Measurement Systems Analysis—How's My Instrumentation?

Measurement Systems

- To judge the quality of process output, it must be measured.
- The measurement system consists of one or more gauges, measurement devices, instruments, analytical methods, or metrology tools.

Measurement Systems Analysis

- ***Measurement Systems Analysis*** (MSA) is the study of variation in a process due to the measurement system.
- MSA can be called the following:
  - Measurement capability analysis (MCA)
  - Gauge study
  - Repeatability and Reproducibility (R&R) study
  - Interlaboratory Uniformity (ILU) study
  - Method validation
  - Round robin study

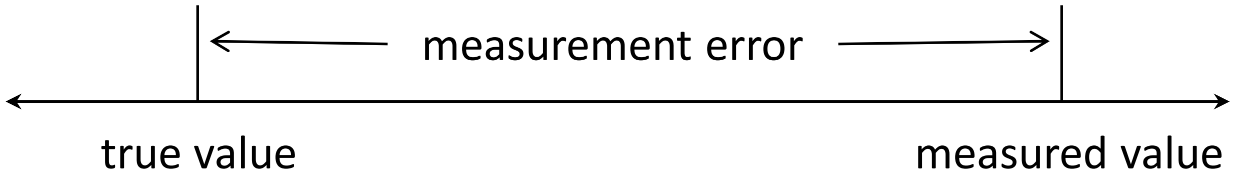
How Measurement Error Affects the Process

$$Y = truth + measurement\ error$$

$$\sigma^2_{process} = \sigma^2_{product} + \sigma^2_{measurement}$$

Measurement Error

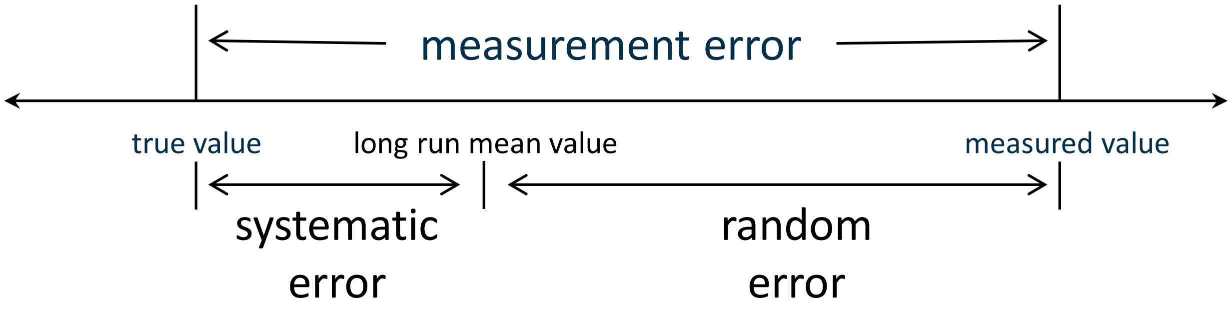
Measurement error is the difference between the measured value of a part and the true value of that part.



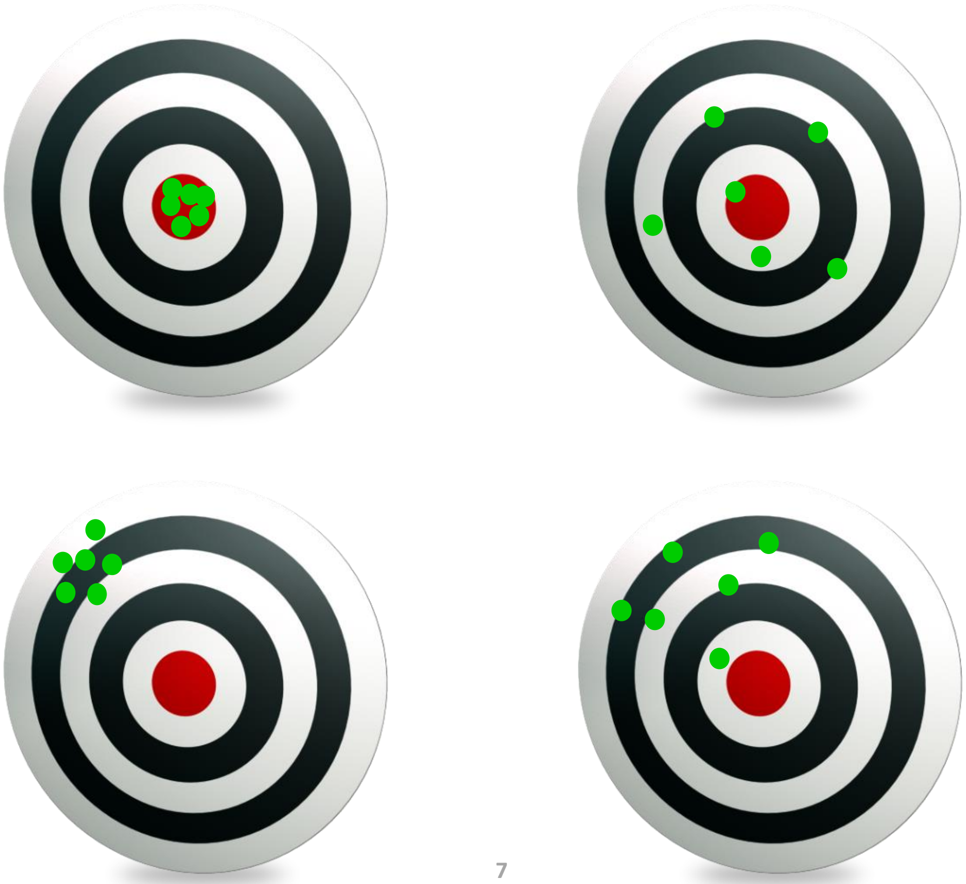
Measurement Error

Measurement error can be split into two parts.

- *Systematic error* is the difference between the long run mean value of the part and the true value.
- *Random error* is the difference between the measured value of the part and the long run mean value.



Systematic and Random Error



Systematic Error

- Studying the *bias* of the measurement system gives information about systematic error.
- Bias measures how far the system is from truth, on the average.
- Bias can be zero, constant, linear, or nonlinear.
- Bias studies need *standard reference material* for which the true value can be assumed known.

Random Error

- Studying the variability of the measurement system gives information about random error.
- The sources of variability in the measurement system are estimated.
  - *Repeatability* – the measurement error associated with multiple measurements taken under identical conditions. It is a lower limit on measurement system error.
  - *Reproducibility* – used to estimate total measurement system variability. It includes variability from any factors under study.

Part Variation

- MSA is typically performed using multiple parts, or standards.
- Variability due to differences in the parts should not be included in the estimate of measurement system variability.
- The same reference standards should be used throughout the study.

Destructive Testing

Destructive gauges cannot make repeated measurements on the same standard part.

- Collect parts that are as close to the same as possible, from the same run, lot, or batch. Assume that there is no part-to-part variation, and confound part variation with measurement system variation to obtain an upper bound on the true variability of the measurement system.
- Find a non-destructive measurement that correlates with the destructive measurement. Build a calibration curve relating the two measurements. Perform the gauge study on the non-destructive measurement. Model the destructive measurement predictions from the calibration curve.

Interactions with Parts

Interactions between the measurement system and the reference standard should be considered to be part of the measurement system variability.

- For example, if the measurement system causes the part's true value to change
- Sometimes considered to be a destructive gauge

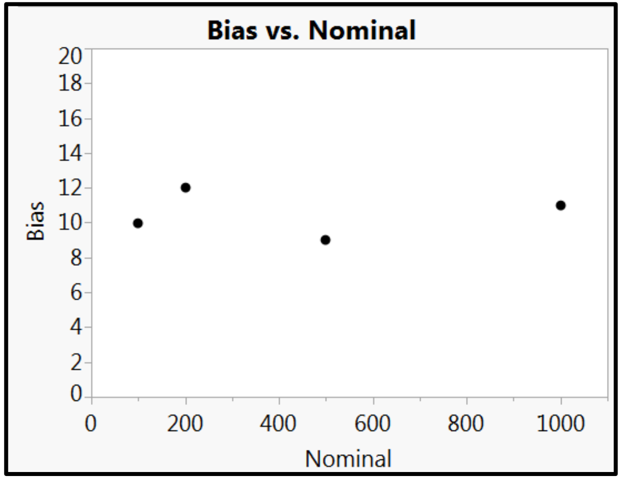
Precision and Accuracy

- *Precision* is an estimate of the total variability of the random error.
  - Often calculated as a multiple of reproducibility (for example, $6\sigma_{R\&R}$).
- *Accuracy* is a qualitative term describing the closeness of the mean of a group of measurements of a standard and the true value of the standard.
  - Depends on both bias and precision.
  - Not used interchangeably with bias or precision.

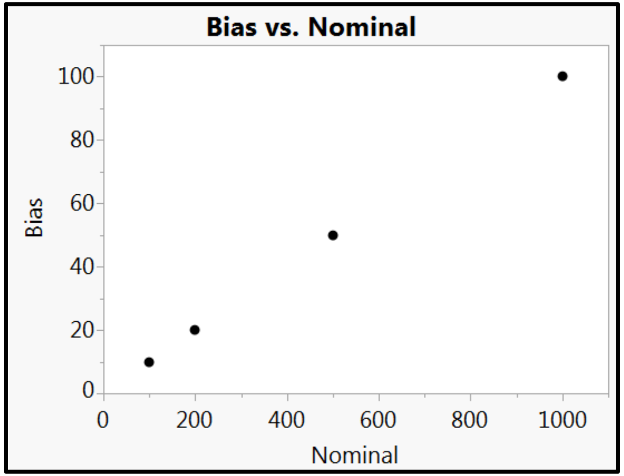
Linearity

Linearity is the absence of variability due to bias over the measurement range.

- Some authors also include the absence of variability due to reproducibility over the measurement range.



Constant bias

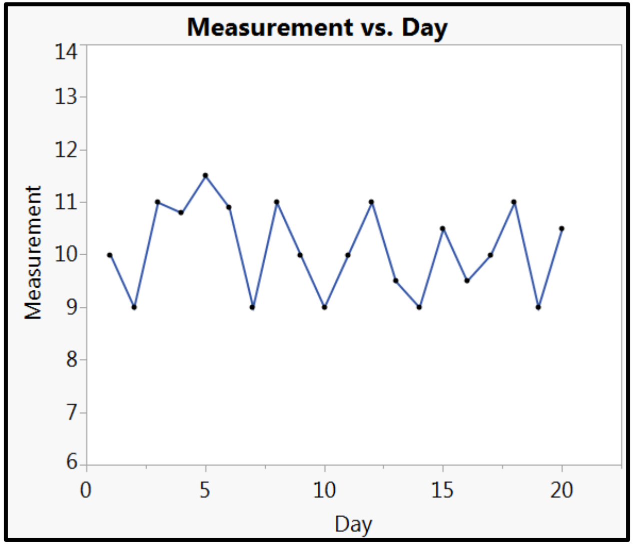


Non-constant bias

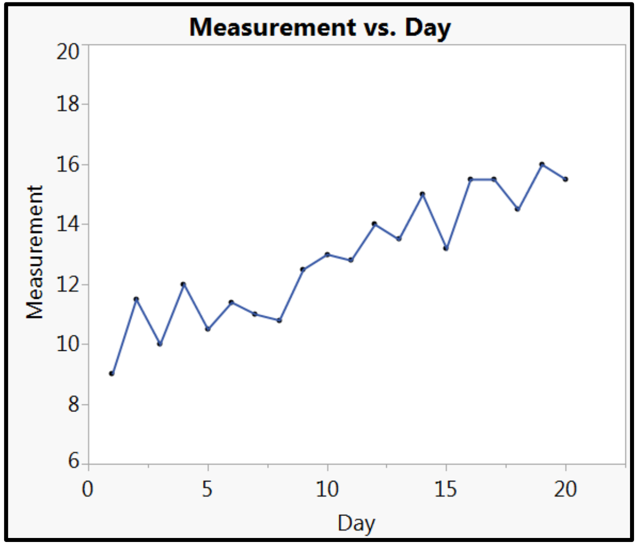
Stability

Stability is the absence of additional variability due to taking measurements over time.

