The Tech Group

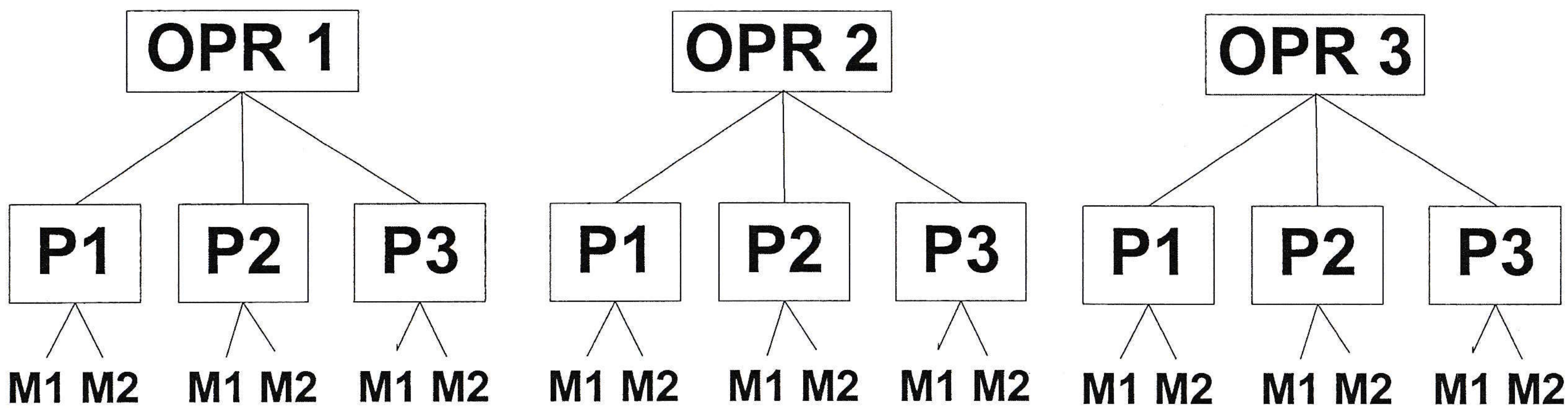
Components of Variation & ANOVA Applications

Lean Sigma

Black Belt Training

The Tech Group

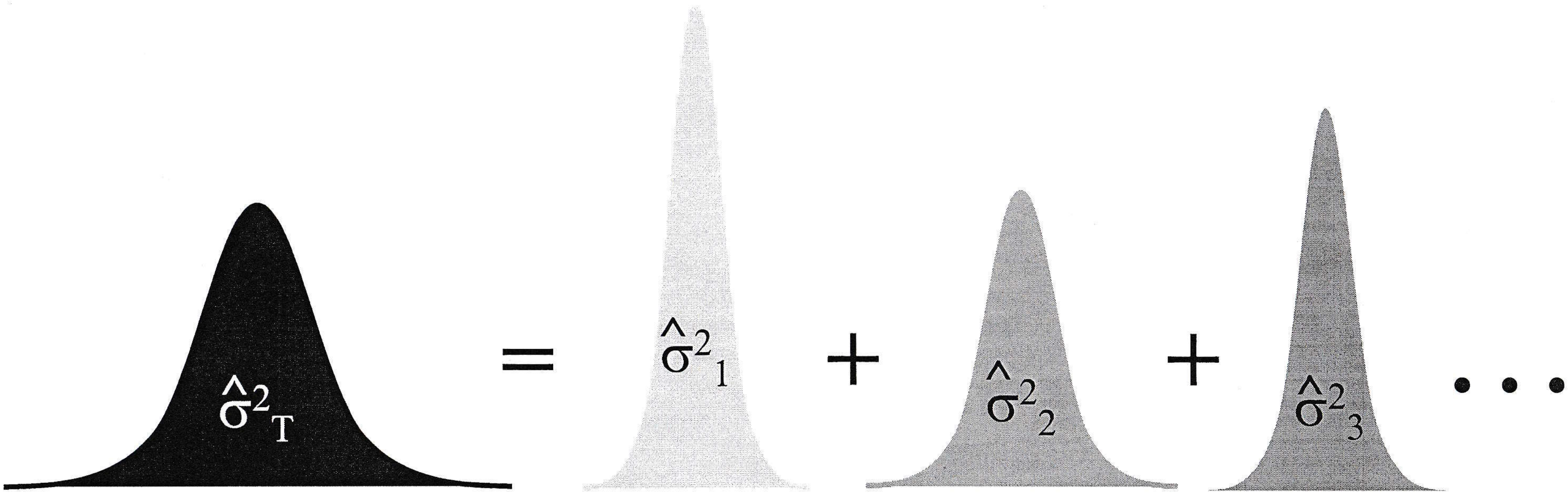
Recall MSE?



What are the components of variation?

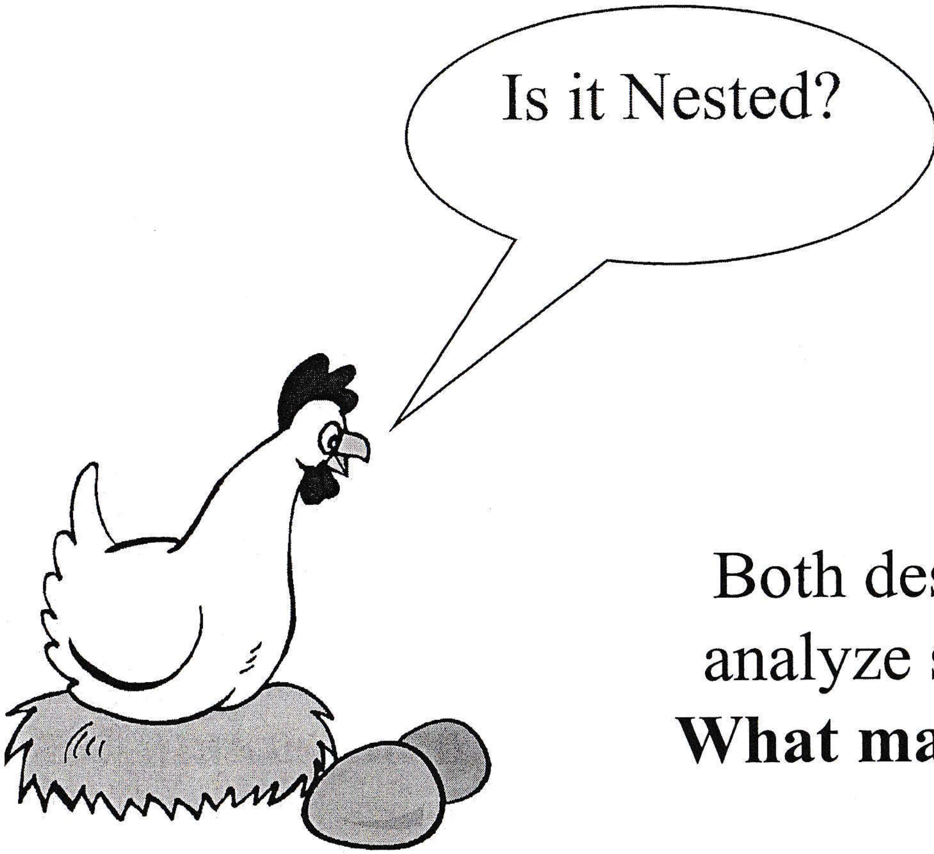
Components of Variation

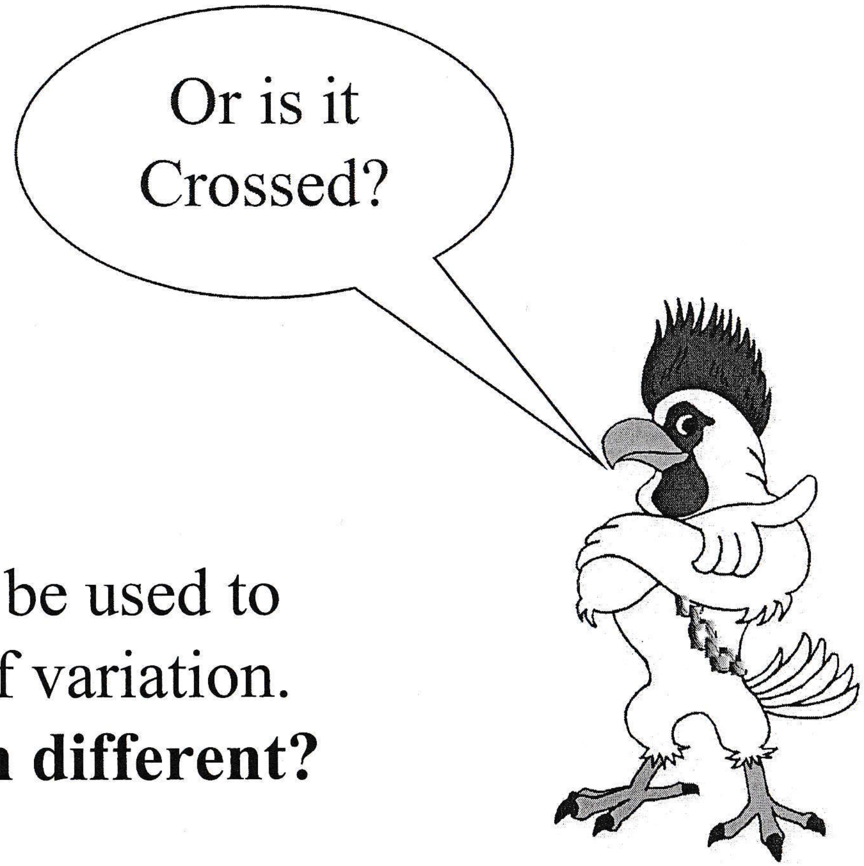
$$\hat{\sigma}^2_T = \hat{\sigma}^2_1 + \hat{\sigma}^2_2 + \hat{\sigma}^2_3 \ldots\ldots \hat{\sigma}^2_n$$



The Tech Group

Nested vs. Crossed Studies





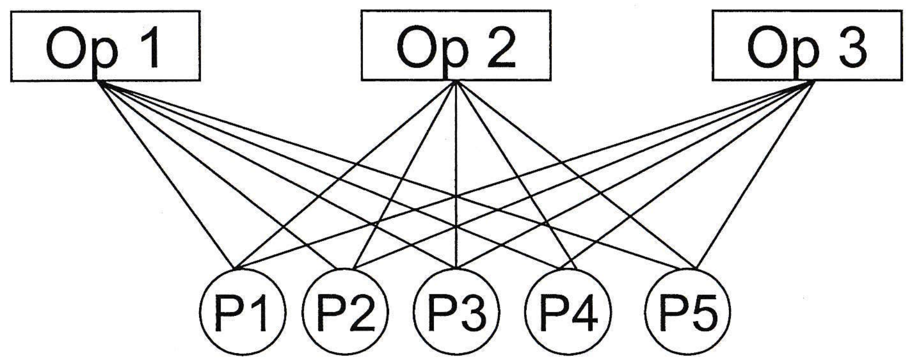
Both designs can be used to
analyze sources of variation.
What makes them different?

