
	Coded	Uncoded
Low	-1	3
mid	0	NA
High	1	4

Tension Pin (C)		
	Coded	Uncoded
Low	-1	2
mid	0	3
High	1	4

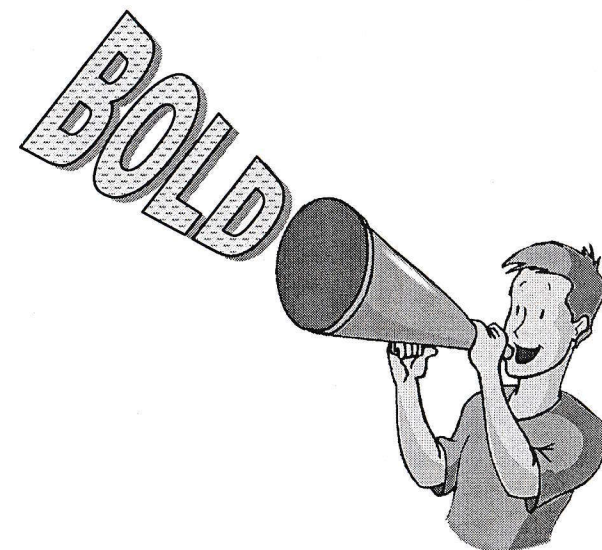
Launch Angle (B)		
	Coded	Uncoded
Low	-1	155
	-0.84	157
	-0.68	159
	-0.52	161
	-0.36	163
	-0.2	165
	-0.12	166
	0.04	168
	0.2	170
High	0.36	172
	0.52	174
	0.68	176
	0.84	178
	1	180

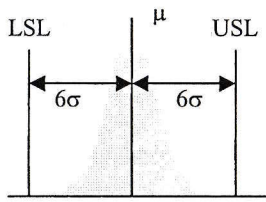


Inference Space

- The area within which you test and draw conclusions based on the results of your experiment.
- Level settings – start **BOLD** and reduce space between levels as you approach optimum conditions.

Narrow	Bold
1 lot of material	Several lots
1 Day	Several weeks
1 Machine	5 Machines
1 Operator	Many Operators
100 – 110 psi	100 – 2000 psi





Replication vs. Repetition

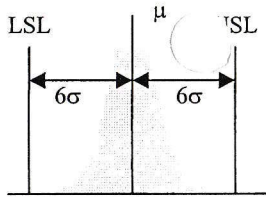
of Replications: Refers to the number of times a treatment combination is set up and run. The data collected from a single sample from each run is called a replication.

Example: A 2^2 , four run experiment is run and response variable measurement are taken on one part for each run. The values of the single set of response variable measurements is one replication. If the treatment combination is changed then set up again and measured (the experiment is done again), the total data collected (8 values, 4 for each experiment) would represent two replications.

of Repeats: Refers to the number of parts taken and measured from each treatment combination.

Example: A 2^2 , four run experiment is run once and response variable measurement are taken. Three parts are measured per run. The total data collected (12 values, 3 from each run) would represent three repetitions.

Repeats are simply additional samples measured per run. Replications require additional runs.

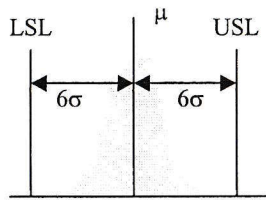


Replication vs. Repetition (con't)

Three Repeats

Worksheet 2 ***										
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	StdOrder	RunOrder	CenterPt	Blocks	A	B	Y1	Y2	Y3	Yavg
1	1	1	1	1	-1	-1	58	58	59	58.3
2	2	2	1	1	1	-1	54	55	53	54.0
3	3	3	1	1	-1	1	63	63	62	62.7
4	4	4	1	1	1	1	46	44	45	45.0
5	5	5	1	1	-1	-1	57	6	5	22.7
6	6	6	1	1	1	-1	52	52	52	52.0
7	7	7	1	1	-1	1	65	66	65	65.3
8	8	8	1	1	1	1	44	43	44	43.7
9	9	9	1	1	-1	-1	55	54	54	54.3
10	10	10	1	1	1	-1	51	52	51	51.3
11	11	11	1	1	-1	1	59	58	60	59.0
12	12	12	1	1	1	1	45	44	45	44.7
13										