- A component of reproducibility.



Stable

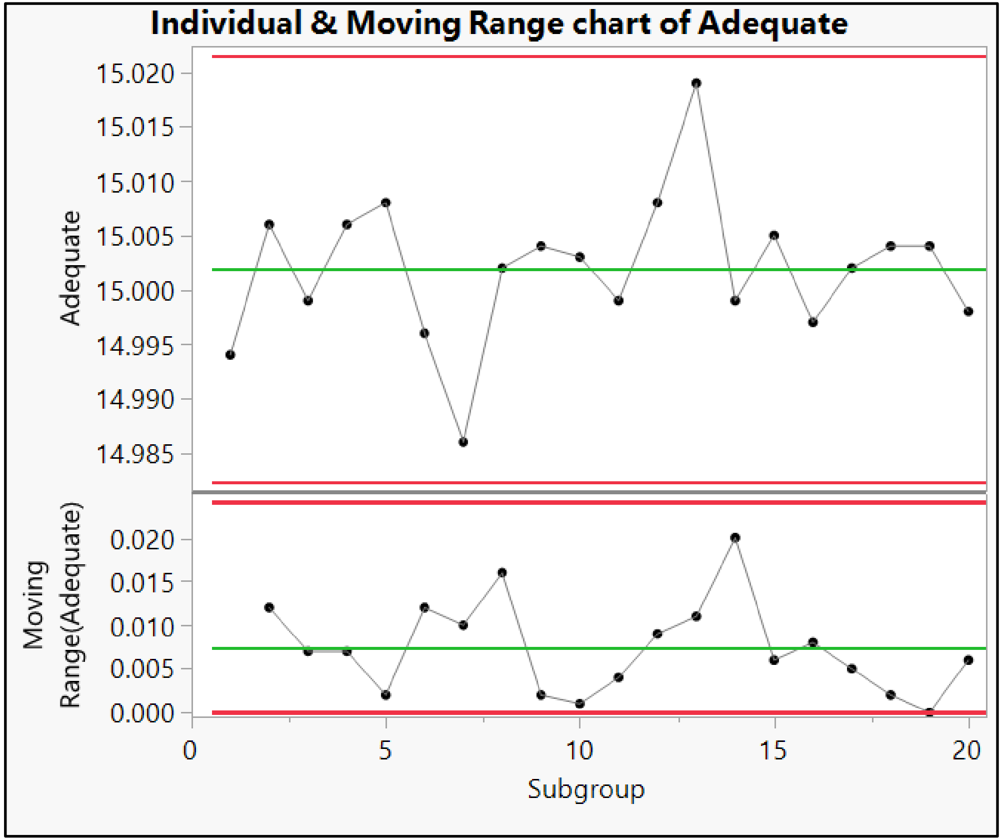


Not stable

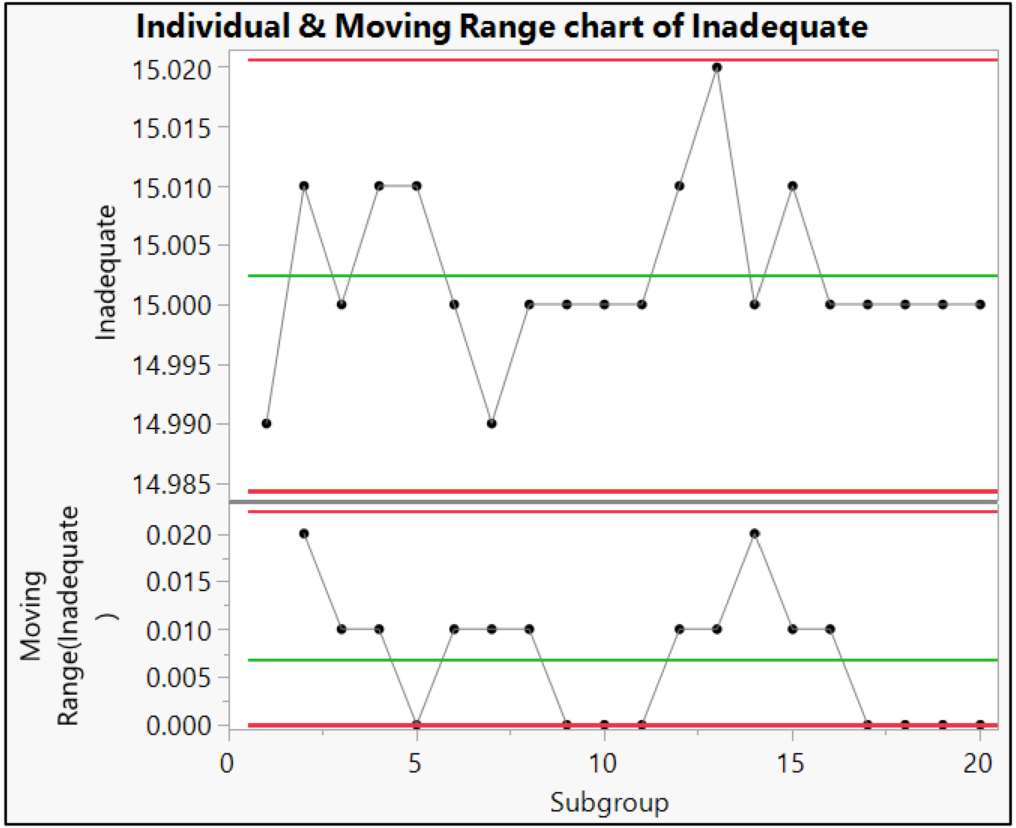
Discrimination

- *Discrimination* is the ability of a measurement system to distinguish differences among parts.
- The *resolution* of a measurement system is how many significant digits the measurement system displays.
- Discrimination might be adequate if there are at least
  - four possible values inside the control limits on an SPC chart for the range or standard deviation
  - five possible values inside the control limits on an SPC chart for individuals or means.

Adequate Discrimination



Inadequate Discrimination



Remedies to Inadequate Discrimination

- Increase measurements by at least one decimal place.
- Use a new measurement system.
  - Compare measurement systems using the techniques of this course.
- Take data over a wider time period to increase the “within” variability.
- Take multiple measurements and average.

Design of Experiments

- A gauge study is a designed experiment.
  - a set of trials run on a process to learn something about that process
- Statistically designed experiments are efficient.
  - more information out of fewer resources

Methodology

- State the objective
- Identify potential sources of variability
- Gather standards
- Schedule trials
- Generate data-entry form
- Collect data (run the experiment)
- Analyze results
- Draw conclusions from analysis

State the Objective

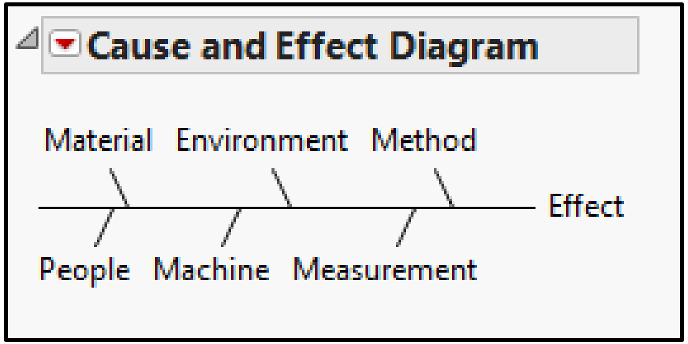
- The objective of a gauge study is typically to quantify the random variability in the measurement system.
- Other objectives are possible.
  - Determine variance components as a guide for measurement process improvement activities.
  - Determine whether excessive process variability is caused by metrology variability.
  - Determine misclassification probabilities in order to set guard band or specification limits.
  - Satisfy customer or auditor requirements.
- The objective statement must be specific.

Identify Potential Sources of Variability

- Assemble a team of those close to the process.
  - process engineers, maintenance, operators, and measurement engineers
- Brainstorm a list of potential sources of variability.
- Identify those on the list that will be deliberately varied during the MSA study.
- Those that are not deliberately varied will be documented and perhaps added to an out-of-control action plan on the gauge.

Organizing Your Factors

- Many measurement systems today are more complex than a single operator using a simple tool to measure the product.
- Other sources of variability need to be included in the study in order to best estimate the total measurement system variability.
- One tool to organize potential factors is a cause-and-effect diagram.



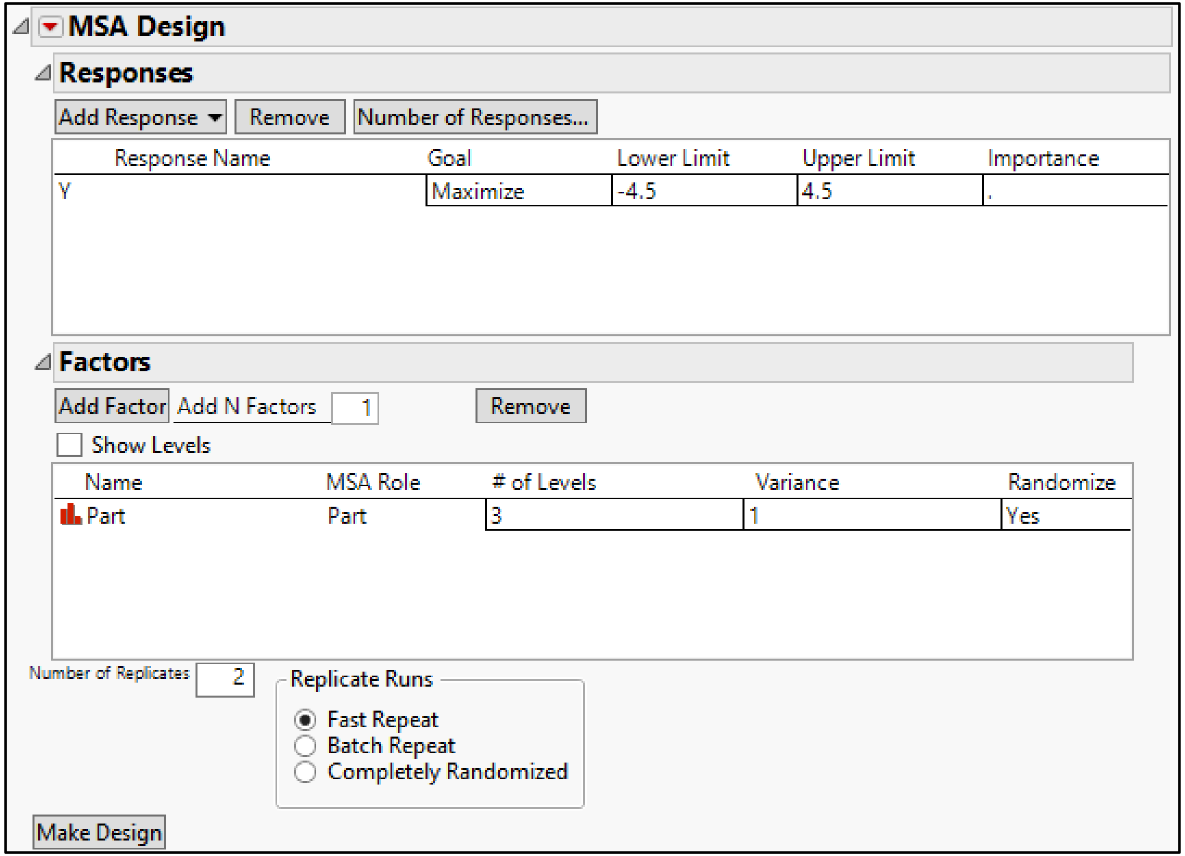
Gather Standards

- Select parts that are representative of production or standards that span the range of measurements and label them.
- What range of responses should be included?
- Does the part have multiple sites that need to be measured? If so, how many sites should be included?
- Are traceable standards available?
- Are appropriate internal standards available?

Schedule Trials

- Consider both sample size and sampling frequency when designing the experiment (more on this later).
- How long can the experiment run?
- How long can the gauge be used for the experiment?

Design the Experiment



Randomization Options in MSA Design

Fast Repeat

- Levels of first factor are randomized.
- Levels of second factor are randomized within first factor.
- Replicates are not randomized within levels of the second factor.