The Tech Group

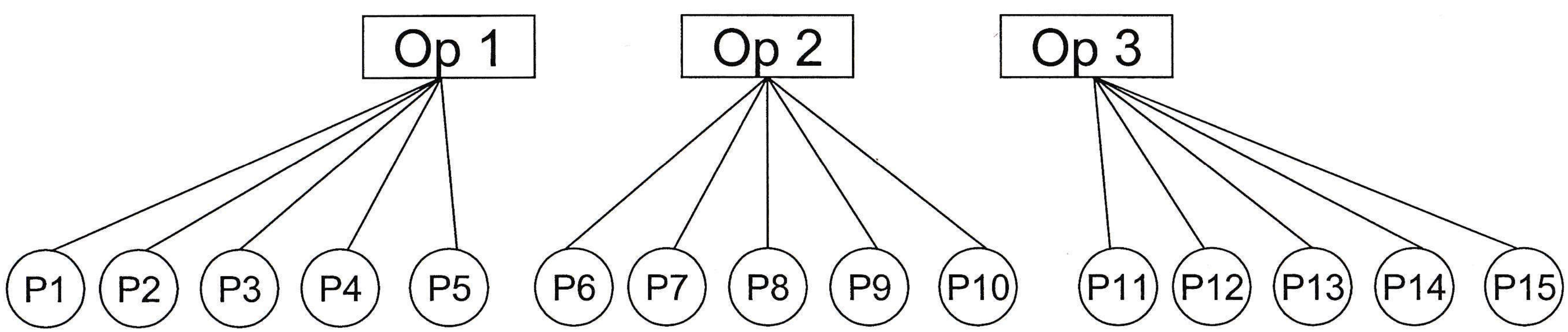
MSE Example

Nested vs. Crossed

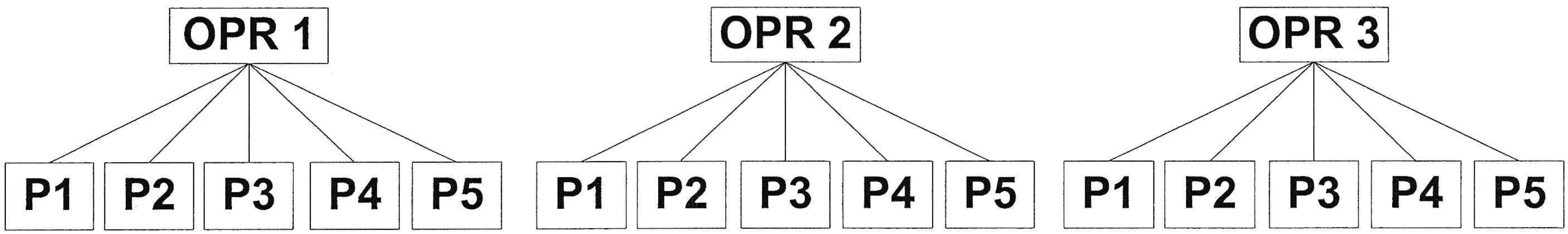
Crossed – Each operator measures the same parts



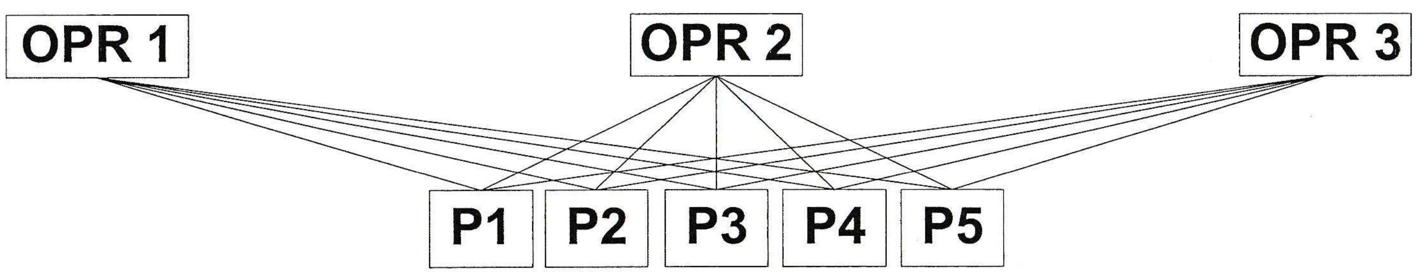
Nested – Each operator measures <u>different</u> parts



Crossed Design
(The MSE from week 1)



Alternate way to graphically illustrate a crossed design:



Most MSE's are "Crossed"

Control Chart Review

The Tech Group

- The Shewhart X-Bar and R control chart is an example of a "Nested" design… Why?

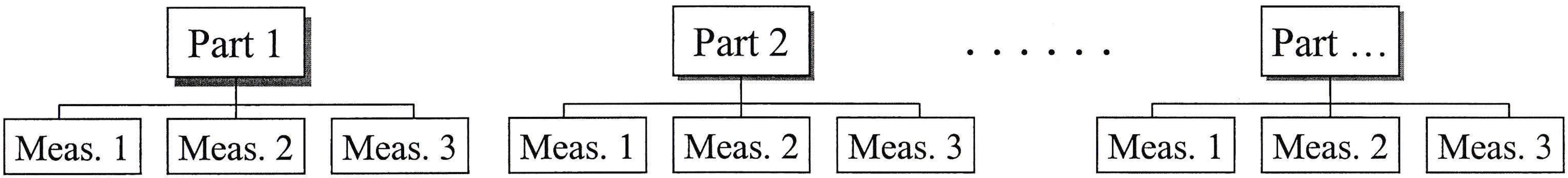
- Shewhart's *Rational Sub-grouping* was used to organize data into two groups;
  - *Within the subgroup*
  - *Between different subgroups*

- *Rational Sub-grouping* strategy leads to four questions:
  1. What changed within "sub-group" ?
  2. What changed between "sub-group" ?
  3. What did not change within "sub-group" ?
  4. What did not change between "sub-group" ?