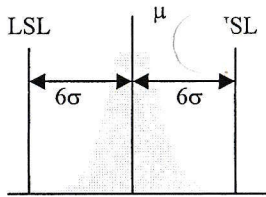
Three Replicates



Replication vs. Repetition (con't)

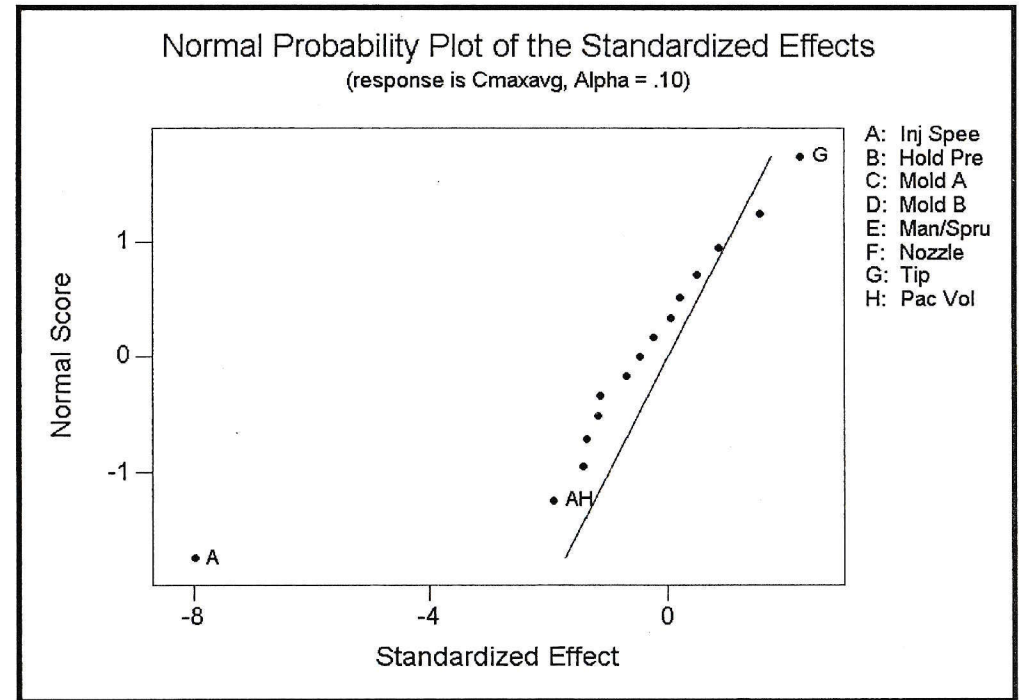
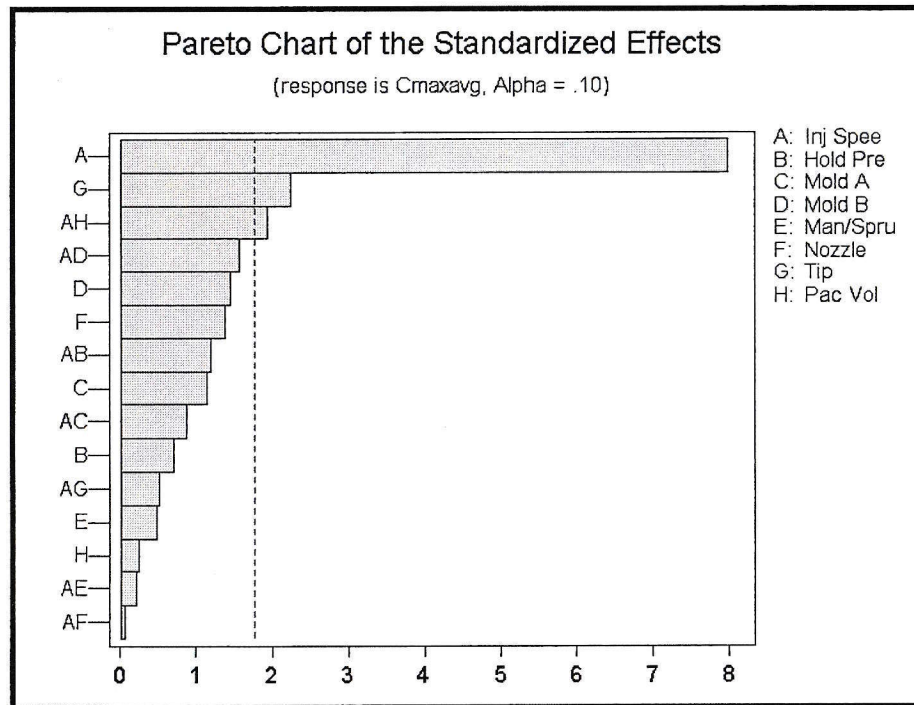
- Significance is determined by comparing variation between runs to variation within runs.
- Replications give us data so we can calculate variation within runs (Setup-to-setup how consistent is a specific set of settings (run) vs. other settings (runs)).
- Replications gives us an “error term” so we can calculate significance, repetitions do not.
- Without an error term we cannot estimate significance.