Run	Part	Operator	Replicates
1	L1	L3	1
2	L1	L3	2
3	L1	L2	1
4	L1	L2	2
5	L1	L1	1
6	L1	L1	2
7	L3	L2	1
8	L3	L2	2
9	L3	L1	1
10	L3	L1	2
11	L3	L3	1
12	L3	L3	2
13	L2	L2	1
14	L2	L2	2
15	L2	L1	1
16	L2	L1	2
17	L2	L3	1
18	L2	L3	2

Randomization Options in MSA Design

Batch Repeat

- Replicates are not randomized.
- Levels of first factor are randomized within replicates.
- Levels of second factor are randomized within first factor.

Run	Part	Operator	Replicates
1	L2	L1	1
2	L2	L3	1
3	L2	L2	1
4	L3	L2	1
5	L3	L1	1
6	L3	L3	1
7	L1	L3	1
8	L1	L2	1
9	L1	L1	1
10	L2	L2	2
11	L2	L1	2
12	L2	L3	2
13	L1	L3	2
14	L1	L1	2
15	L1	L2	2
16	L3	L1	2
17	L3	L2	2
18	L3	L3	2

Randomization Options in MSA Design

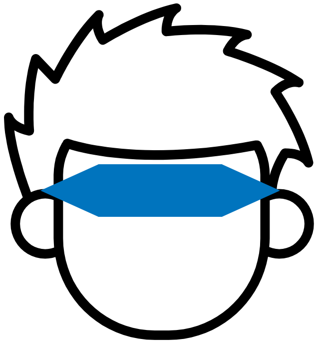
Completely Randomized

- All combinations of levels of first and second factors and replicates are randomized.
- Replicates are numbered in order.

Run	Part	Operator	Replicates
1	L3	L1	1
2	L1	L2	1
3	L3	L2	1
4	L3	L3	1
5	L3	L1	2
6	L3	L2	2
7	L1	L1	1
8	L2	L3	1
9	L2	L2	1
10	L1	L2	2
11	L2	L3	2
12	L2	L1	1
13	L3	L3	2
14	L1	L3	1
15	L1	L3	2
16	L2	L2	2
17	L1	L1	2
18	L2	L1	2

Collect Data and Run the Experiment

- Randomize the order of measurements as much as possible to protect against systematic biases of the measurements.
- If Operator is a factor, have the operators make blind measurements if possible. An operator should not know the identity of the part he is measuring.

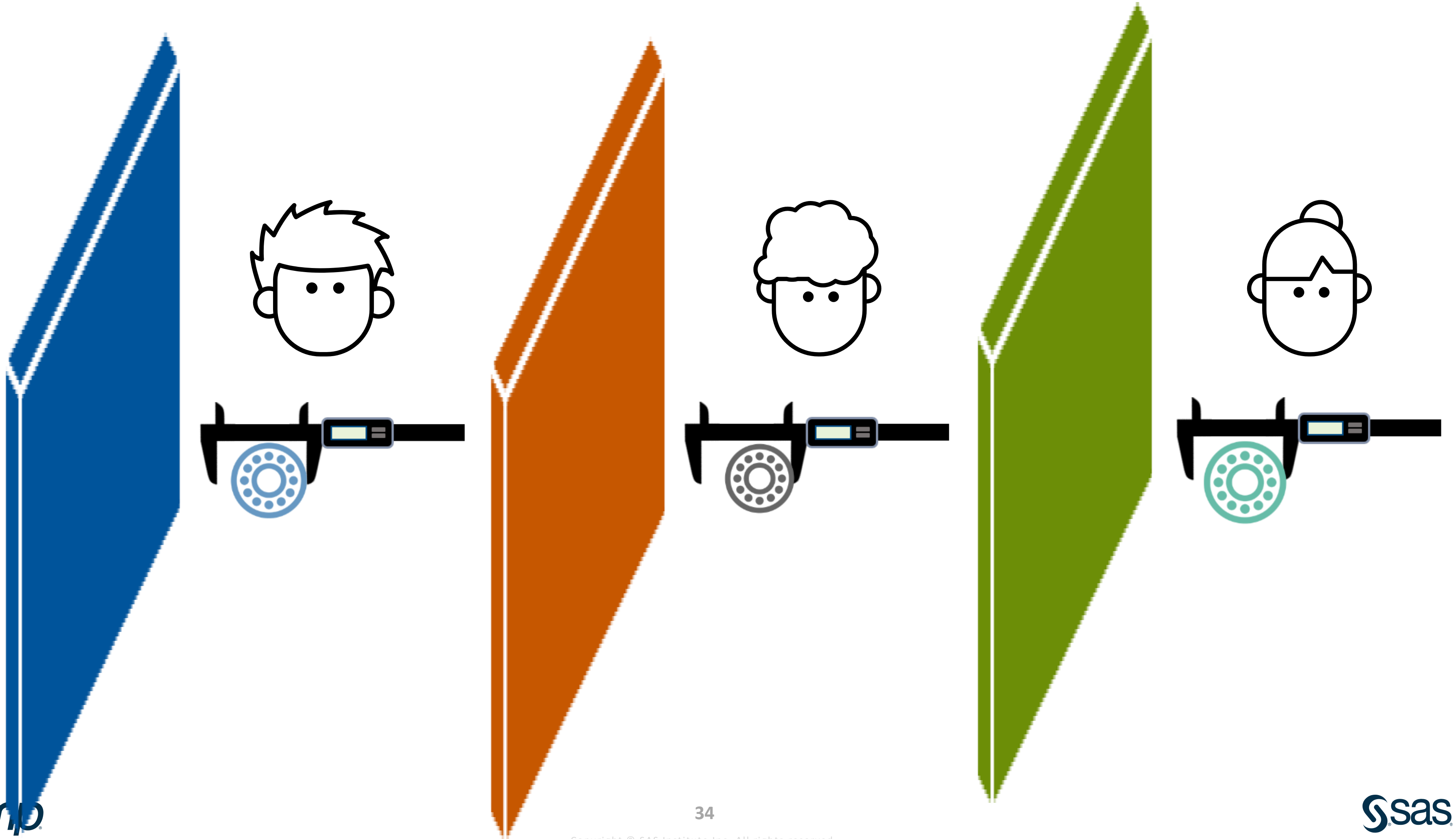


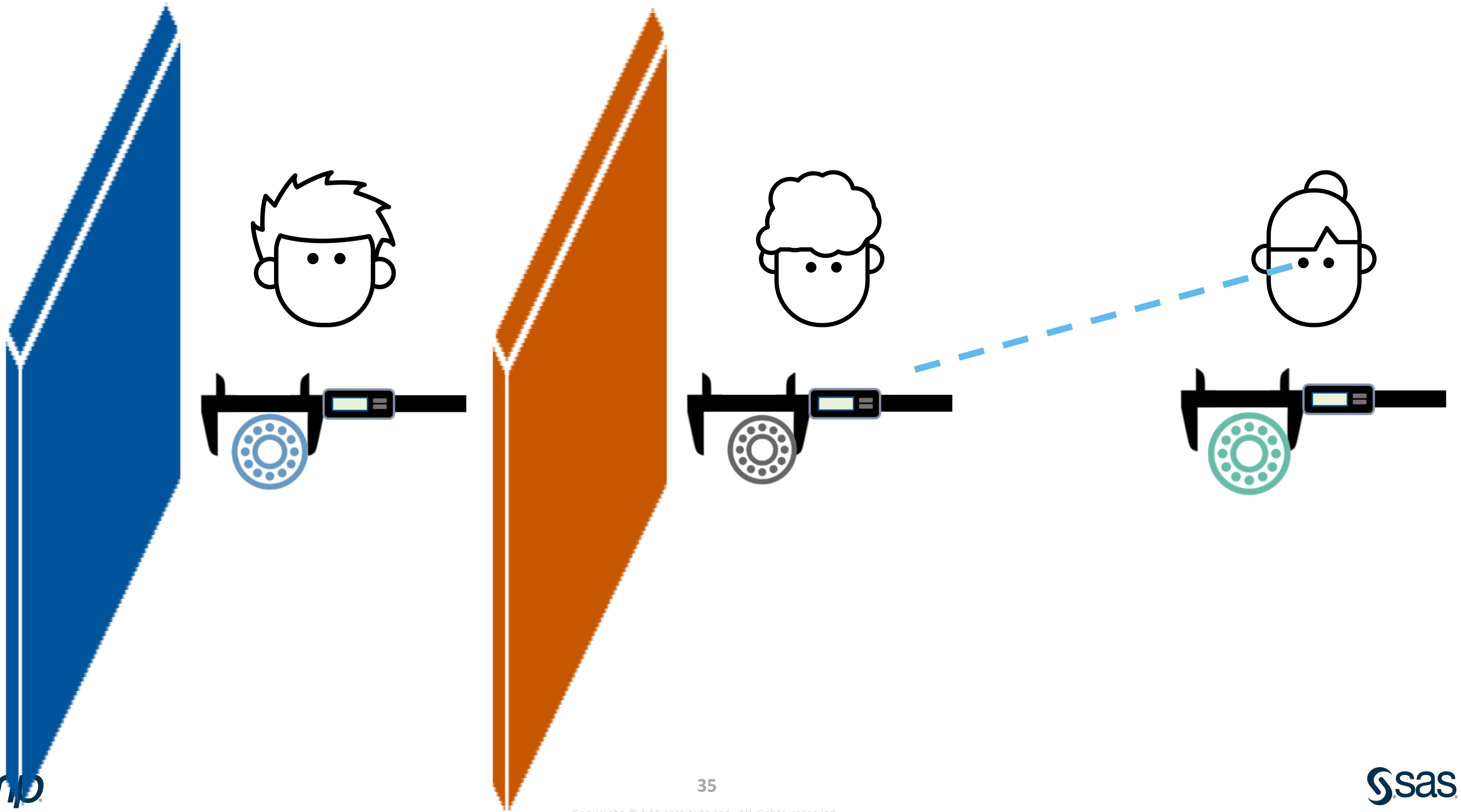
Method of Data Collection

- How you conduct the experiment is critical to the analysis and interpretation of the analysis.
- Suppose you want to have each operator measure each part twice to assess if they can get the same answer twice.
- Do they measure it twice before moving on to the next part, or do all operators measure all parts and then come back to each part a second time?
- Can the operators talk to each other about their process for taking the measurements?
- How long is the time between the first and second measurement? A few minutes, a few hours, the next day?

Data Collection

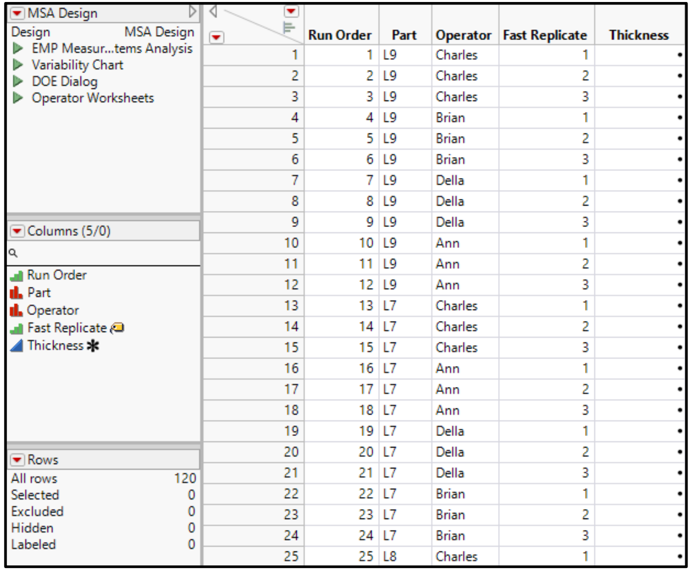
- Executing the data collection successfully requires careful planning.
- The samples need to be blindly presented to the operator for measurement.
- Someone (you) needs to record the data.
- Randomize the runs wherever possible.