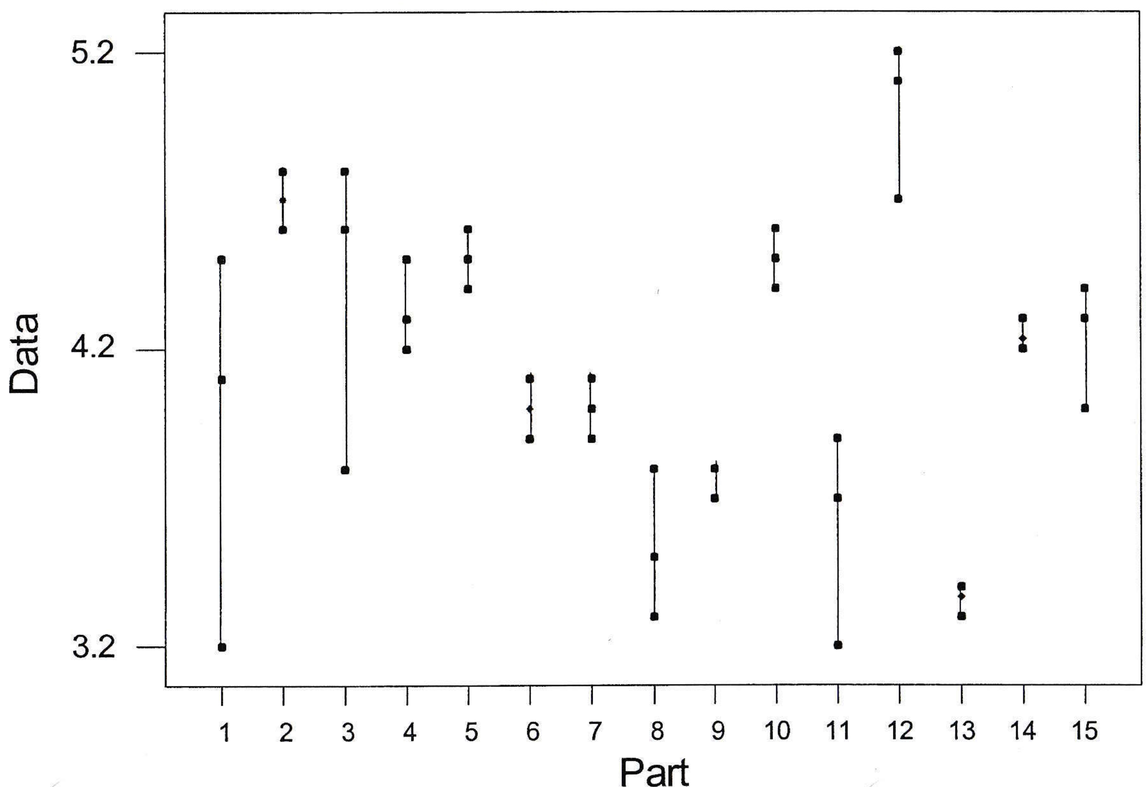
Sub-Grouping Example

The Tech Group

- **Xbar & R chart schematic:**



- **Dot frequency diagram for a control chart:**



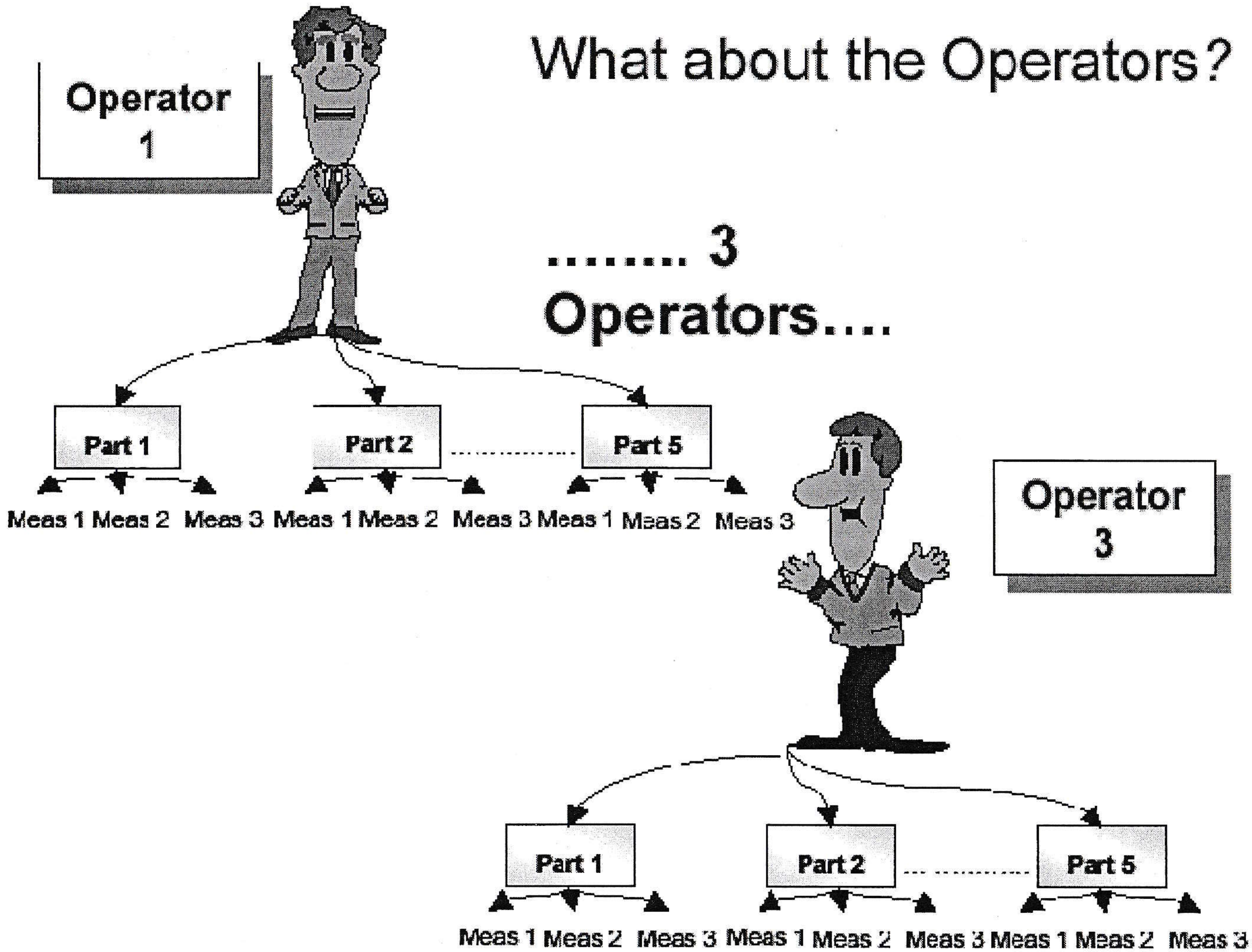
What is the biggest source of variation?

Extending the Sub-Grouping Concept

The Tech Group

What about the Operators?

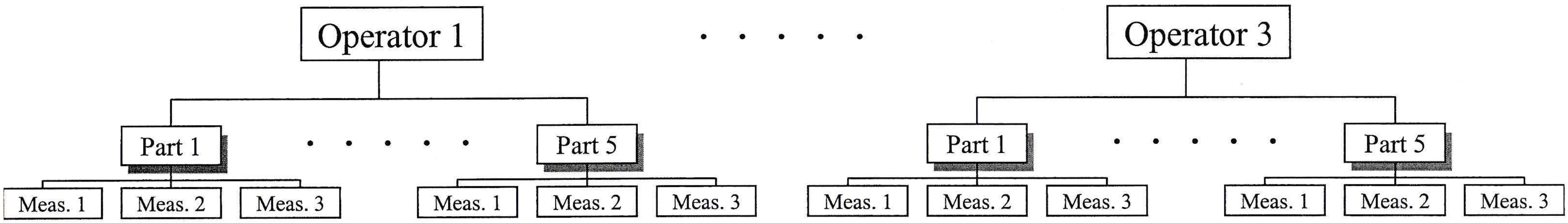
.... 3 Operators....

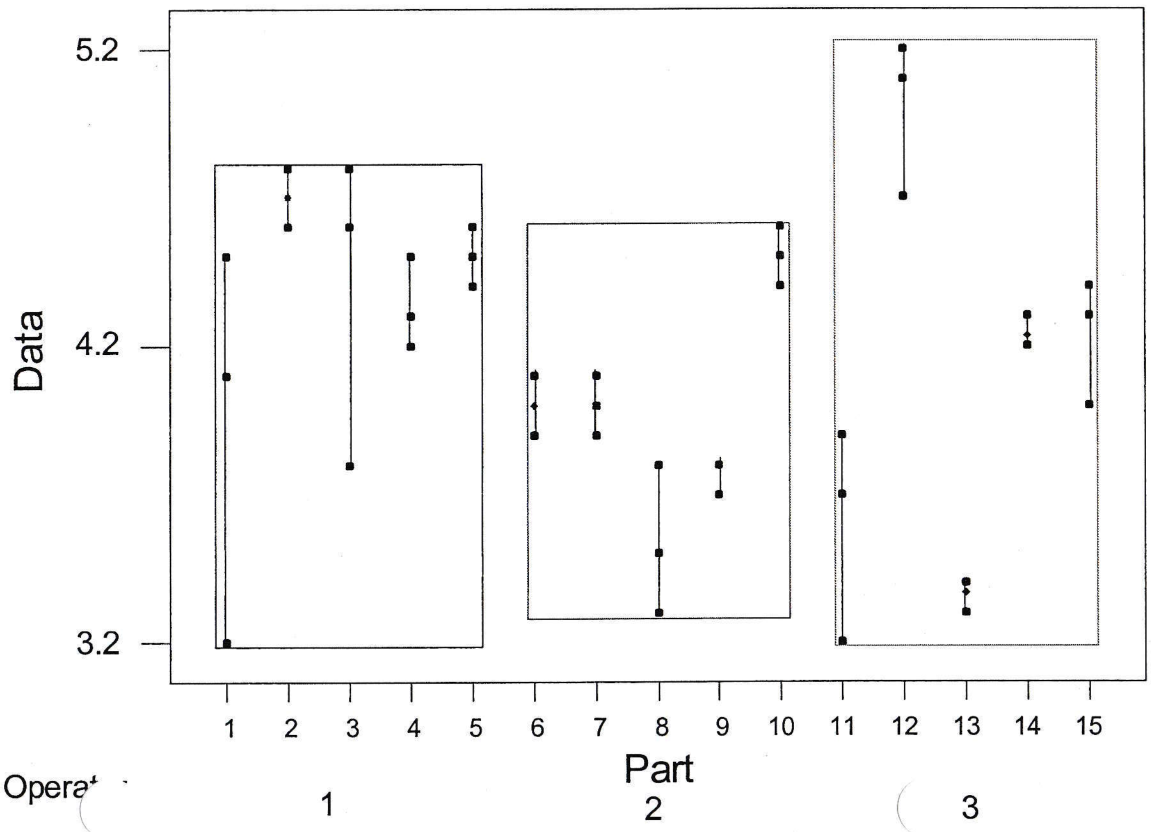


Extending the Sub-Grouping Concept

The Tech Group

- **Xbar & R chart schematic:**



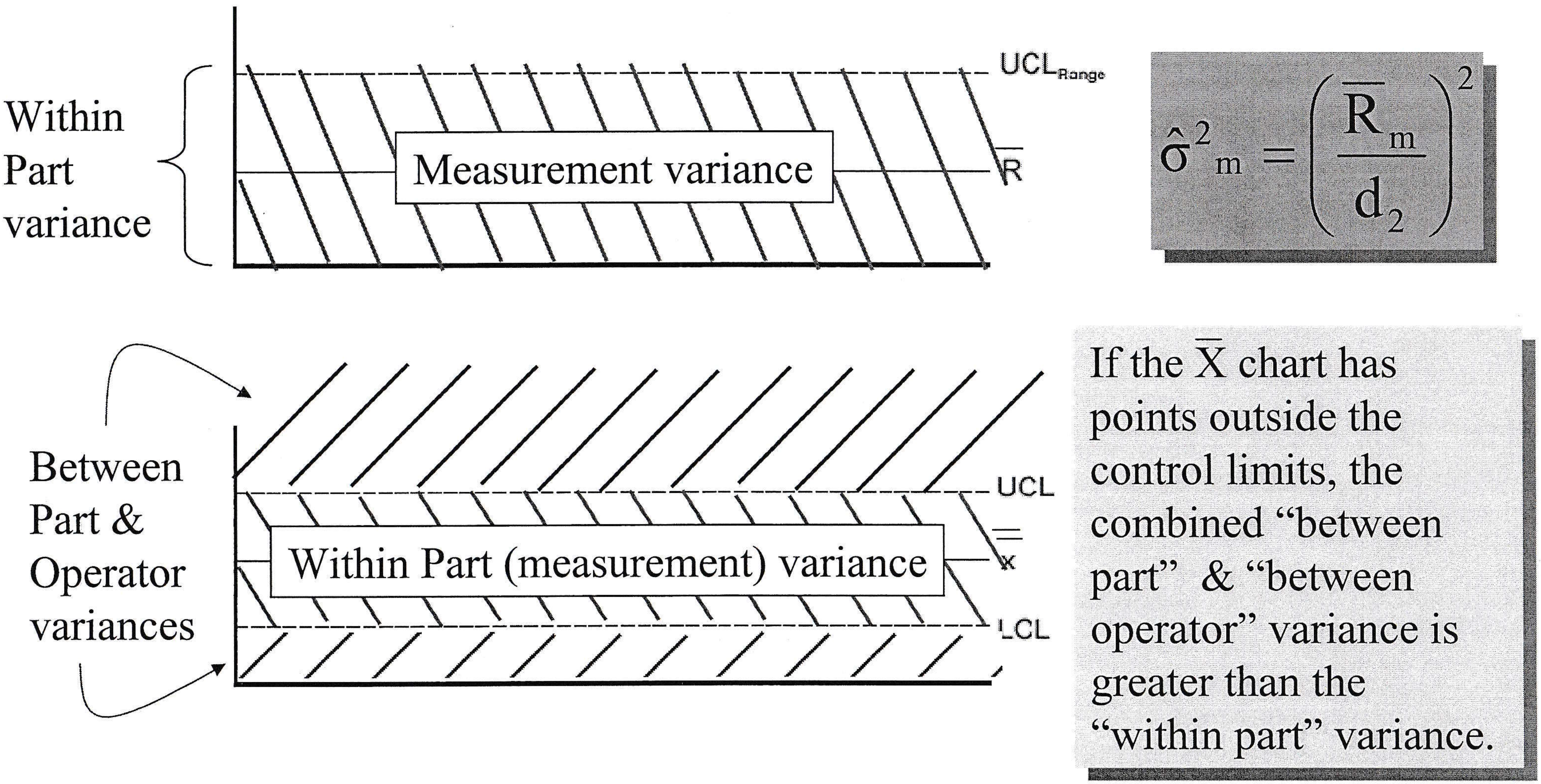


What is the biggest source of variation?