Error is our friend!

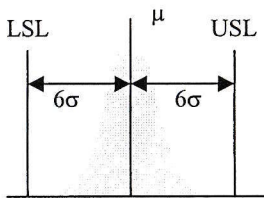


Interpretation of Significance

After we have established an error term we can determine significance by looking at the following graphs:



Which effects are significant based on the data above?



Session Window Results

The information below shows significance of effects. Also, coefficients shown can be used to create the $Y=f(x)$ predictive equation.

Fractional Factorial Fit: Cmaxavg versus Inj Speed, Hold Press, ...

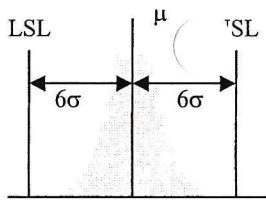
Estimated Effects and Coefficients for Cmaxavg (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		-25.17	11.4823	-2.19	0.047
Inj Spee	-0.95	-0.47	0.0592	-7.99	0.000
Hold Pre	-0.21	-0.11	0.1534	-0.69	0.500
Mold A	-0.13	-0.07	0.0586	-1.13	0.279
Mold B	-0.19	-0.09	0.0659	-1.44	0.175
Man/Spru	-0.06	-0.03	0.0638	-0.47	0.648
Nozzle	-0.30	-0.15	0.1075	-1.37	0.192
Tip	0.28	0.14	0.0617	2.24	0.043
Inj Spee*Hold Pre	-0.14	-0.07	0.0607	-1.18	0.258
Inj Spee*Mold A	0.11	0.05	0.0616	0.87	0.400
Inj Spee*Mold B	0.27	0.13	0.0859	1.56	0.142
Inj Spee*Man/Spru	0.03	0.01	0.0617	0.21	0.837
Inj Spee*Nozzle	0.01	0.00	0.0653	0.06	0.957
Inj Spee*Tip	0.07	0.03	0.0643	0.51	0.620

Significant Effects

Analysis of Variance for Cmaxavg (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	8	7.4720	7.7279	0.9660	9.10	0.000
2-Way Interactions	7	0.7876	0.9140	0.1306	1.23	0.354
Residual Error	18	1.3804	1.3804	0.1062		
Total	33	11.0974				

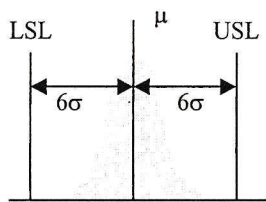


Randomization

- Randomization is the cornerstone underlying the use of statistical methods in DOE's.
- Statistical methods require that the variation (noise or error) be independently distributed random variables.
- Averages out the effects of extraneous factors that may be present.