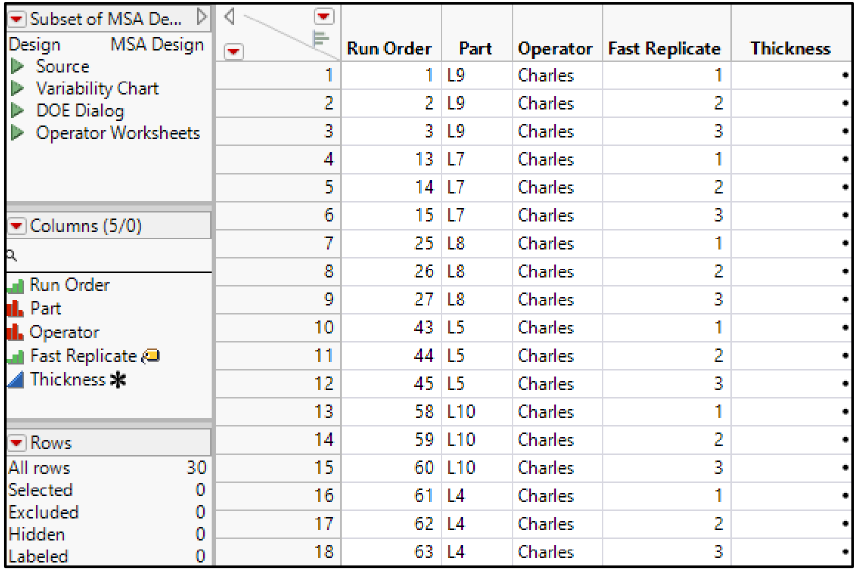


Operator Worksheets

- MSA Design allows for creation of operator worksheets.
- Make sure someone (you) fills out the worksheet, so the operator does not know which part they are measuring.



	Run Order	Part	Operator	Fast Replicate	Thickness
1	1	L9	Charles	1	•
2	2	L9	Charles	2	•
3	3	L9	Charles	3	•
4	4	L9	Brian	1	•
5	5	L9	Brian	2	•
6	6	L9	Brian	3	•
7	7	L9	Della	1	•
8	8	L9	Della	2	•
9	9	L9	Della	3	•
10	10	L9	Ann	1	•
11	11	L9	Ann	2	•
12	12	L9	Ann	3	•
13	13	L7	Charles	1	•
14	14	L7	Charles	2	•
15	15	L7	Charles	3	•
16	16	L7	Ann	1	•
17	17	L7	Ann	2	•
18	18	L7	Ann	3	•
19	19	L7	Della	1	•
20	20	L7	Della	2	•
21	21	L7	Della	3	•
22	22	L7	Brian	1	•
23	23	L7	Brian	2	•
24	24	L7	Brian	3	•
25	25	L8	Charles	1	•



	Run Order	Part	Operator	Fast Replicate	Thickness
1	1	L9	Charles	1	•
2	2	L9	Charles	2	•
3	3	L9	Charles	3	•
4	13	L7	Charles	1	•
5	14	L7	Charles	2	•
6	15	L7	Charles	3	•
7	25	L8	Charles	1	•
8	26	L8	Charles	2	•
9	27	L8	Charles	3	•
10	43	L5	Charles	1	•
11	44	L5	Charles	2	•
12	45	L5	Charles	3	•
13	58	L10	Charles	1	•
14	59	L10	Charles	2	•
15	60	L10	Charles	3	•
16	61	L4	Charles	1	•
17	62	L4	Charles	2	•
18	63	L4	Charles	3	•

Analyze Results

- Once the data have been collected, they can be analyzed using a statistical model.
- The model results will help you determine the variance of all measured sources of variability.
- Knowing the variance components will help you to conclude if the measurement process is capable or sufficient for the product that it is measuring.

Modeling Variability Depends on Design

Suppose you want to estimate random variability from *k* operators measuring *m* parts.

- How do you choose the operators? Randomly or not?
- How do you choose the parts? Randomly or not?
- Does each operator measure the same set of parts or a different set of parts?
- Does the operator measure the same part once or more than once?

The answers to these questions help determine the statistical model and which variance components can be estimated.

Random or Fixed Factors

If your choice of operators for the MSA was

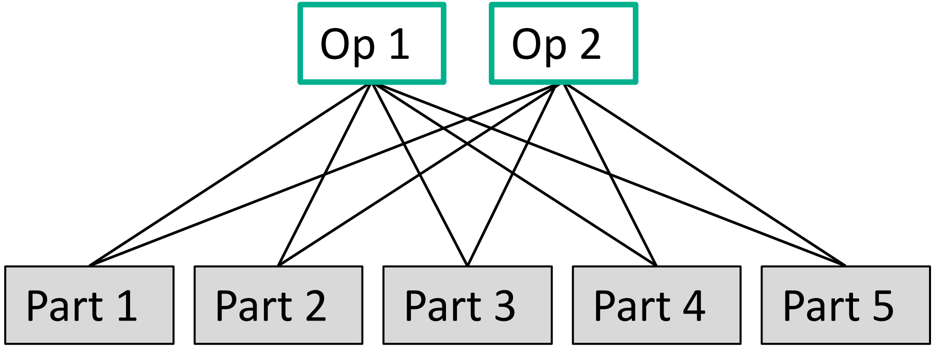
- the only ones available, or chosen to represent levels of another variable such as training level, the factor is fixed
- a random selection from the pool of operators who run the process, the factor is random.

Nested versus Crossed

- In the statistical model, two factors can be *nested* or *crossed*.
- The data table will look the same for both nested or crossed factors.
- The difference is in how the experiment is conducted.

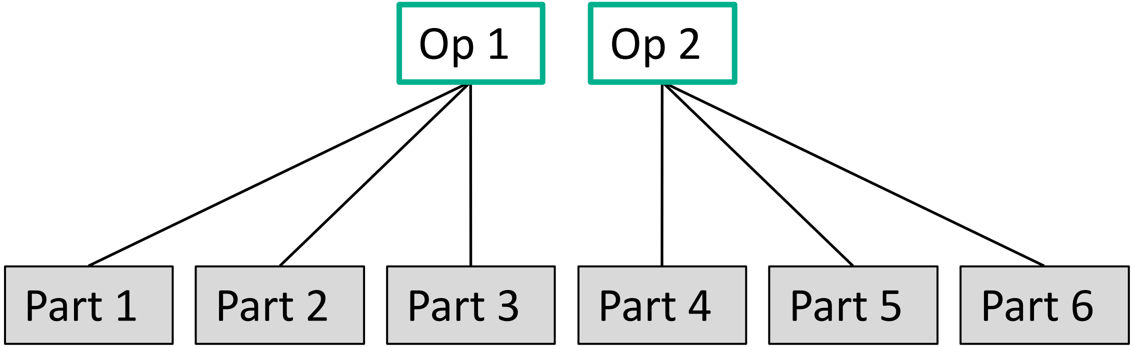
Crossed Factors

If each **Operator** measures the same set of parts, then **Part** is ***crossed*** with **Operator**.



Nested Factors

If each **Operator** measures a different set of parts, then **Part** is ***nested*** within **Operator**.



Note that **Part 4** could be called **Part 1** for **Operator 2**.
It is a different part than **Part 1** for **Operator 1**.

Make Conclusions from Analysis

- Communicate your results both statistically and graphically.
- Improve the measurement process if necessary.

Calculate Metrics

Figures of merit are calculated from the variance components.

Figure of Merit	Formula
% Gauge R&R	$100 \times 6\hat{\sigma}_{gauge} / 6\hat{\sigma}_{total}$
Precision to Part Variation	$6\hat{\sigma}_{gauge} / 6\hat{\sigma}_{part}$
Precision/Tolerance Ratio	$6\hat{\sigma}_{gauge} / (USL - LSL)$
Intraclass Correlation Coefficient	$\hat{\sigma}^2_{part} / (\hat{\sigma}^2_{part} + \hat{\sigma}^2_{gauge})$

Figures of Merit

Precision to Tolerance Ratio

How much of the allowable specification limits is taken up by measurement error?

P/T	Interpretation
< 10%	Acceptable
10% – 30%	May be acceptable
> 30%	Not acceptable

$$P / T = \frac{6\hat{s}_{gauge}}{USL - LSL}$$

Intraclass Correlation Coefficient

What proportion of total variance is Part variance?

ICC	Class
> 0.8	First class
0.5 – 0.8	Second class
0.2 – 0.5	Third class
< 0.2	Fourth class

$$ICC = \frac{s^2_{part}}{s^2_{part} + s^2_{gauge}}$$

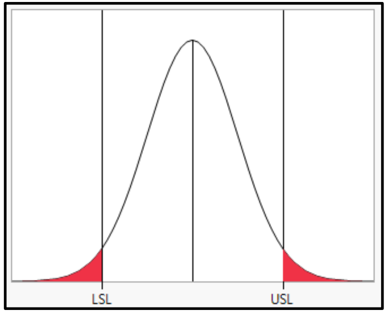
Process Capability Index

A third figure of merit measures the relationship between the specification limits and process variation: C_p.

$$C_p = \frac{USL - LSL}{6\hat{s}_{process}}$$

$$s^2_{process} = s^2_{part} + s^2_{gauge}$$

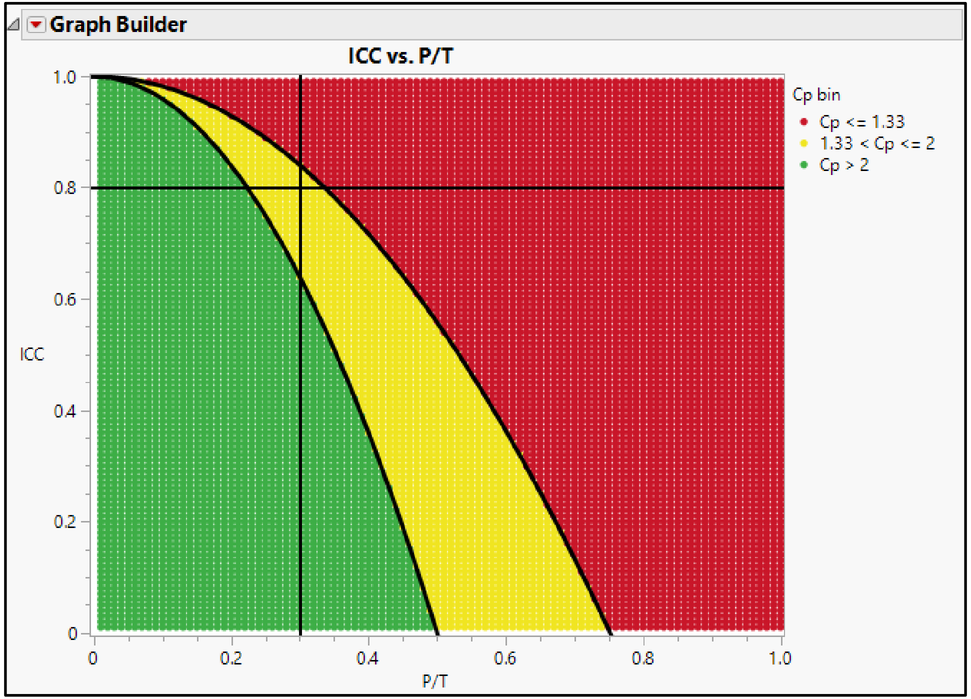
How much of the allowable specification limits is taken up by process variation (which includes measurement system variation)?



C_p	Interpretation
> 2	Excellent
1.33 – 2	Acceptable
< 1.33	Not acceptable

Relationships

$$C_p = \frac{\sqrt{1 - ICC}}{P/T}$$



Misclassification Probabilities

Because the measurement system adds variability to the process, there is a chance that

- a good part can be judged to be a bad part
- a bad part can be judged to be a good part.

Misclassification Probabilities

Description	Probability
P(Good part is falsely rejected)	0.07991
P(Bad part is falsely accepted)	0.3188
P(Part is good and is rejected)	0.07548
P(Part is bad and is accepted)	0.01767
P(Part is good)	0.94458

Three Analysis Platforms

Three main approaches to performing MSA are implemented in JMP.

- MSA platform: EMP method
- Variability platform: Gauge R&R method
- Fit Model platform: Either method with more complicated models

MSA Platform

- Features:
  - Calculations based on Wheeler (2006) *Evaluating the Measurement Process III*
  - Analysis of Means charts
  - Shift detection profiler
- Advantages:
  - This method has a complete system in one platform.
  - It has good visualization tools, easy to use and interpret.
- Disadvantages:
  - Confusion can result from using control charts in unconventional ways.
  - This approach handles only random factors.