The Tech Group

COV - Measurement Variance

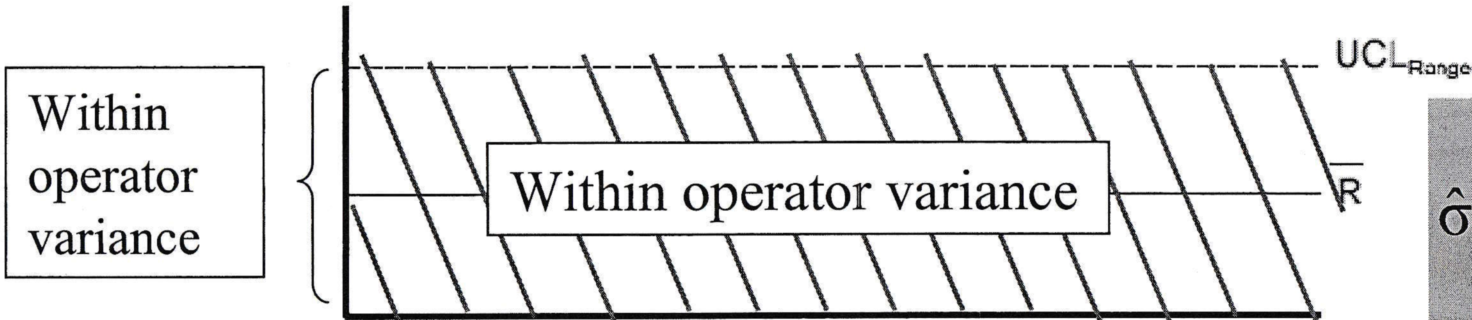
$\hat{\sigma}^2_{measurement}$



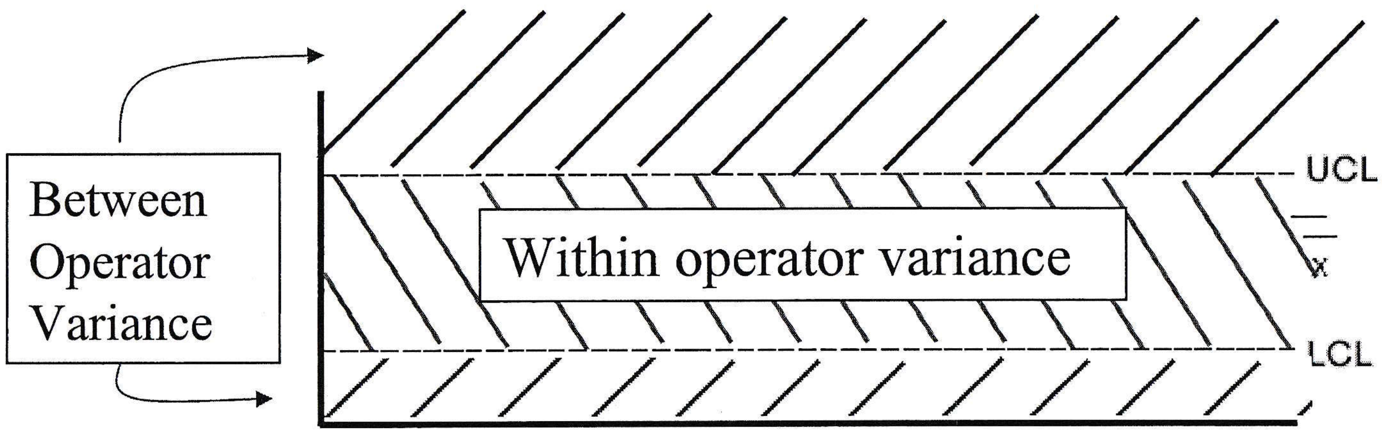
COV - Part Variance

The Tech Group

$\hat{\sigma}^{2}_{part}$



$$\hat{\sigma}^2_p = \left(\frac{\overline{R}_p}{d_2}\right)^2 - \frac{\hat{\sigma}^2_m}{2}$$



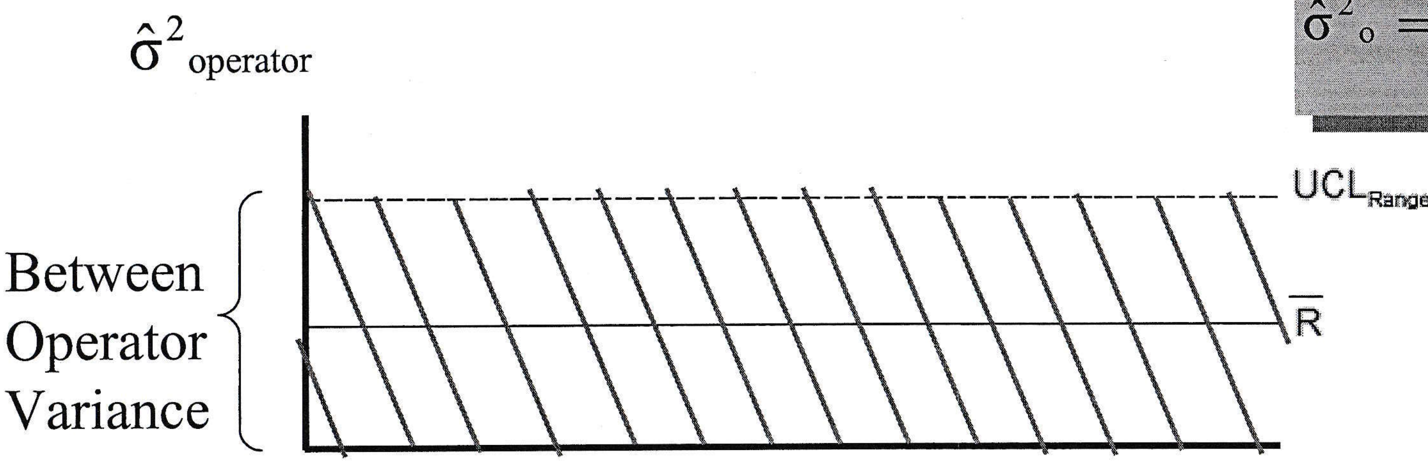
If $\overline{X}$ chart has points outside the control limits, the between operator variance is greater than the within operator variance.

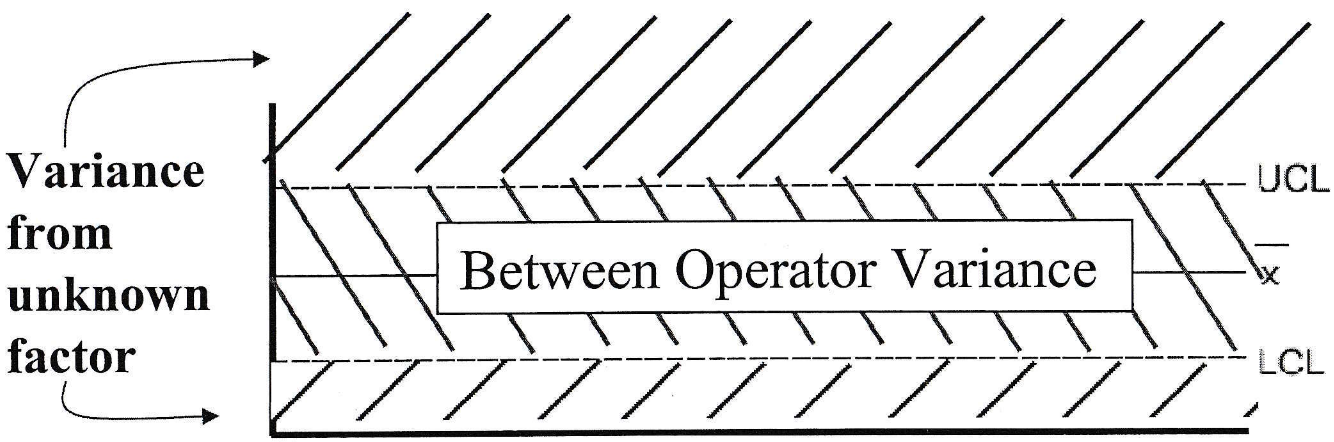
What is included in “within operator” variance?

The Tech Group

COV – Operator Variance

$$\hat{\sigma}^2_o = \left(\frac{MR_o}{d_2}\right)^2 - \frac{\hat{\sigma}^2_p}{2} - \frac{\hat{\sigma}^2_m}{4}$$





If $\overline{X}$ points go outside the control limits, there is some significant source of variation above operator that was overlooked.

Contribution by Component

The Tech Group

Recall, the total variation is expressed by:

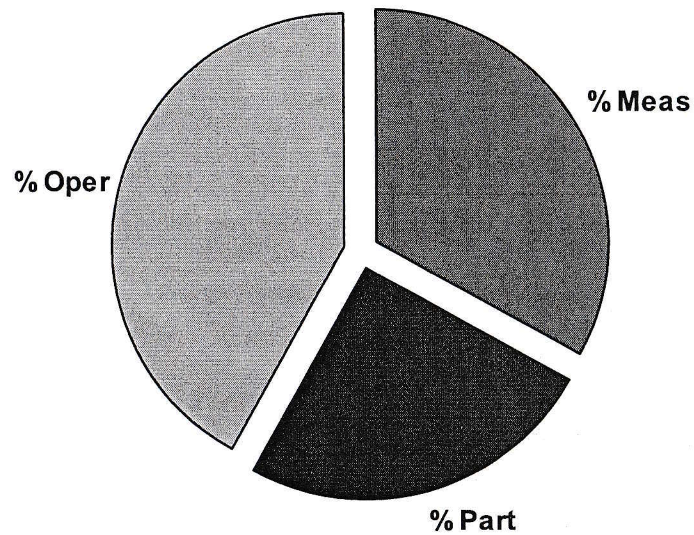
$$\hat{\sigma}^2_{Total} = \hat{\sigma}^2_{Meas} + \hat{\sigma}^2_{Part} + \hat{\sigma}^2_{Oper}$$

The percent contribution for each component is determined as shown:

$$\%\,\hat{\sigma}^2_{Meas} = \frac{\hat{\sigma}^2_{Meas}}{\hat{\sigma}^2_{Total}} \times 100$$

$$\%\,\hat{\sigma}^2_{Part} = \frac{\hat{\sigma}^2_{Part}}{\hat{\sigma}^2_{Total}} \times 100$$

$$\%\,\hat{\sigma}^2_{Oper} = \frac{\hat{\sigma}^2_{Oper}}{\hat{\sigma}^2_{Total}} \times 100$$



Control Chart Nested Studies

The Tech Group

What does this method do?

Partitions the overall process variation, into portions with assignable cause.

For instance, the total process variation might be attributable to a (1) within piece component, (2) a between piece and within lot component and a (3) between lot component of variation.

Why do it?

To evaluate the **stability** and **magnitude** of the various components of variation and therefore provide focus for the process improvement initiative.

Control Chart Nested Studies

The Tech Group

How to do it?

1. Identify the hierarchical causes (Factors). For instance, measurement (error), between parts within an operator, between operators, etc.

2. Design a sampling plan to collect data.

3. Analyze the data
    - Multi-vari charts
    - Control chart
    - Calculate the variance components

Black Belt Project #2

The Tech Group

- A Black Belt (BB) is assigned to a process.
- There is little scrap from this process, but the customer is beginning to demand reduced variation with one key product characteristic.
- The BB begins to map the process in order to develop an understanding of x's, y's, (including the Y that is important to the customer) and to start to uncover any hidden factory. Historical data is, unfortunately, not available.
- Thirty to forty of these parts are produced hourly. The BB wishes to characterize process variation and do an MSE. He samples two parts per hour for the 5 remaining hours of the day and has an operator measure each part three times.