For example: Slight differences in temperature that may occur throughout a DOE.

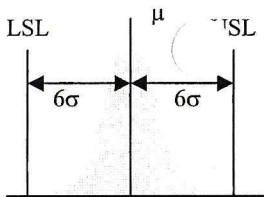
Randomize or risk incorrect results !



DOE Planning Template

Planning is the single most important part of a DOE. An important tool to help facilitate the planning and document the expectations is the DOE Planning Template (shown on left).

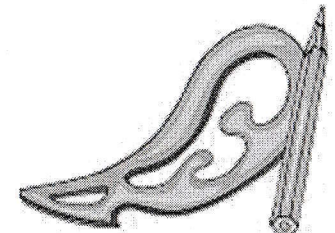
DOE Planning Template		
Date:	1/9/2003	
Name:	Tree-Tip Cassette SoD7452 Mold #2	
Title:	Mold Optimization for Sink	
Objective:	To determine what factors are significant producing sink and related response variables. It is hoped we will be able to demonstrate a process that will produce sink so eDart parameters can be evaluated for control purposes. Ultimately this information will also be used going forward to optimize the process when material change is put in place.	
Background:	Mold #2 is a 16 cavity hot tip mold that produces parts with sink. Although this appears to be an isolated event, this DOE is being conducted to find root cause and possible controls and ultimately ensure an optimum process is developed.	
Assumptions:	Lot-to-lot variation will not be significant. Temperature changes will stabilize in 30 minutes or less.	
Response Variables:	Measurement Techniques:	
OD, OOL, OOH	Perispectba plus	
Sink, Stragling	Perispectba plus	
eDart measurement		
Factor:	Level (-1)	Level (+1)
Injection speed (no profile)	3 in/in	7.5 in/in
Mold Temp - A (Front)	60	120
Mold Temp - B (Back)	60	100
Mold Pressure	500 psi	700 psi
Pack/hold pressure limit due to cycle stage 2	1 in/sec	2.5 in/sec
Barrel temps (dry profile w/nozzle)	350	460
Main hot & Sprue Temp	440	490
Tip settings (16)	440	490
Resin block with resin - day1 & day2	PP1105	PP1605
*(Highlighted rows are most time intensive with run change)		
Noise:	Control Method:	
Time between runs (while temperatures stabilize)	Specify and document	
Measurement error	MSE	
Machine variation	Existing controls (PM, eDart monitoring, training)	
Temp & humidity		
Replication:	Yes - 2 replicates	
Randomization:	Yes	
Design Matrix:	2 ⁵ (2 level, 5 factor, 1/16 fractional + 1 center = 17 runs x 2 (replications) = 34 runs)	
Analysis:	This is screening DOE that will lead to critical parameters	
Est Time/Cost:	Two days (16 hrs), 1 Process Tech, 1 QA, 1 Proj Eng	
	Resin = \$300; Proc Tech + press = \$1,920; QE+PE = \$300; Total = \$3,520	
Comments/Notes:		
Will validate water flow through the mold and document.		
Middle and Rear barrel zones will tack at 40 degrees below nozzle on high settings.		
Two center processes will be conducted, one with each resin.		
Will block with resin		

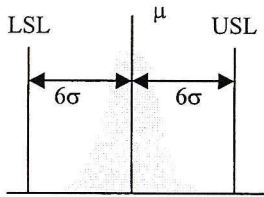


DOE Planning Template (con't)

 Tech Group DOE Planning Sheet							
The Tech Group WE PROVIDE INNOVATIVE SOLUTIONS							
Name:		Date:		Title:			
Phone No:		Location:		Project:			
Objective:							
Background Information:							
Assumptions:							

The first portion of the DOE Planning Template provides areas to collect general information.