Variability Platform

- Features:
  - Variability chart
  - Variance components analysis
- Advantages:
  - Output addresses many quality system requirements.
  - It is able to analyze attribute gauge studies.
- Disadvantages:
  - It is easy to misapply output.
  - This approach handles only random factors.

Fit Model

- Features:
  - All-purpose linear model fitting platform
- Advantages:
  - Variance components from any gauge study model can be estimated, including fixed and random factors.
- Disadvantages:
  - There is no built-in data visualization.
  - MSA summary metrics are not calculated automatically.
  - Power and complexity are unnecessary for the most common scenarios.

After the Study

After a measurement systems analysis study is completed, the following actions can be taken:

- improve the measurement process bias or variation, or both
- adjust manufacturing specification limits to ensure a high probability of meeting customer specification
- control the measurement process using control charts

Repeatability

Repeatability

Repeatability is the variability associated with repeated measurements taken under practically identical conditions.

- Same gauge
- Same setup
- Same laboratory
- Same operator
- Same part
- Shortest possible time period

You are often more interested in *test-retest error*: the variability associated with resetting the gauge and re-measuring the part.

Repeatability

Perhaps the simplest way to estimate repeatability is to measure one part on the gauge a number of times.

Model

A possible model is

$$y_i = \mu + R_i$$

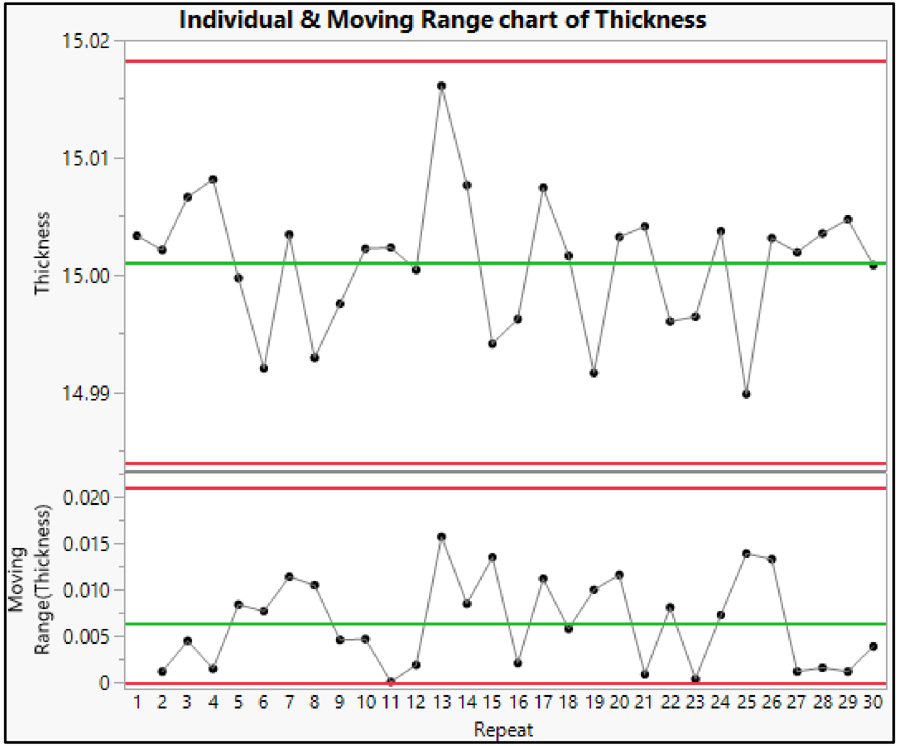
where

R_i is the effect of **Repeat** i

$$R_i \sim N(0, \sigma_R^2)$$

Graphical Analysis

An X-MR chart verifies that repeated measurements come from the same population.



Statistical Analysis

Gauge variability is estimated by the standard deviation of the repeated measurements.

EMP Results

EMP Test	Results	Description
Test-Retest Error	0.0057	Within Error
Degrees of Freedom	29	Amount of information used to estimate within error
Probable Error	0.0039	Median error for a single measurement
Intraclass Correlation	0	Proportion of variation attributed to part variation

Part Variability

- The study should use multiple parts or standards.
- This practice is to ensure applicability of the results of the study to all manufactured parts.
- The parts should be randomly chosen to span the range of possibilities, including outside of specification.
- Part variability will not be included in the variability attributed to the gauge.

Repeatability over Multiple Parts

- A more complex study is to use multiple parts and measure each part a number of times.
- From this study, you can estimate two sources of variability.
  - Part-to-part
  - Repeat-to-repeat

Model

A possible model is

$$y_{ij} = \mu + P_i + R_{j(i)}$$

where

P_i is the effect of **Part** i

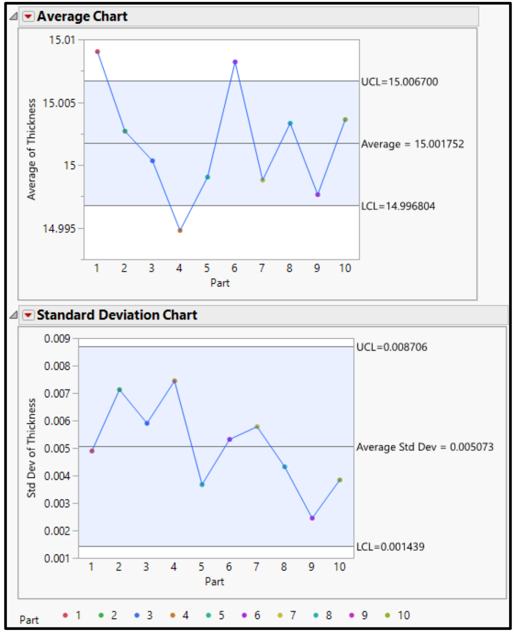
$R_{j(i)}$ is the effect of **Repeat** j for **Part** i

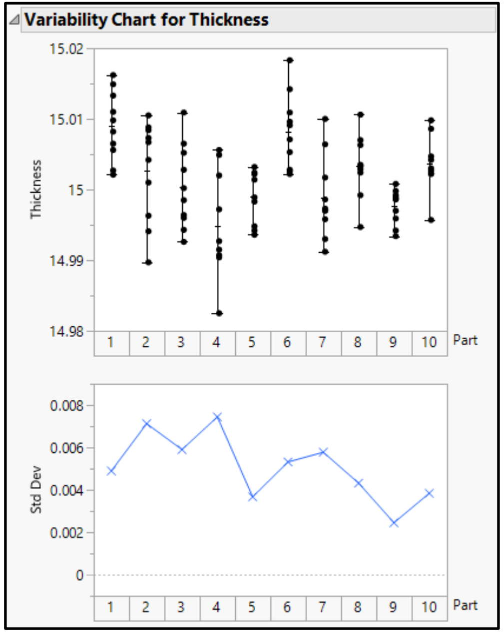
$$P_i \sim N(0, \sigma_P^2)$$

$$R_{j(i)} \sim N(0, \sigma_R^2)$$

Graphical Analysis

- Average Chart shows the mean for each part.
- Standard Deviation Chart shows the variability for each part.
- Variability Chart helps compare between-part and within-part variation.





Statistical Analysis

To estimate repeatability using multiple parts, find the repeatability for each part and combine together.

EMP Results

EMP Test	Results	Description
Test-Retest Error	0.0053	Within Error
Degrees of Freedom	90	Amount of information used to estimate within error
Probable Error	0.0036	Median error for a single measurement
Intraclass Correlation	0.3885	Proportion of variation attributed to part variation

System	Classification
Current	Third Class

Monitor Classification Legend

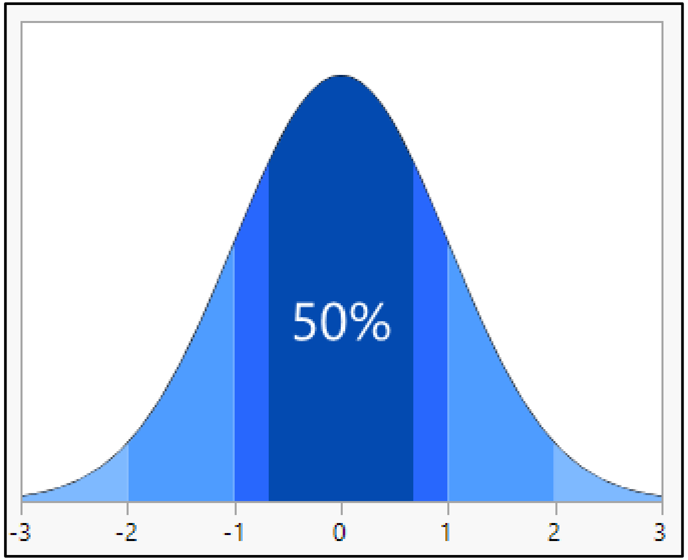
Classification	Intraclass Correlation	Attenuation of Process Signal	Probability of Warning, Test 1 Only*	Probability of Warning, Tests 1-4*
First Class	0.80 - 1.00	Less than 11%	0.99 - 1.00	1.00
Second Class	0.50 - 0.80	11% - 29%	0.88 - 0.99	1.00
Third Class	0.20 - 0.50	29% - 55%	0.40 - 0.88	0.92 - 1.00
Fourth Class	0.00 - 0.20	More than 55%	0.03 - 0.40	0.08 - 0.92

* Probability of warning for a 3 standard error shift within 10 subgroups using Wheeler's tests, which correspond to Nelson's tests 1, 2, 5, and 6.

Statistical Analysis

Probable error is the median error for one measurement.

$$Probable\ error = 0.675\hat{\sigma}_{repeat}$$



EMP Results

EMP Test	Results	Description
Test-Retest Error	0.0053	Within Error
Degrees of Freedom	90	Amount of information used to estimate within error
Probable Error	0.0036	Median error for a single measurement
Intraclass Correlation	0.3885	Proportion of variation attributed to part variation

System	Classification
Current	Third Class

Monitor Classification Legend

Classification	Intraclass Correlation	Attenuation of Process Signal	Probability of Warning, Test 1 Only*	Probability of Warning, Tests 1-4*
First Class	0.80 - 1.00	Less than 11%	0.99 - 1.00	1.00
Second Class	0.50 - 0.80	11% - 29%	0.88 - 0.99	1.00
Third Class	0.20 - 0.50	29% - 55%	0.40 - 0.88	0.92 - 1.00
Fourth Class	0.00 - 0.20	More than 55%	0.03 - 0.40	0.08 - 0.92

* Probability of warning for a 3 standard error shift within 10 subgroups using Wheeler's tests, which correspond to Nelson's tests 1, 2, 5, and 6.

Statistical Analysis

Intraclass correlation coefficient (ICC) is used to measure the quality of the measurement system. Higher is better.

$$\hat{ICC} = \frac{\hat{\sigma}^2_{part}}{\hat{\sigma}^2_{part} + \hat{\sigma}^2_{gauge}}$$

EMP Results

EMP Test	Results	Description
Test-Retest Error	0.0053	Within Error
Degrees of Freedom	90	Amount of information used to estimate within error
Probable Error	0.0036	Median error for a single measurement
Intraclass Correlation	0.3885	Proportion of variation attributed to part variation

System	Classification
Current	Third Class

Monitor Classification Legend

Classification	Intraclass Correlation	Attenuation of Process Signal	Probability of Warning, Test 1 Only*	Probability of Warning, Tests 1-4*
First Class	0.80 - 1.00	Less than 11%	0.99 - 1.00	1.00
Second Class	0.50 - 0.80	11% - 29%	0.88 - 0.99	1.00
Third Class	0.20 - 0.50	29% - 55%	0.40 - 0.88	0.92 - 1.00
Fourth Class	0.00 - 0.20	More than 55%	0.03 - 0.40	0.08 - 0.92

* Probability of warning for a 3 standard error shift within 10 subgroups using Wheeler's tests, which correspond to Nelson's tests 1, 2, 5, and 6.

Statistical Analysis

Effective Resolution measures the discrimination of the measurement system. Possible actions are Drop a digit, Consider dropping a digit, Use as is, and Consider adding a digit, based on the probable error and measurement increment.