Assignment: The data from the BB's study is shown in a table on the next page.

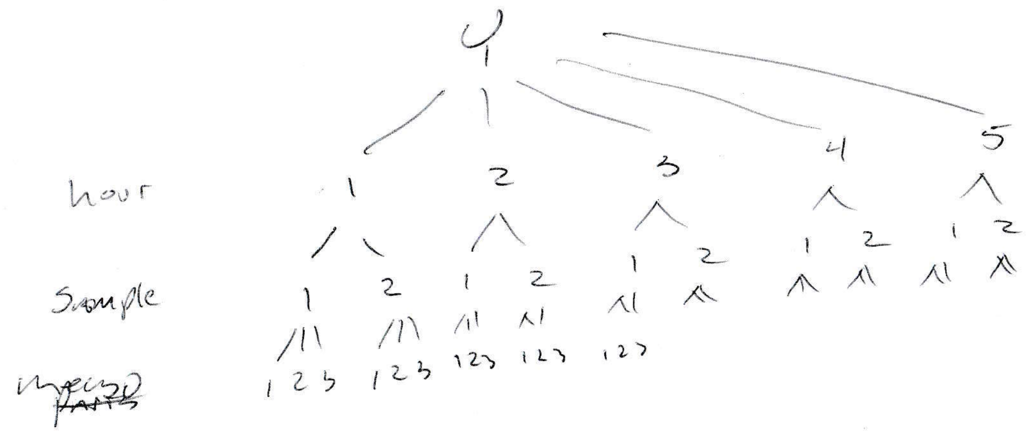
- Draw the "Tree" schematic showing the design.
- Finish the Data Tables.
- Interpret the Dot-Frequency Diagram.
- Complete the assignments as given.

Hour	Part	Meas	Y
1	1	1	1.45
1	1	2	2.33
1	1	3	2.26
1	2	1	3.56
1	2	2	3.68
1	2	3	3.63
2	3	1	4.74
2	3	2	4.76
2	3	3	5.09
2	4	1	5.5
2	4	2	5.77
2	4	3	5.39
3	5	1	4.2
3	5	2	4.14
3	5	3	4.16
3	6	1	2.84
3	6	2	3.1
3	6	3	2.77
4	7	1	3.85
4	7	2	4.02
4	7	3	3.67
4	8	1	2.46
4	8	2	2.16
4	8	3	2.39
5	9	1	6.27
5	9	2	5.72
5	9	3	5.81
5	10	1	3.81
5	10	2	3.87
5	10	3	4.08

Data Table

The Tech Group

Draw the tree schematic below:

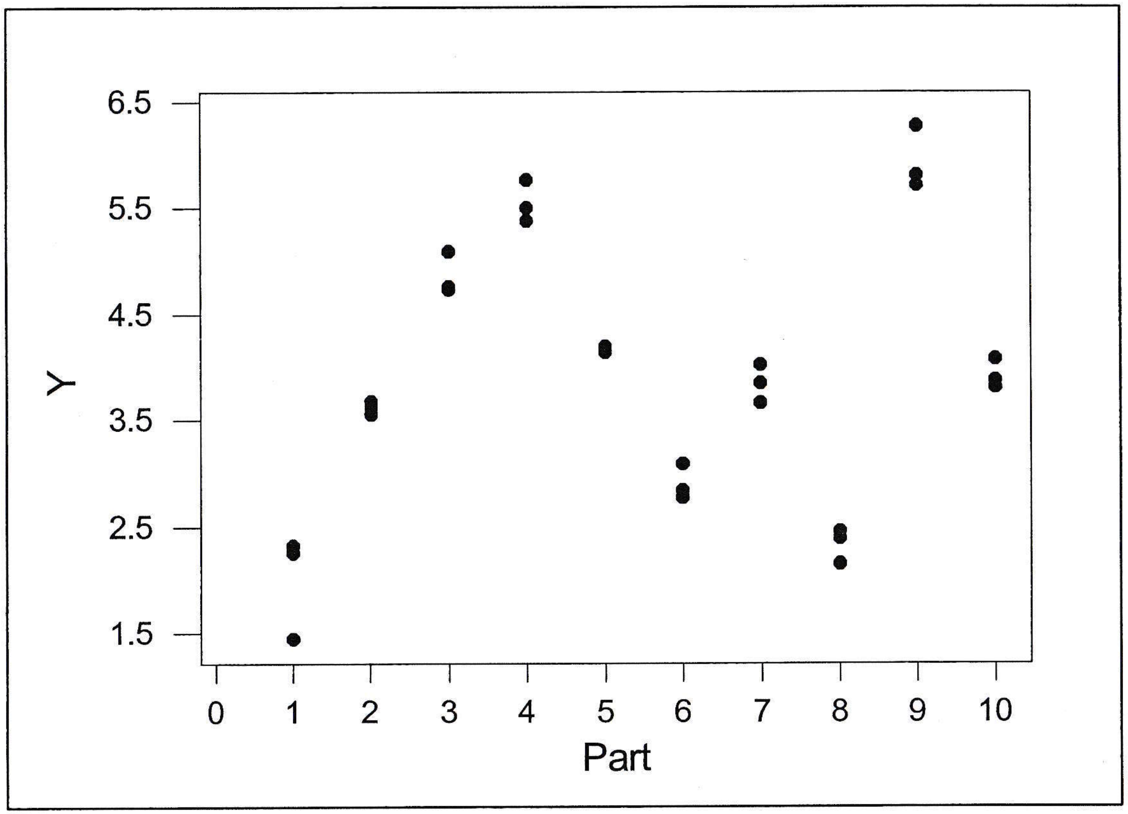


Dot Frequency Diagram

The Tech Group

Question 1: Is the largest component of variation associated with the measurement process or the manufacturing process?

Question 2: What additional information would be provided by control charts?



Assignment:

Plot the points and calculate the control limits for this data on the graphs shown on the following pages

(The control chart equations and constants located in the Appendix).

Data Table

Hour	Part	Meas	Y	R_M	X_{barM}
1	1	1	1.45		
1	1	2	2.33	0.88	2.01
1	1	3	2.26		
1	2	1	3.56		
1	2	2	3.68	0.12	3.623
1	2	3	3.63		
2	3	1	4.74		
2	3	2	4.76	.35	4.863
2	3	3	5.09		
2	4	1	5.5		
2	4	2	5.77	.35	5.553
2	4	3	5.39		
3	5	1	4.2		
3	5	2	4.14	.06	4.167
3	5	3	4.16		
3	6	1	2.84		
3	6	2	3.1	.33	2.903
3	6	3	2.77		
4	7	1	3.85		
4	7	2	4.02	.35	3.817
4	7	3	3.67		
4	8	1	2.46		
4	8	2	2.16	.30	2.337
4	8	3	2.39		
5	9	1	6.27		
5	9	2	5.72	.55	5.933
5	9	3	5.81		
5	10	1	3.81		
5	10	2	3.87	.27	3.92
5	10	3	4.08		