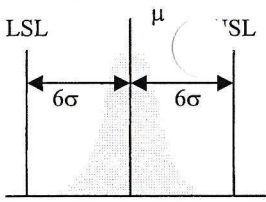




DOE Planning Template (con't)

Response Variables:		Measurement Techniques:	
Factors:		Level (-1)	Level (+1)
Fixed Process Parameters:			
Noise Variables:		Control Method:	
Design Matrix:			

The second portion of the DOE Planning Template provides areas to complete information about the process and the design matrix.



Class Problem #1

Label-Me, a U.S. manufacturer of labels, is in the final steps of developing a new label product.

Historically, Label-Me has experienced problems with the thickness of their label backings. Label-Me investigates the possible process configurations that may contribute to the thickness problem. Label-Me decides they need to select a new paper supplier. Also, since the labels can be printed at one of two plants, they need to select a plant.

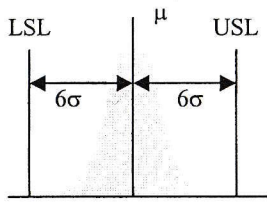
Two factors are considered worthy of investigation to eliminate the thickness problem. They are:

L= Paper Supplier (Supplier A and Supplier B)

P= Plant 1 and Plant 2 (P1 and P2)

They decide to run a factorial experiment with factors L and P (in first and second columns respectively). The experiment is run over the course of one week.

A 2 x 2 factorial design was run with three replicates for each experimental condition (treatment combination).

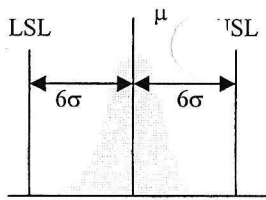


Class Problem #1 (con't)

Specification = $0.062 \pm .006$

Historical data on similar labels: CR=0.50
Cpk=0.66

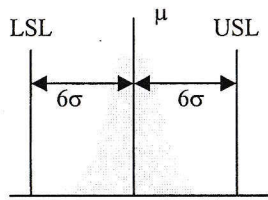
1. What, historically has been the problem (centering, variation)?
2. How does the statement of the problem affect the experiment strategy?
3. Knowing there are two variables of interest, Paper Supplier (L) & Plant (P), what questions should be asked?
4. How should the data be organized to meet the objectives?
5. What assumptions have been made in running the experiment?
6. How might these assumptions affect our conclusions?