Effective Resolution

Source		Value	Description
Probable Error	(PE)	0.0036	Median error for a single measurement
Current Measurement Increment	(MI)	0.0001	Measurement increment estimated from data (in tenths)
Lower Bound Increment	(0.1*PE)	0.0004	Measurement increment should not be below this value
Smallest Effective Increment	(0.22*PE)	0.0008	Measurement increment is more effective above this value
Largest Effective Increment	(2.2*PE)	0.0078	Measurement increment is more effective below this value

Action: Drop a digit

Reason: The measurement increment of 0.0001 is below the lowest measurement increment bound and should be adjusted to record fewer digits.

Statistical Analysis

Shift Detection Profiler

Shift Detection Profiler

In-Control Part Std Dev	0.0042126
In-Control Chart Sigma	0.0067585
False Alarm Probability	0.0266723

Reproducibility

- Reproducibility is the variability associated with the measurement system under different, but typical, conditions.
  - Different time periods
  - Different operators
  - Different gauges
  - Different setup procedures
  - Different environmental conditions

Operators

- Run the same study using multiple operators.
- Two possibilities exist.
  - Operators are chosen randomly from a pool of operators.
  - Operators are either the only ones available or are chosen to represent levels of another variable, such as training level.

Model for Random Operators

A possible main effects model is

$$y_{ijk} = \mu + P_i + O_j + R_{k(ij)}$$

where

P_i is the effect of **Part** i

O_j is the effect of **Operator** j

$R_{k(ij)}$ is the effect of **Repeat** k for each **Part** i and **Operator** j

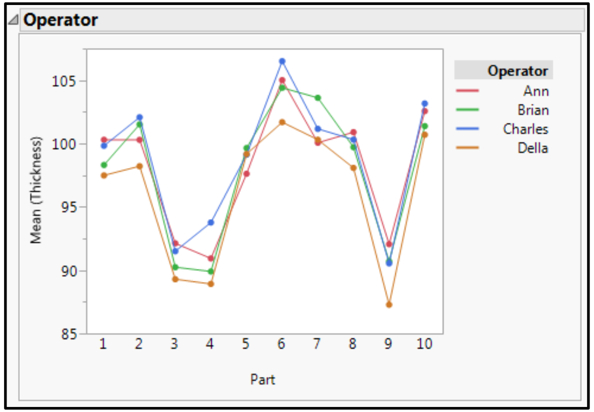
$$P_i \sim N(0, \sigma_P^2)$$

$$O_j \sim N(0, \sigma_O^2)$$

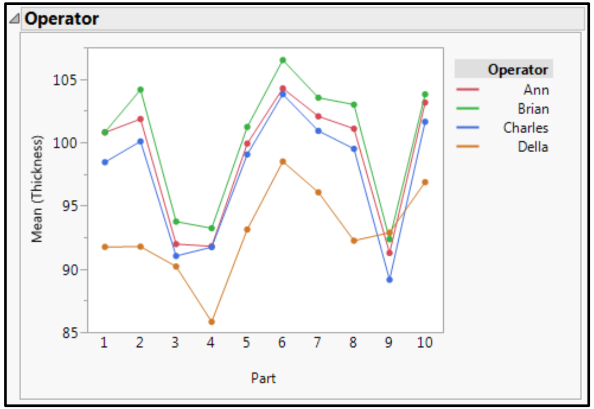
$$R_{k(ij)} \sim N(0, \sigma_R^2)$$

Crossed Factors

- If each **Operator** measures each **Part**, then the model supports estimation of the **Operator*Part** interaction.
- A two-factor interaction exists if the effect of one factor depends on the level of the other factor. Interactions appear as non-parallel lines on the parallelism plot.



no interaction



Operator*Part interaction

Model for Random Operators Crossed with Parts

A possible crossed model is

$$y_{ijk} = \mu + P_i + O_j + PO_{ij} + R_{k(ij)}$$

where

P_i is the effect of **Part** *i*

O_j is the effect of **Operator** *j*

PO_{ij} is the interaction effect of **Part** *i* and **Operator** *j*

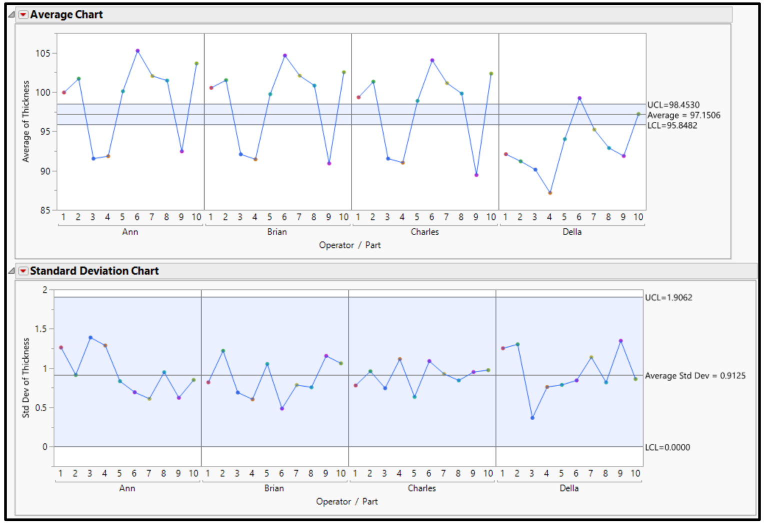
$R_{k(ij)}$ is the effect of **Repeat** *k*

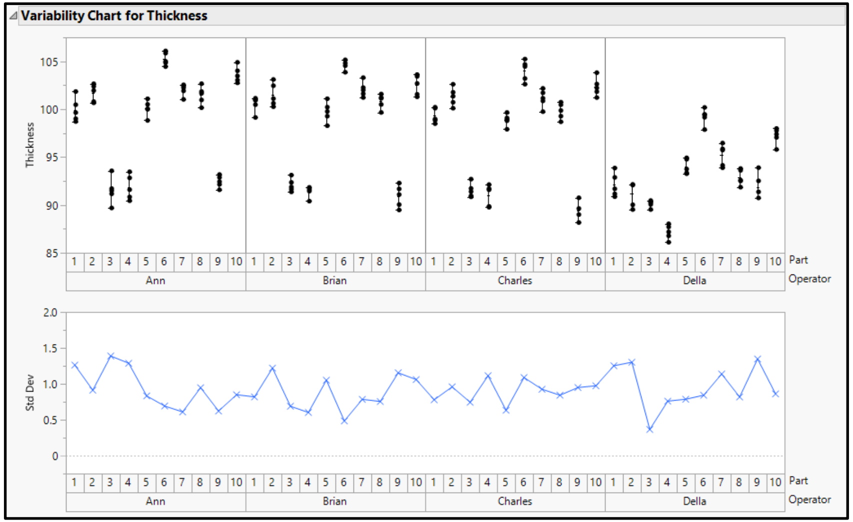
$$P_i \sim N(0, \sigma_P^2) \quad PO_{ij} \sim N(0, \sigma_{PO}^2)$$

$$O_j \sim N(0, \sigma_O^2) \quad R_{k(ij)} \sim N(0, \sigma_R^2)$$

Graphical Analysis

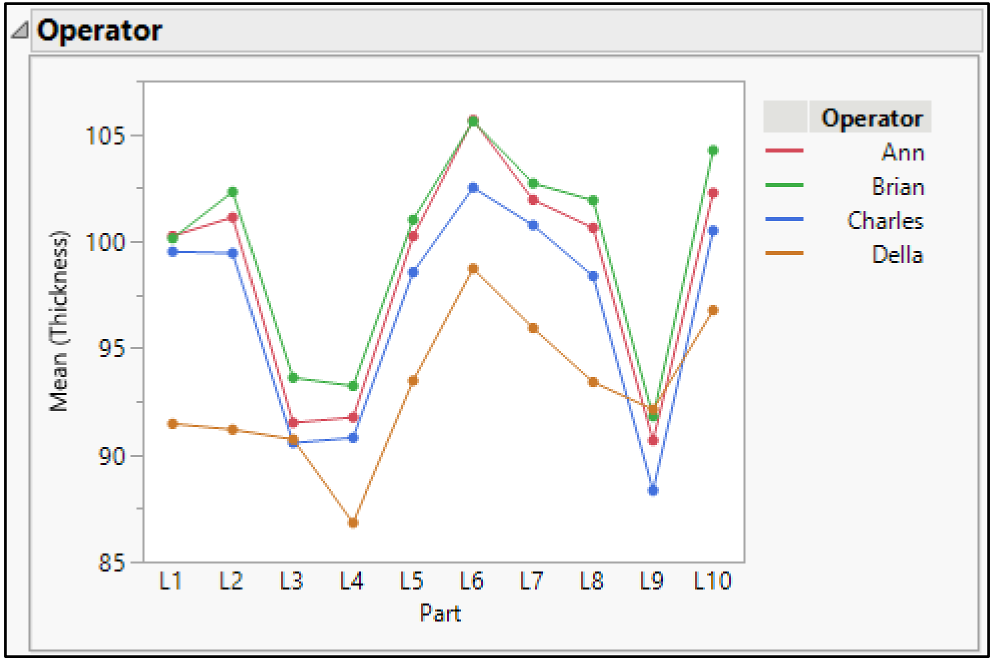
- Average and Standard Deviation Charts show the mean and variability of each part for each operator.
- Variability Chart shows how much variation is due to **Operator**, **Part**, and **Repeat**.





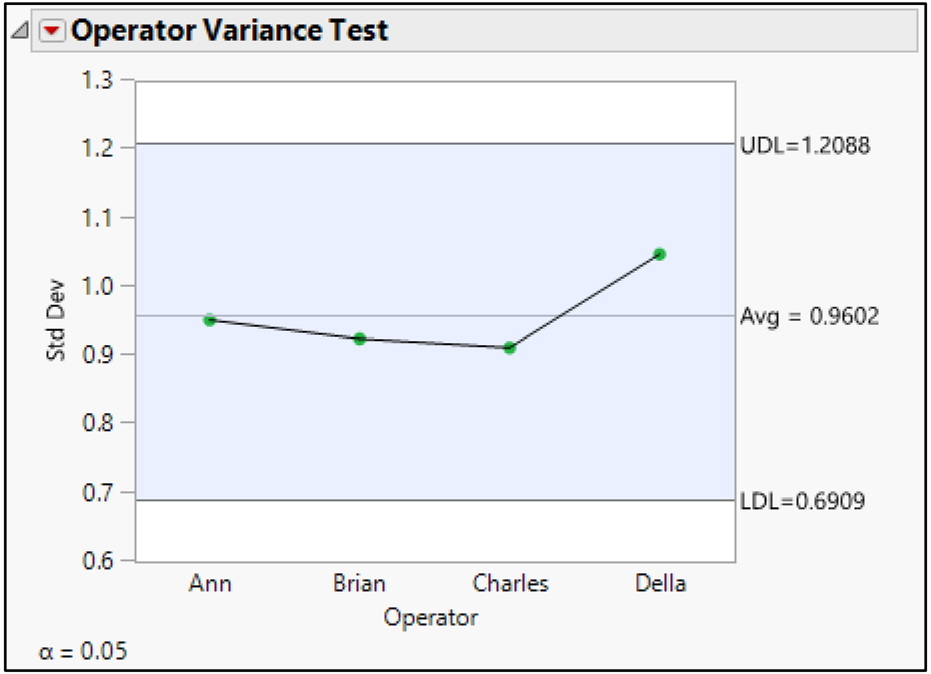
Graphical Analysis

A parallelism plot shows whether there is an **Operator*Part** interaction.



Graphical Analysis

Test-retest error comparison plot shows the variability for each operator.



Statistical Analysis

- The estimate of repeatability is found by combining the estimates of repeatability for each operator and part.
- The report also displays the estimate of probable error and intraclass correlation coefficient.