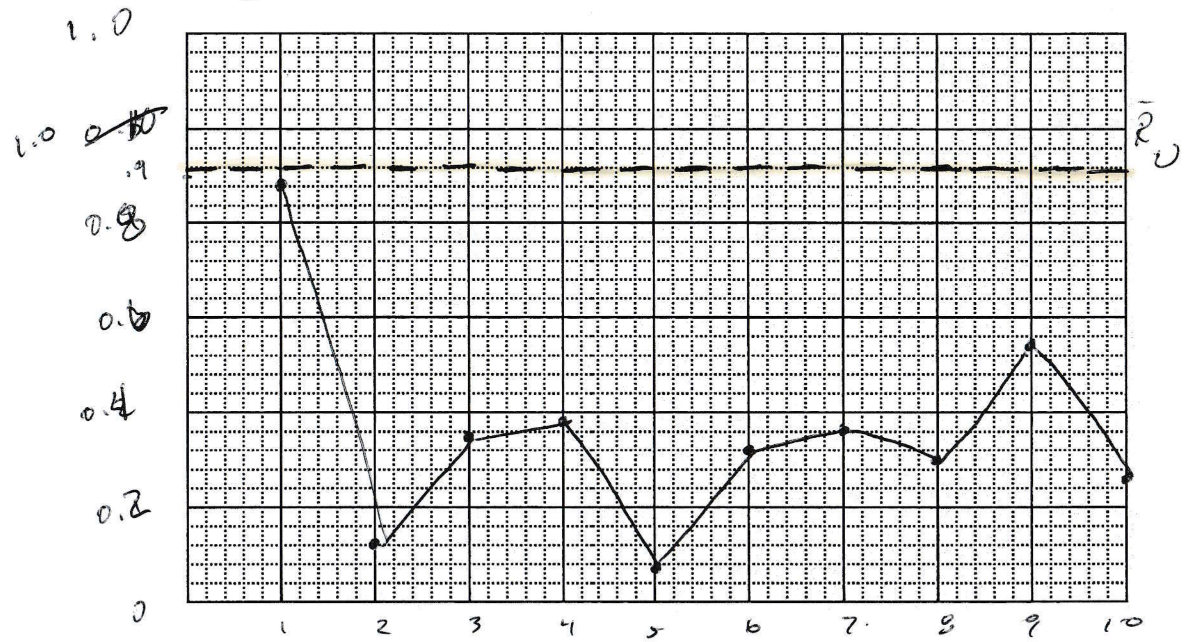
$\overline{R}_m$ = 0.359

$\overline{\overline{X}}_m$ = 3.916

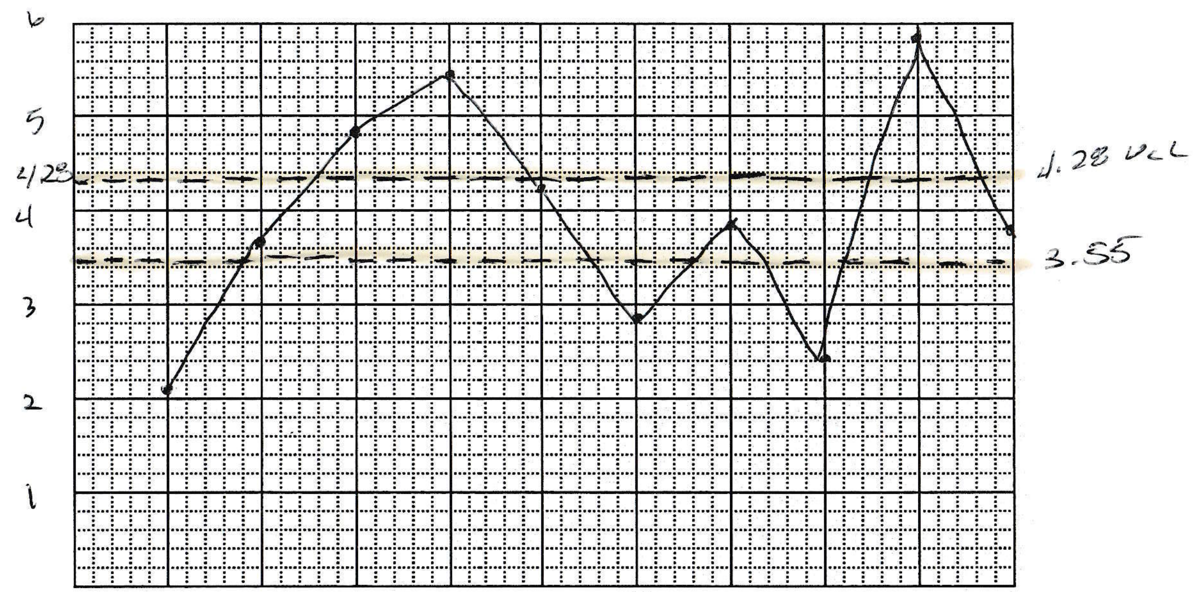
The Tech Group

Control Chart for Measurement Variance

Range



Average



$\bar{R}_U$ $\quad 2.574(\cancel{3.916}\ 0.359) = 0.924$

R_L $\quad 0$

$\bar{\bar{X}} \pm \bar{R}\,(1.023)$ $\quad .367257$

$3.916 + 0.359(1.023)$

$3.916 - 0.359(1.023)$

Project #2 Questions

The Tech Group

- Is the measurement process stable?
- What sources of variation contribute to differences in the ranges?
- What sources of variation contribute to differences in the averages?
- Is the variation in the manufacturing process greater than that due to the measurement process?
- Is the Part to Part variation stable?
- How would you investigate the Part to Part issue?
- What role does the process map have in developing a sampling strategy?

Data Table for "Part" Variance

$UCL_{\bar{R}} = 4.63$
$UCL_{\bar{\bar{X}}} = 6.58$
$LCL_{\bar{\bar{X}}} = 1.25$

$UCL_{\bar{R}} = 3.268\,(\bar{R})$

~~UC~~ $= 3.916 \pm 1.880\,(\bar{R})$

$\bar{R}_p = $ 1.419

$\bar{\bar{X}}_p = $ 3.916

H	Part	Meas	Y	R_M	X_{barM}	R_P	X_{barP}
1	1	1	1.45				
1	1	2	2.33	0.88	2.01		
1	1	3	2.26				
1	2	1	3.56			1.61	2.82
1	2	2	3.68	0.12	3.62		
1	2	3	3.63				
2	3	1	4.74				
2	3	2	4.76	0.35	4.86		5.205
2	3	3	5.09				
2	4	1	5.5			.69	
2	4	2	5.77	0.38	5.55		
2	4	3	5.39				
3	5	1	4.2				3.535
3	5	2	4.14	0.06	4.17	1.27	
3	5	3	4.16				
3	6	1	2.84				
3	6	2	3.1	0.33	2.90		
3	6	3	2.77				
4	7	1	3.85				3.095
4	7	2	4.02	0.35	3.85	1.51	
4	7	3	3.67				
4	8	1	2.46				
4	8	2	2.16	0.30	2.34		
4	8	3	2.39				
5	9	1	6.27			2.01	
5	9	2	5.72	0.55	5.93		4.925
5	9	3	5.81				
5	10	1	3.81				
5	10	2	3.87	0.27	3.92		
5	10	3	4.08				

4.5 UCL$_{\bar{R}}$ 4.63

4.0

3.5

The Tech Group

Control Chart for Part Variance

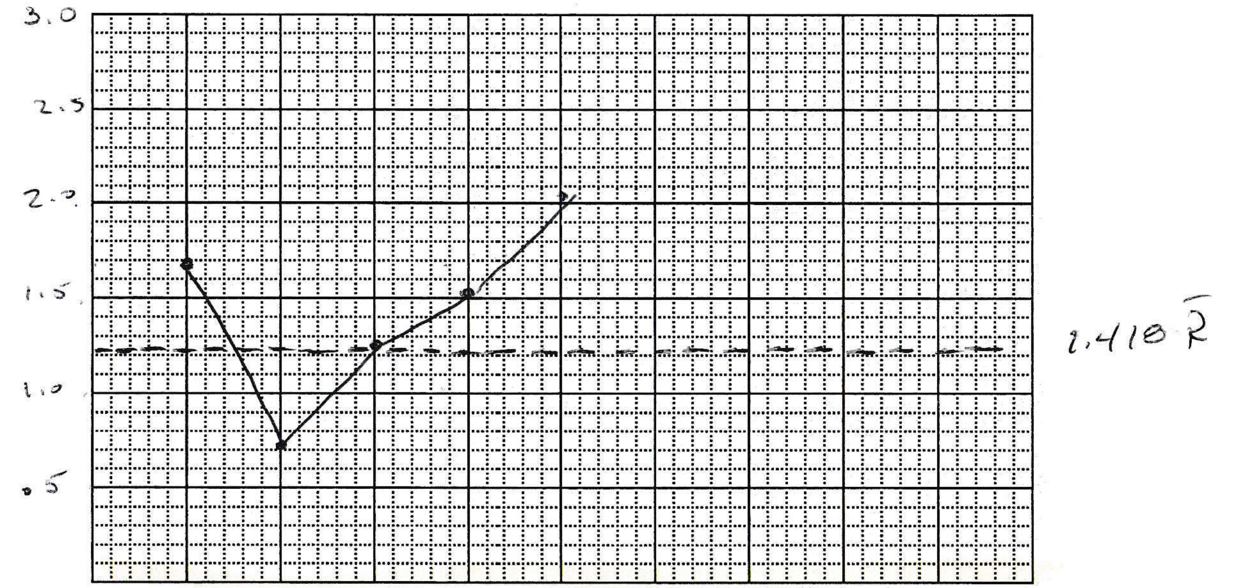
max ~~.88~~ 2.01
min ~~.06~~ .69

min 2.82

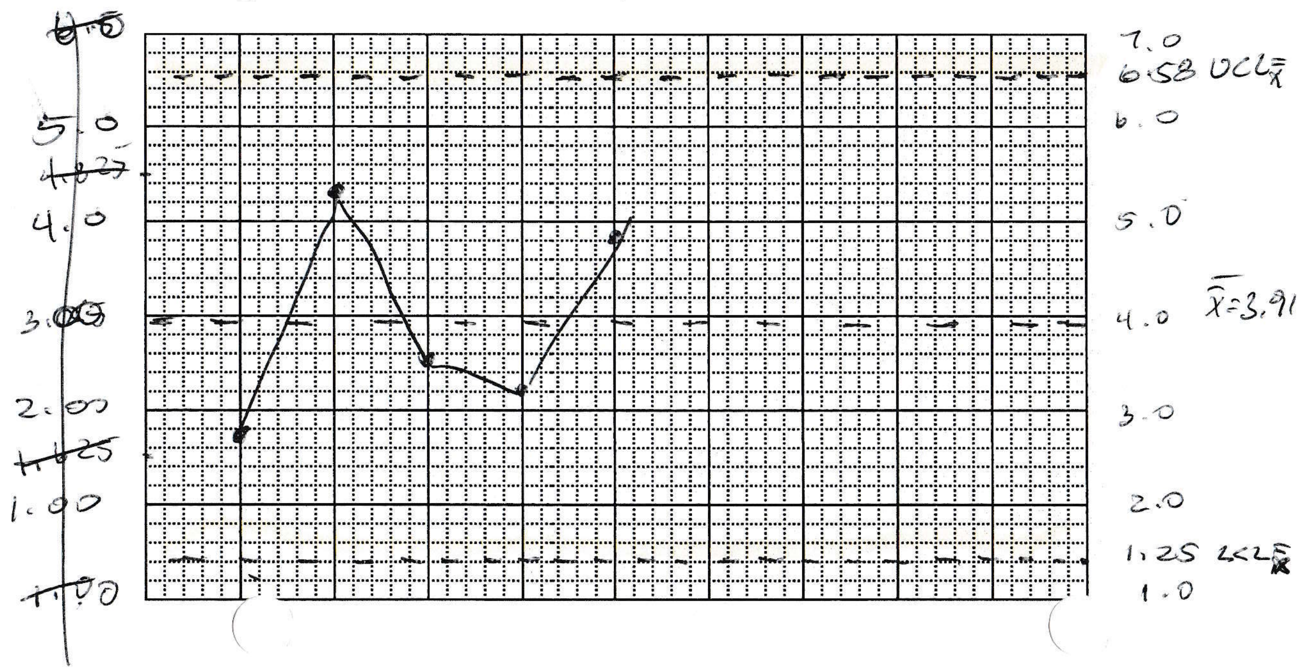
max 5.21

.082

Range



Average



Data Table for “Hour” Variance

	H_0 is True	H_0 is False
Accept H_0	No error ~~True~~	β
Reject H_0	α	No error True

$m\overline{R}_H$ = 1.5875

$\overline{X}_H$ = ~~3.908~~ 3.918

$\sqrt{2} = 3($

Hour	P	Meas	Y	R_M	X_{barM}	R_P	X_{barP}	m	I_{HR}
1	1	1	1.45	0.88	2.01	1.61	2.82		2.82
1	1	2	2.33						
1	1	3	2.26						
1	2	1	3.56	0.12	3.62				
1	2	2	3.68						
1	2	3	3.63						
2	3	1	4.74	0.35	4.86	0.69	5.21	2.39	5.21
2	3	2	4.76						
2	3	3	5.09						
2	4	1	5.5	0.38	5.55				
2	4	2	5.77						
2	4	3	5.39						
3	5	1	4.2	0.06	4.17	1.26	3.54	1.67	3.54
3	5	2	4.14						
3	5	3	4.16						
3	6	1	2.84	0.33	2.90				
3	6	2	3.1						
3	6	3	2.77						
4	7	1	3.85	0.35	3.85	1.51	3.09	0.45	3.09
4	7	2	4.02						
4	7	3	3.67						
4	8	1	2.46	0.30	2.34				
4	8	2	2.16						
4	8	3	2.39						
5	9	1	6.27	0.55	5.93	2.01	4.93	1.84	4.93
5	9	2	5.72						
5	9	3	5.81						
5	10	1	3.81	0.27	3.92				
5	10	2	3.87						
5	10	3	4.08						