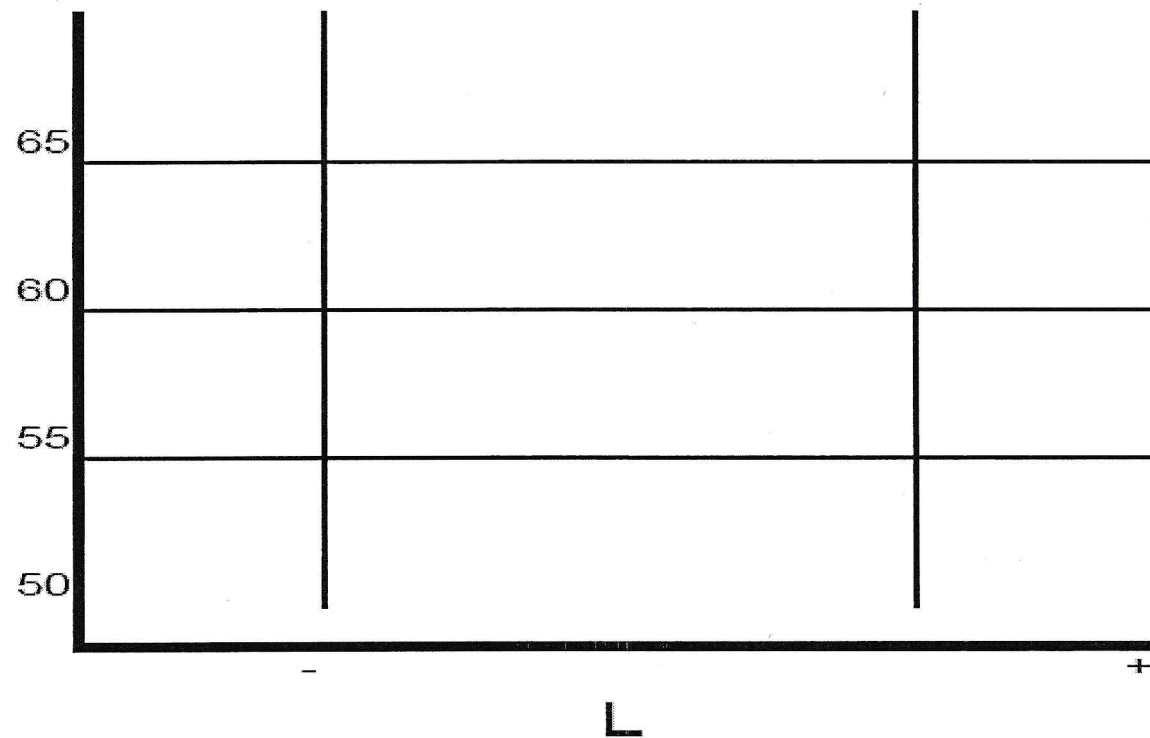
Class Problem #1 (con't)

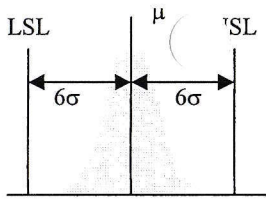
Response					
L	P	LP	(0.0xx)	X	R
-	-	+	62,65,59		
+	-	-	54,55,55		
-	+	-	61,66,58		
+	+	+	58,58,59		

7. Based on the experimental results shown above, draw and analyze an interaction plot (use the following page).
8. What else could be analyzed from the experiment results?
9. What are the effects of L, P and the LP interaction?
 L=
 P=
 LP=
10. What conclusions could be made?



Interaction Plot





Class Problem #2

A machine shop provides parts for the automotive industry. They want to improve the average and consistency of product yields (first pass) in the next generation of parts, and thereby, reduce production costs. Current process yields average 83%. After logical problem solving activities, 2 factors are deemed most likely to cause low yields. An experiment was set up for the factors:

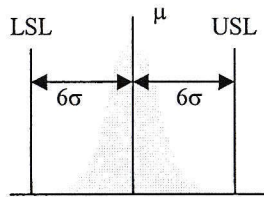
Cutting fluid (C): - Water Base
+ Oil Base

Steel (S): - 310 Stainless
+ 302 Stainless

A 2 x 2 factorial design was run with four replicates for each treatment combination.

Questions:

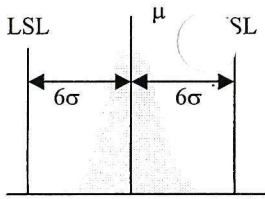
1. How could this experiment be replicated?
2. How could this experiment be repeated?
3. How should this experiment be run?



Class Problem #2 (con't)

C	S	CS	Response (%Yield)	X avg	R range
-	-	+	76,80,81,83		
+	-	-	80,84,88,89		
-	+	-	80,80,82,86		
+	+	+	86,90,92,92		

1. What comments can be made about the response variable?
2. What observations can be made about the data?
3. What factors may have influenced the results?
4. How could the data be analyzed?



Summary

- Designed experiment's are a powerful way to solve complex problems with many variables.
- What are factors and levels and how do you pick them?
- Go bold with factor levels. What could happen if you go too bold? Not bold enough?
- Why do we want an error term in our experiment? How do you get it?
- What is the difference between replication and repetition?
- What is the definition of an interaction?
- Why is it important to randomize?