EMP Results

EMP Test	Results	Description
Test-Retest Error	0.9602	Within Error
Degrees of Freedom	160	Amount of information used to estimate within error
Probable Error	0.6476	Median error for a single measurement
Intraclass Correlation (no bias)	0.9543	Proportion of variation attributed to part variation without including bias factors
Intraclass Correlation (with bias)	0.6815	Proportion of variation attributed to part variation with bias factors
Intraclass Correlation (with bias and interactions)	0.6179	Proportion of variation attributed to part variation with bias factors and interactions
Bias Impact	0.2728	Amount by which the bias factors reduce the intraclass correlation
Bias and Interaction Impact	0.3364	Amount by which the bias factors and interactions reduce the intraclass correlation

System	Classification
Current (with bias)	Second Class
Current (with bias and interactions)	Second Class
Potential (no bias)	First Class

Model for Parts Nested within Random Operators

A possible nested model is

$$y_{ijk} = \mu + O_j + P_{i(j)} + R_{k(ij)}$$

where

O_j is the effect of **Operator** j

$P_{i(j)}$ is the effect of **Part** i nested within **Operator** j

$R_{k(ij)}$ is the effect of **Repeat** k

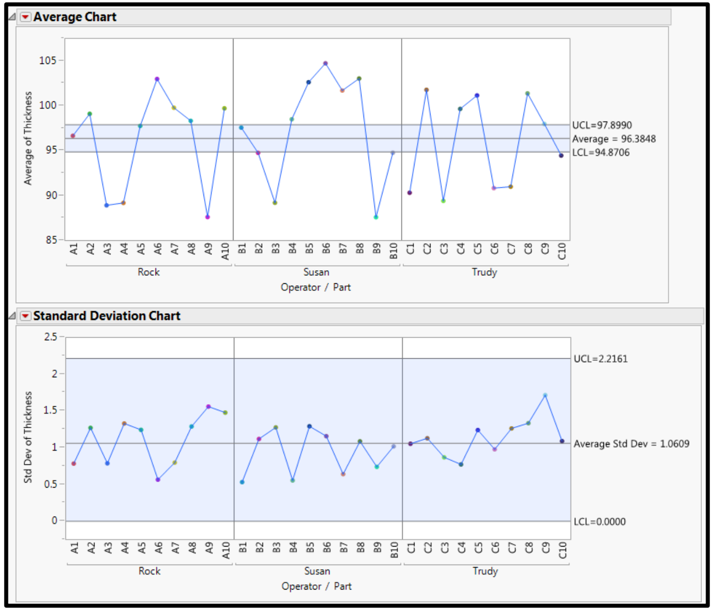
$$O_j \sim N(0, \sigma_O^2)$$

$$P_{i(j)} \sim N(0, \sigma_P^2)$$

$$R_{k(ij)} \sim N(0, \sigma_R^2)$$

Graphical Analysis

- Average Chart shows the variability of parts for each operator.
- Standard Deviation Chart shows the variability of each operator.



Statistical Analysis

- The estimate of repeatability is found by combining together the estimates of repeatability for each operator and part.
- The report also displays the estimate of probable error and intraclass correlation coefficient.

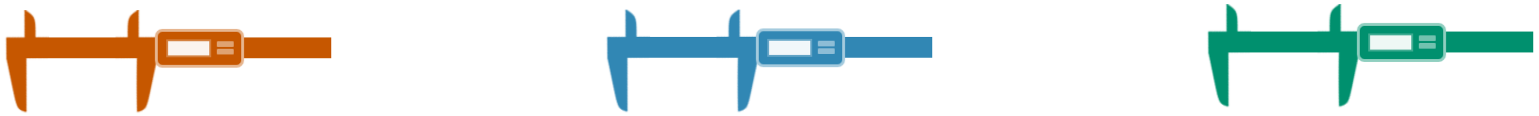
EMP Results

EMP Test	Results	Description
Test-Retest Error	1.1029	Within Error
Degrees of Freedom	120	Amount of information used to estimate within error
Probable Error	0.7439	Median error for a single measurement
Intraclass Correlation (no bias)	0.9586	Proportion of variation attributed to part variation without including bias factors
Intraclass Correlation (with bias)	0.9586	Proportion of variation attributed to part variation with bias factors
Bias Impact	0	Amount by which the bias factors reduce the intraclass correlation

System	Classification
Current (with bias)	First Class
Potential (no bias)	First Class

Reproducibility with Fixed Effects

- Suppose you want to compare the measurement variability of three gauges.
- These are the only gauges you have, and it does not make sense to think about them as a random sample of all gauges.



- You have four operators who are selected at random from a pool of operators.
- The operators will measure all parts on all gauges twice each day for five days.

Statistical Model

A possible model is

$$y_{ijklm} = m + P_i + G_j + O_k + D_l + interactions + R_{m(ijkl)}$$

where

P_i is the effect of **Part** i

G_j is the effect of **Gauge** j

O_k is the effect of **Operator** k

D_l is the effect of **Day** l

$R_{m(ijkl)}$ is the effect of **Repeat** l.

$$P_i \sim N(0, s_P^2)$$

$$\overset{3}{\underset{j=1}{\text{å}}} G_j = 0$$

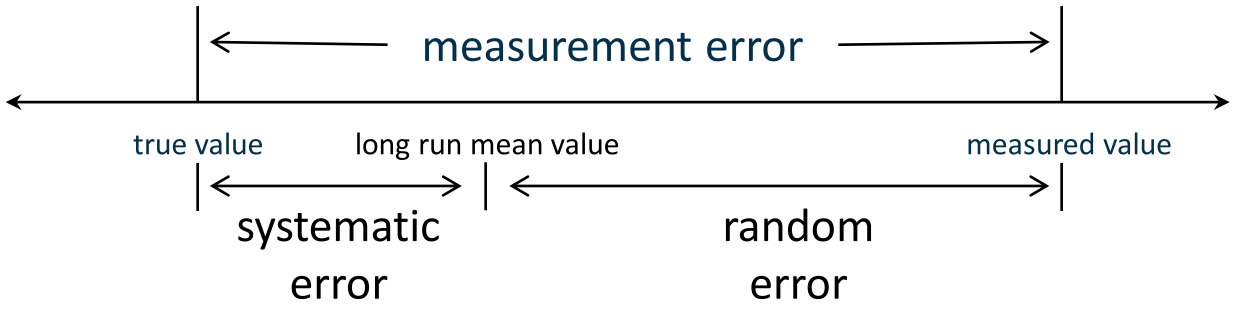
$$O_k \sim N(0, s_O^2)$$

$$D_l \sim N(0, s_D^2)$$

All interactions effects are random.

Measurement Error

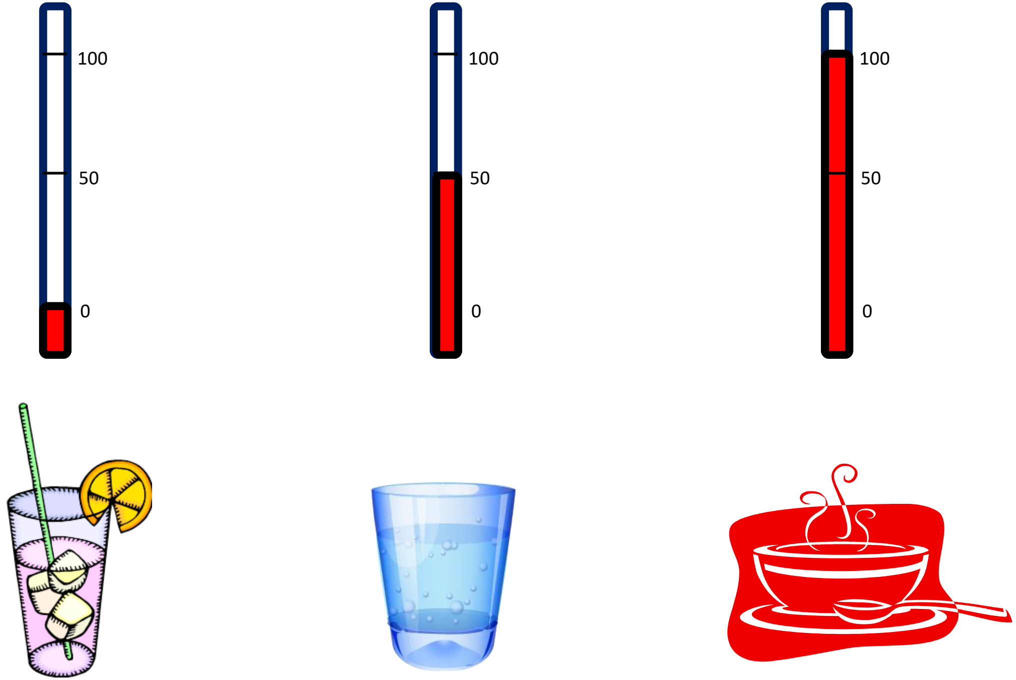
- Measurement error can be split into two parts.
  - *Systematic error* is the difference between the long run mean value of the part and the true value.
  - *Random error* is the difference between the measured value of the part and the long run mean value.



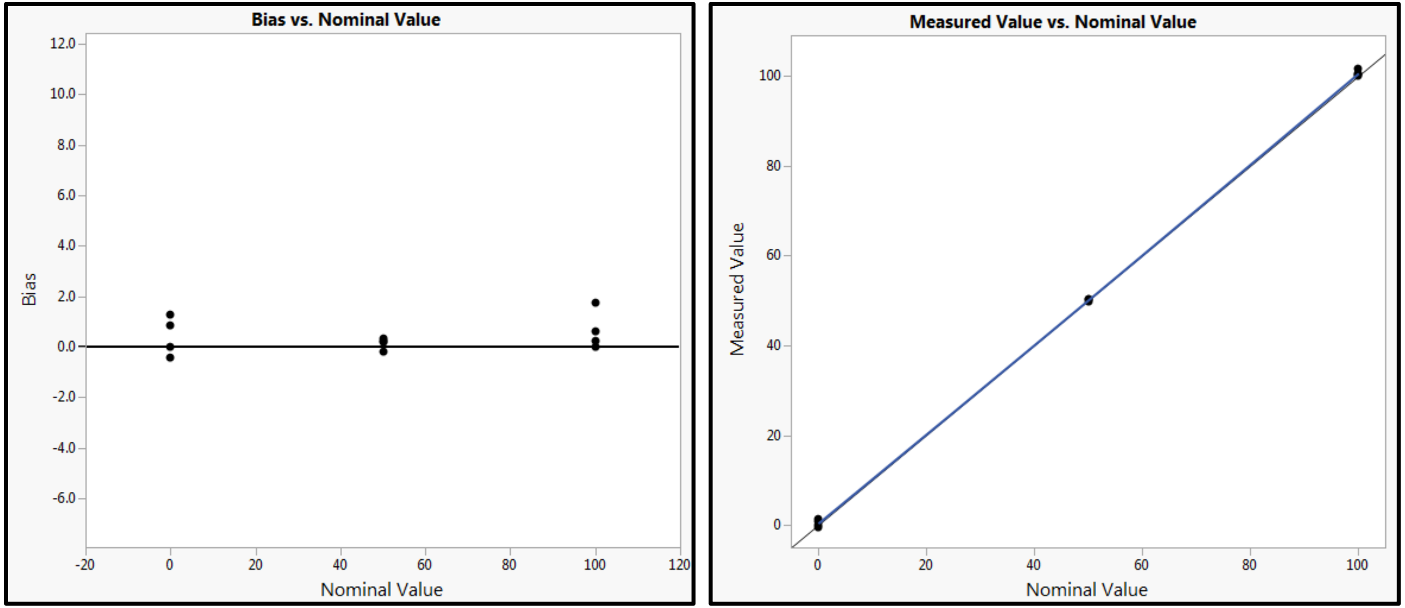
Systematic Error

- Studying systematic error provides information about the ***bias*** of the measurement system.
- Bias measures how far the system is from truth, on the average.
- Bias can be zero, constant, linear, or nonlinear.
- Bias studies need ***standard reference material*** for which the true value can be assumed known.