Control Chart for Hour Variance

$\bar{R}$ min 0.45
max 2.39

$\bar{R} = 1.59$

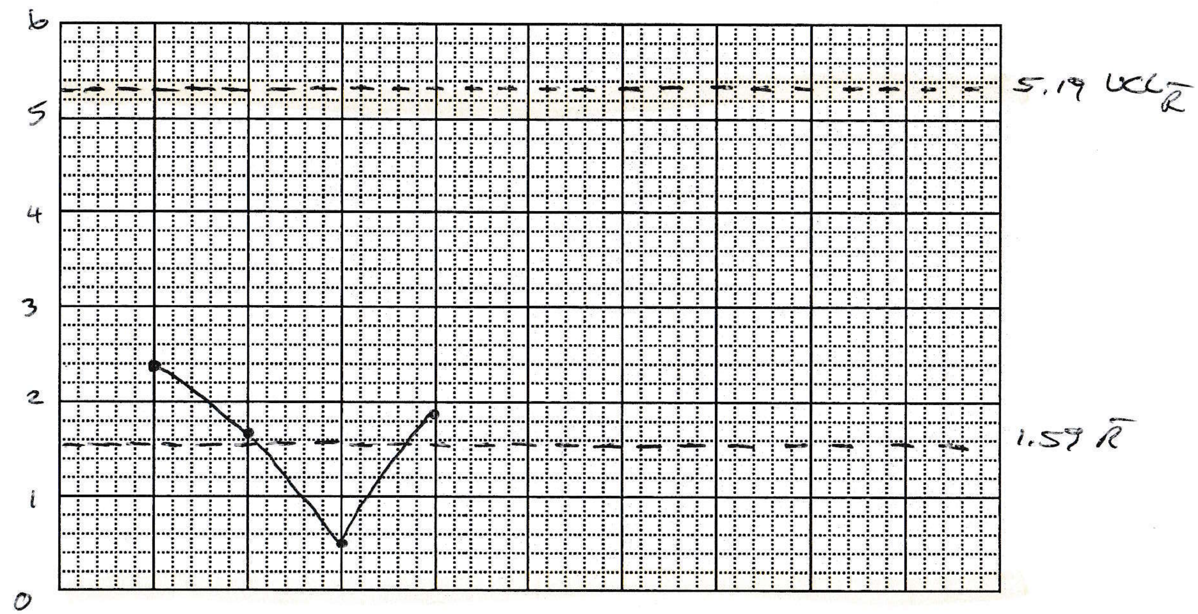
$\bar{\bar{X}} = 3.92$
min 2.82
max 5.21

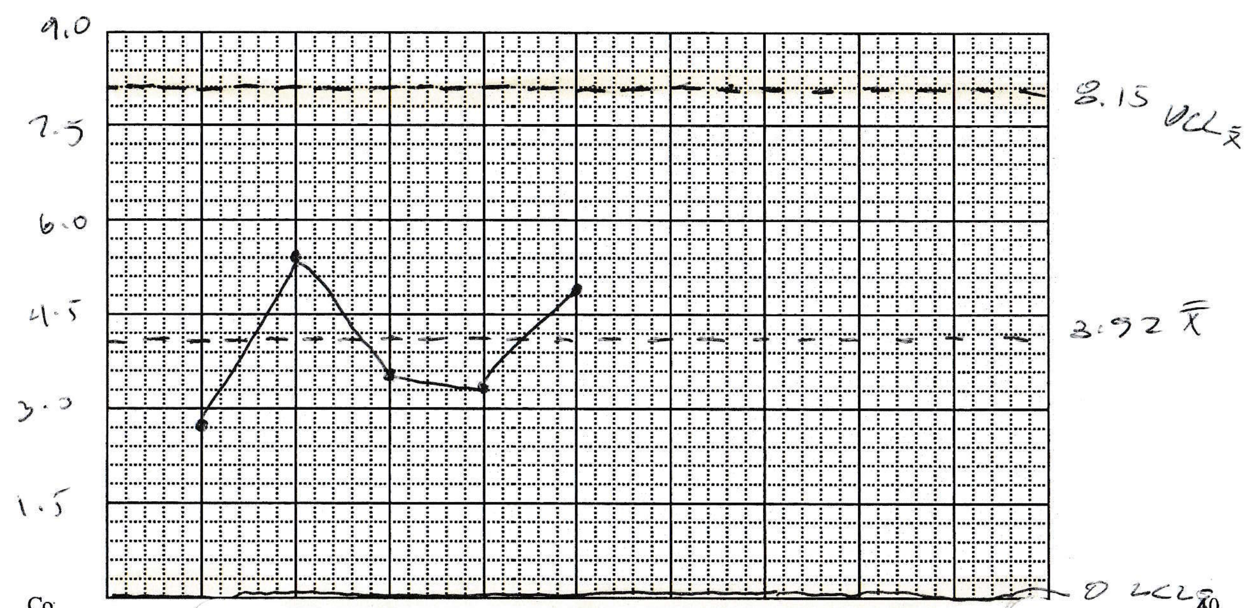
$CL = \bar{X} \pm \frac{\bar{R}}{d_2}$

$= 3.92 \pm 3\left(\frac{1.59}{1.128}\right)$

$= UCL_{\bar{X}}$ ~~5.33~~ 8.15
$LCL_{\bar{X}}$ ~~2.51~~ 0

$UCL_{\bar{R}}$ 1.59(3.267)
5.19





Control Chart for Hour Variance

$\bar{R}$ min 0.45
max 2.39

$\bar{R} = 1.59$

$\bar{\bar{X}} = 3.92$
min 2.82
max 5.21

$CL = \bar{X} \pm \frac{\bar{R}}{d_2}$

$= 3.92 \pm 3\left(\frac{1.59}{1.128}\right)$

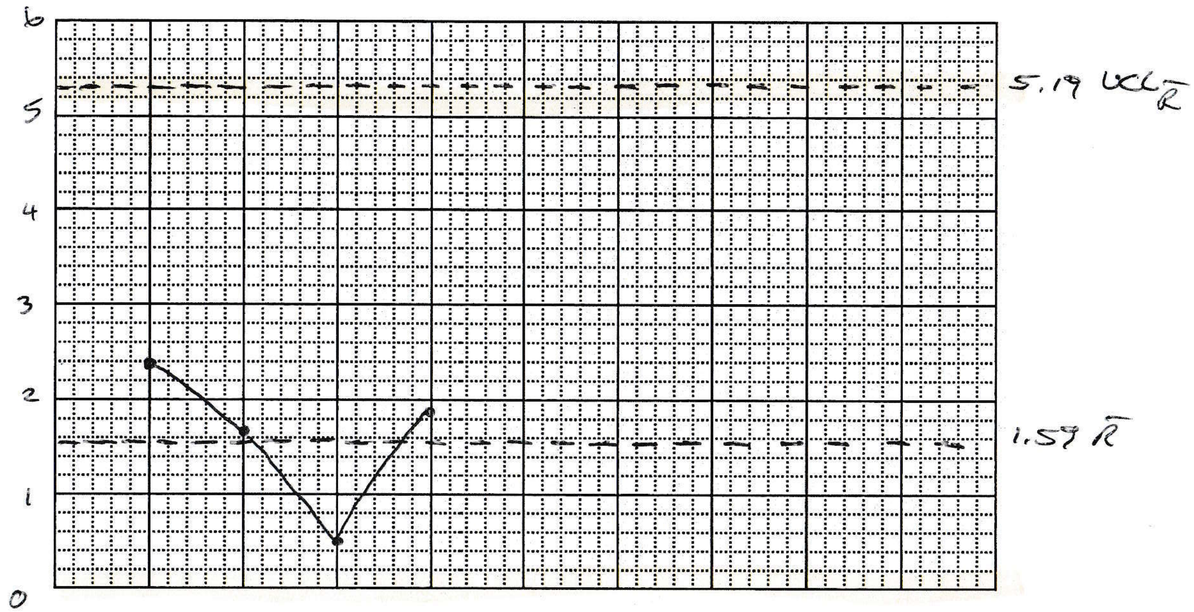
$= UCL_{\bar{X}}$ ~~5.33~~ 8.15

$LCL_{\bar{X}}$ ~~2.51~~ 0

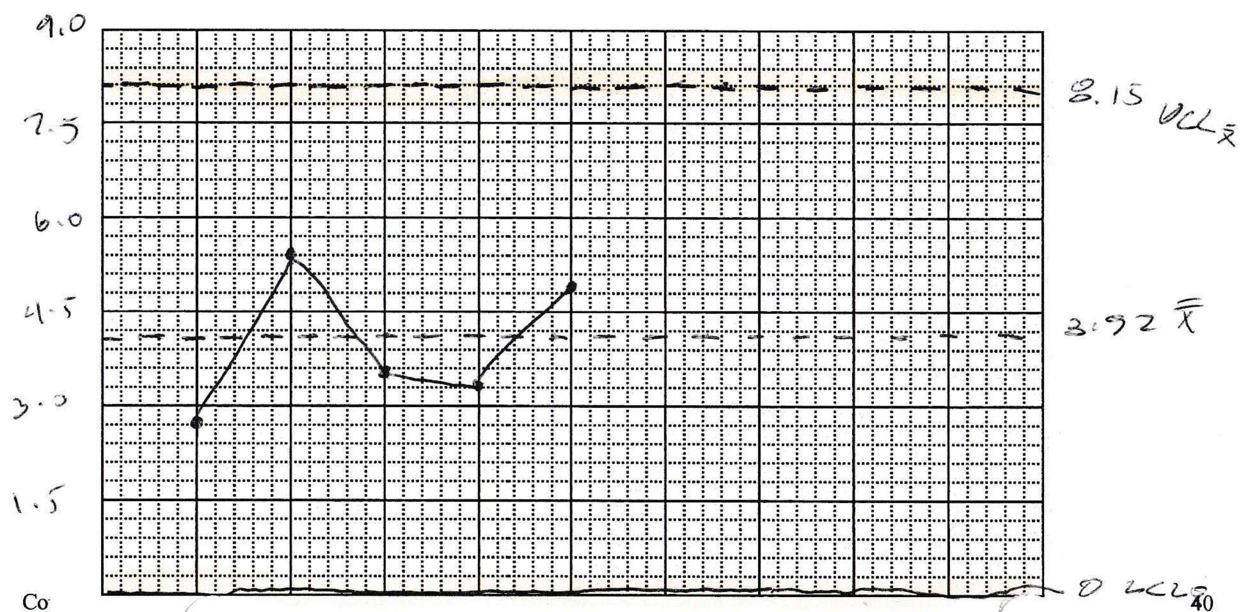
$UCL_{\bar{R}} = 1.59(3.267)$

5.19

Range



Average



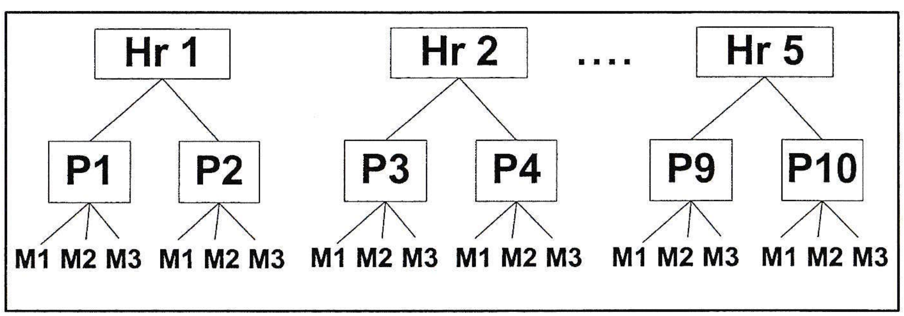
Co

Calculations

$$\hat{\sigma}^2_m = \left(\frac{\overline{R}_m}{d_2}\right)^2 = \left(\frac{.359}{1.693}\right)^2 = .04$$

$$\hat{\sigma}^2_p = \left(\frac{\overline{R}_p}{d_2}\right)^2 - \frac{\hat{\sigma}^2_m}{3} = \left(\frac{1.417}{1.128}\right)^2 - \frac{.04}{3} = 1.56$$

$$\hat{\sigma}^2_H = \left(\frac{MR_H}{d_2}\right)^2 - \frac{\hat{\sigma}^2_p}{2} - \frac{\hat{\sigma}^2_m}{6} = \left(\frac{1.585}{1.128}\right)^2 - \frac{1.56}{2} - \frac{.04}{6} = 1.19$$



Variance Summary

Meas.	0.04	2%
Part	1.56	56%
Hr	1.19	42%
Total:	2.79	100%

ANOVA Method

The Tech Group

- ANOVA Assumptions
  - Independence of observations – **Important**
  - Homogeneous variance (Homoscedasticity)

    The variance within a factor are not substantially different from another. The F test is fairly robust against inequality of variances if the sample sizes are equal. (4x rule)
  - Normal Distribution (Dependent variable)
- Risks
  - Type 1 Error: Falsely conclude there is an effect.

Fully Nested ANOVA

The Tech Group

MiniTab Results:

Analysis of Variance for Y

Source	DF	SS	MS	F	P
Hour	4	28.3268	7.0817	2.146	0.212
Part	5	16.4967	3.2993	64.277	0.000
Error	20	1.0266	0.0513		
Total	29	45.8501			

Variance Components

Source	Var Comp.	% of Total	StDev
Hour	0.630	35.73	0.794
Part	1.083	61.36	1.041
Error	0.051	2.91	0.227
Total	1.764		1.328

Comparison of Methods

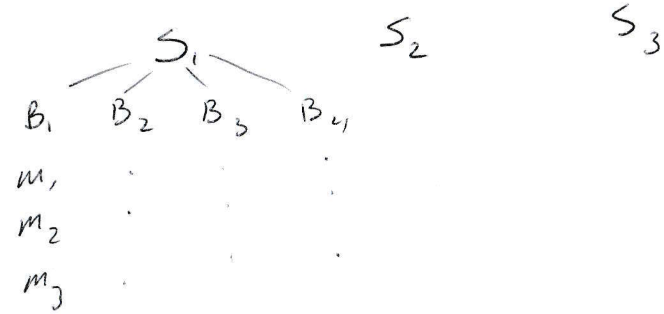
Source of Variation	R_{bar}/d_2	ANOVA
Hour	42%	35.7%
Part	56%	61.4%
Meas.	2%	2.9%
Total:	100%	100%

The Tech Group

COV Nested Exercise

A company buys batches of raw material from 3 suppliers. The variability in the purity of this raw material is adversely affecting the production of the finished product. A Black Belt would like to learn if this variability is due to differences between the suppliers. Four batches of raw material are selected from each supplier and 3 sample measurements are made on each batch.

- What are our factors?
- Draw the design tree (hierarchy).
- How would you analyze this data?



ANOVA Table

The Tech Group

Nested ANOVA: Purity versus Supplier, Batch

Analysis of Variance for Purity

Source	DF	SS	MS	F	P
Supplier	2	1.2606	0.6303	0.055	0.946
Batch	9	102.4883	11.3876	85.053	0.000
Error	24	3.2133	0.1339		
Total	35	106.9622			

Variance Components

Source	Var Comp.	% of Total	StDev
Supplier	-0.896*	0.00	0.000
Batch	3.751	96.55	1.937
Error	0.134	3.45	0.366
Total	3.885		1.971

* Value is negative, and is estimated by zero.

Expected Mean Squares

```
 1 Supplier  1.00(3) +  3.00(2) + 12.00(1)
 2 Batch     1.00(3) +  3.00(2)
 3 Error     1.00(3)
```

What do these results tell us?

Nested Study Analyzed as Crossed (con't)

ANOVA: Purity versus Supplier, Batch

Factor	Type	Levels	Values			
Supplier	fixed	3	1	2	3	
Batch	random	4	1	2	3	4

Analysis of Variance for Purity

Source	DF	SS	MS	F	P
Supplier	2	1.2606	0.6303	0.05	0.951
Batch	3	27.3133	9.1044	0.73	0.572
Supplier*Batch	6	75.1750	12.5292	93.58	0.000
Error	24	3.2133	0.1339		
Total	35	106.9622			

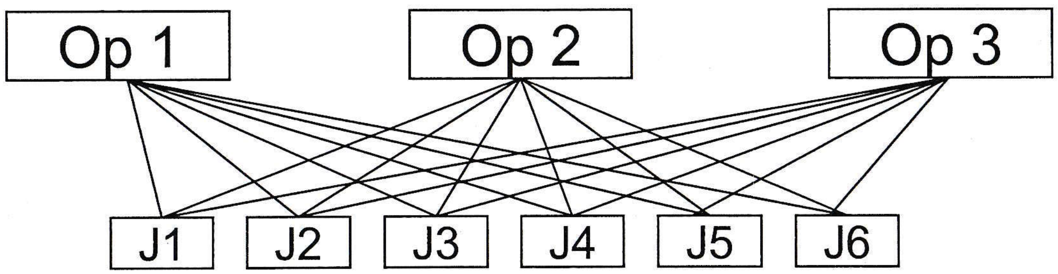
What is the practical significance of batch crossed with supplier?

What is a supplier by batch interaction?

Crossed ANOVA

The Tech Group

A Black Belt is investigating the possibility of assigning time standards to several job tasks. From videotape analysis, six job tasks are selected and each job is given to 3 operators (20 different operators can perform these tasks). Each operator completes each of the jobs twice at different times during the week.



How would you analyze the data?

Based on the analysis, what actions would you recommend regarding assigning time standards for each job task?

Crossed ANOVA

The Tech Group

ANOVA: Time versus Operator, Job

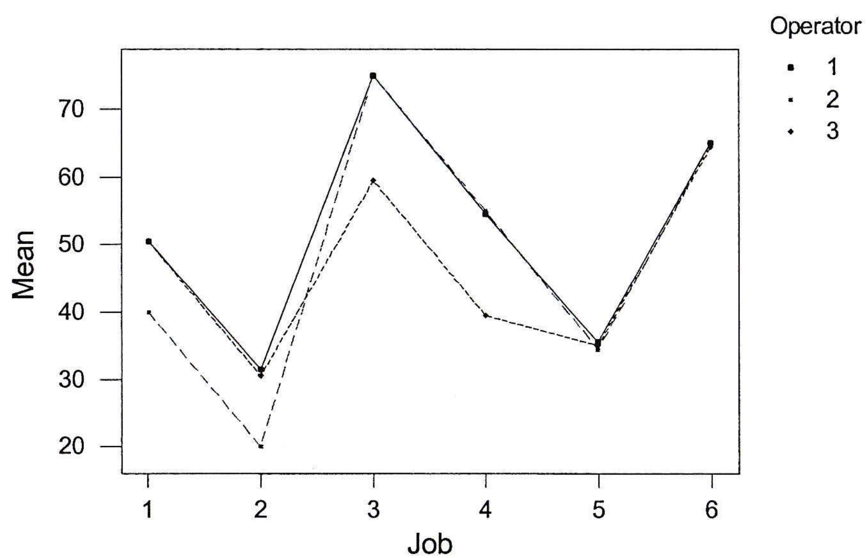
Factor	Type	Levels	Values					
Operator	random	3	1	2	3			
Job	random	6	1	2	3	4	5	6

Analysis of Variance for Time

Source	DF	SS	MS	F	P
Operator	2	184.72	92.36	1.22	0.335
Job	5	8127.56	1625.51	21.48	0.000
Operator*Job	10	756.61	75.66	17.24	0.000
Error	18	79.00	4.39		
Total	35	9147.89			

Source	Variance	% Variance
1 Operator	1.392	0.5%
2 Job	258.308	86.2%
3 Operator*Job	35.636	11.9%
4 Error	4.389	1.5%

Interaction Plot - Data Means for Time



Random Factors Analyzed as Fixed

The Tech Group

ANOVA: Time versus Operator, Job

Factor	Type	Levels	Values					
Operator	fixed	3	1	2	3			
Job	fixed	6	1	2	3	4	5	6

Analysis of Variance for Time

Source	DF	SS	MS	F	P
Operator	2	184.72	92.36	21.04	0.000
Job	5	8127.56	1625.51	370.37	0.000
Operator*Job	10	756.61	75.66	17.24	0.000
Error	18	79.00	4.39		
Total	35	9147.89			

"Operator" is now important !

Crossed Study without the Interaction Term

The Tech Group

ANOVA: Time versus Operator, Job

Factor	Type	Levels	Values					
Operator	random	3	1	2	3			
Job	random	6	1	2	3	4	5	6

Analysis of Variance for Time

Source	DF	SS	MS	F	P
Operator	2	184.7	92.4	3.09	0.061
Job	5	8127.6	1625.5	54.47	0.000
Error	28	835.6	29.8		
Total	35	9147.9			

Source	Variance	% Variance
1 Operator	5.210	1.7%
2 Job	265.945	88.4%
3 Error	29.843	9.9%

With the Interaction

Source	Variance	% Variance
1 Operator	1.392	0.5%
2 Job	258.308	86.2%
3 Operator*Job	35.636	11.9%
4 Error	4.389	1.5%

Control Chart Constants

Sample Size	A_2	D_3	D_4	d_2
2	1.880	---	3.268	1.128
3	1.023	---	2.574	1.693
4	0.729	---	2.282	2.059
5	0.577	---	2.114	2.326
6	0.483	---	2.004	2.534
7	0.419	0.076	1.924	2.704
8	0.373	0.136	1.864	2.847
9	0.337	0.184	1.816	2.970
10	0.308	0.223	1.777	3.078

X_{bar} & R Control Chart Formulas for Calculating Control Limits

$$CL_{\bar{x}} = \bar{\bar{X}} +/- A_2\bar{R}$$

$$UCL_{\bar{R}} = D_4\bar{R}$$

$$LCL_{\bar{R}} = D_3\bar{R}$$ (for n>6)

Sample Size	A_2	D_3	D_4
2	1.880	—	3.268
3	1.023	—	2.574
4	0.729	—	2.282
5	0.577	—	2.114
6	0.483	—	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777