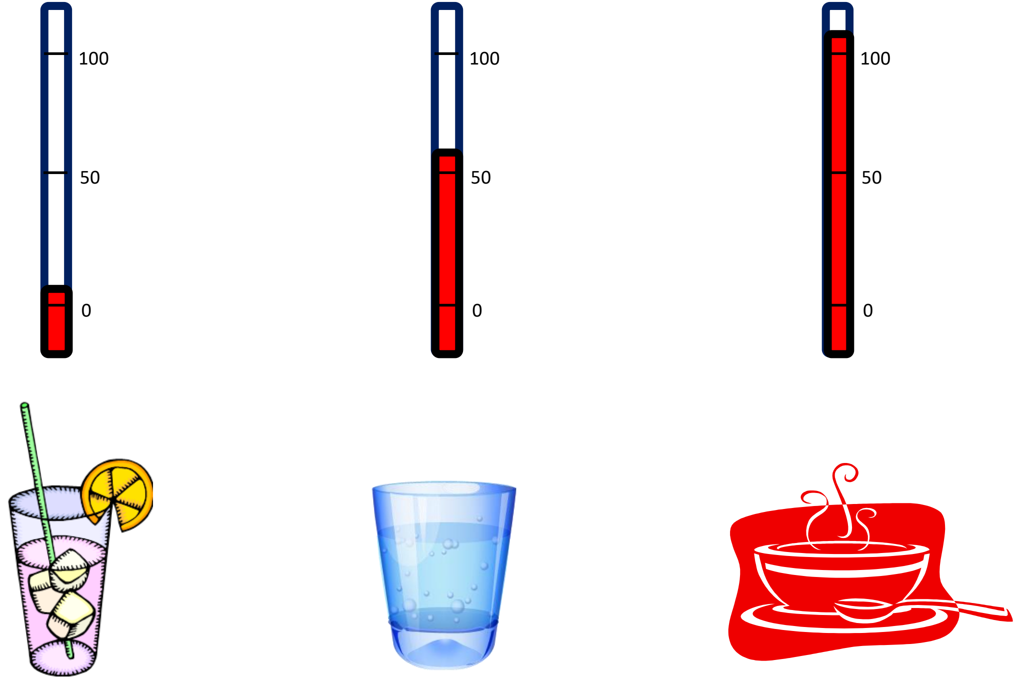
Zero Bias



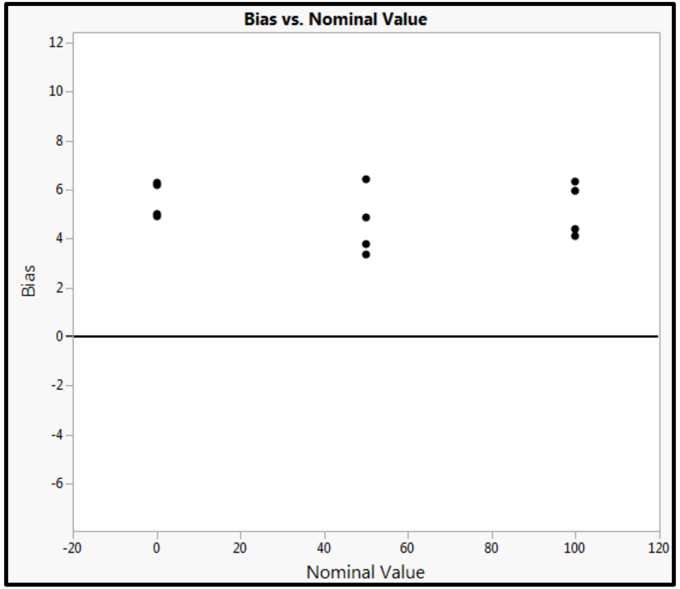
Zero Bias

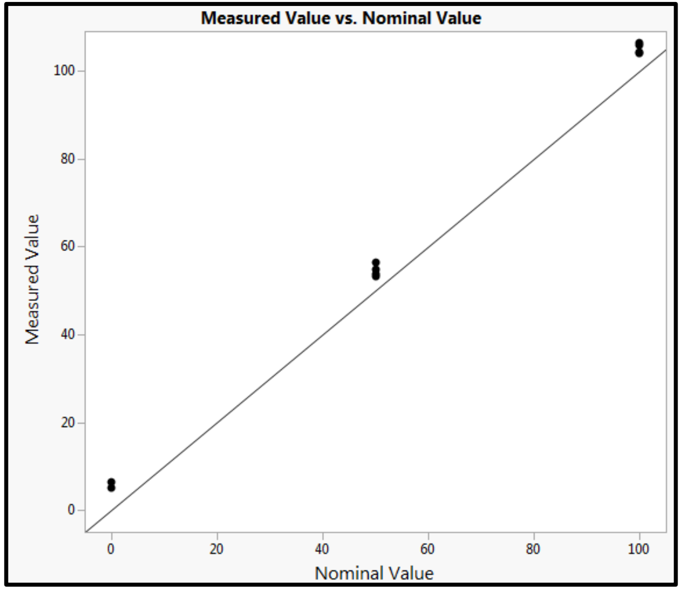

Bias vs. Nominal Value
Bias
12.0
10.0
8.0
6.0
4.0
2.0
0.0
-2.0
-4.0
-6.0
-20
0
20
40
60
80
100
120
Nominal Value
Measured Value vs. Nominal Value
Measured Value
100
80
60
40
20
0
0
20
40
60
80
100
Nominal Value

Constant Bias

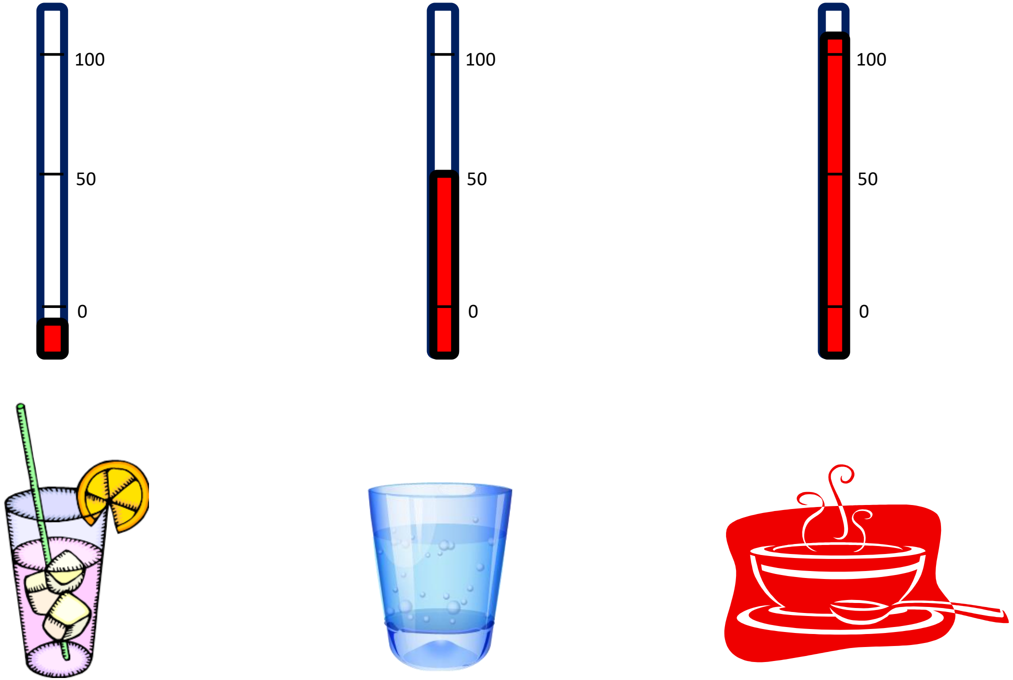


Constant Bias

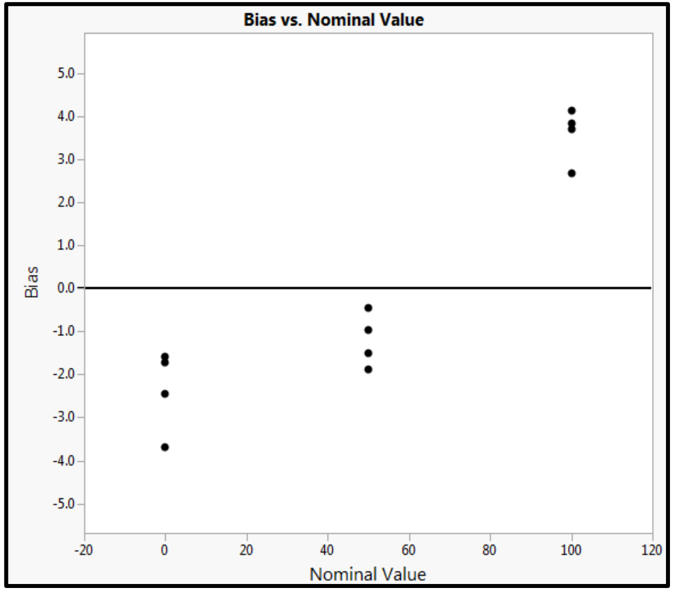


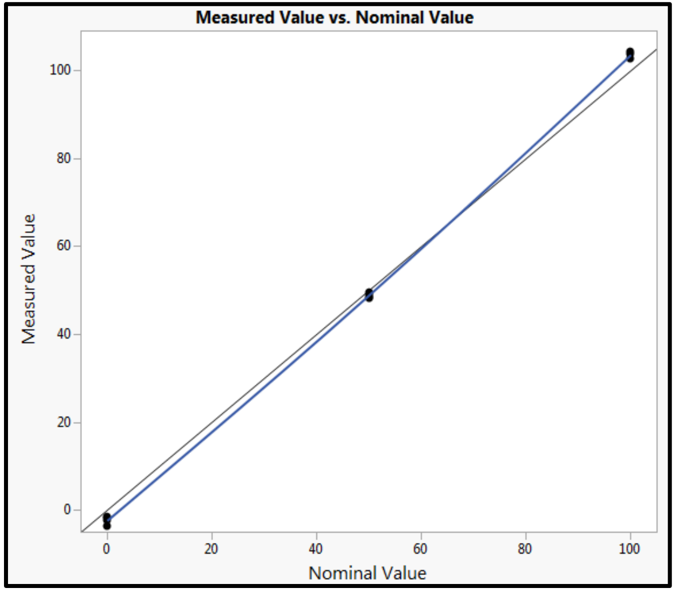


Linear Bias

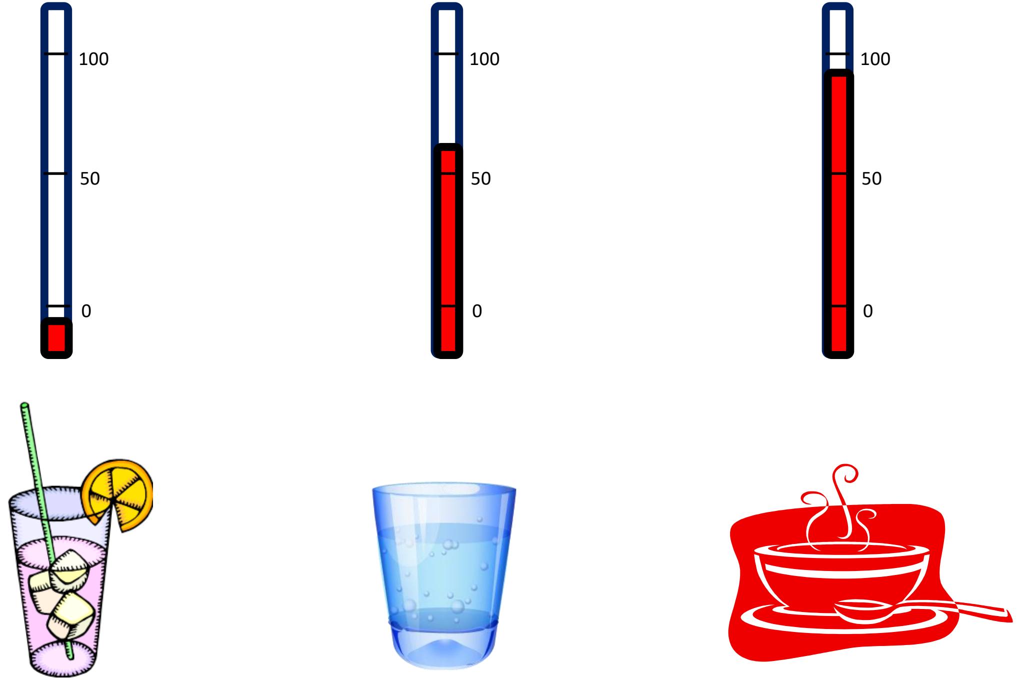


Linear Bias

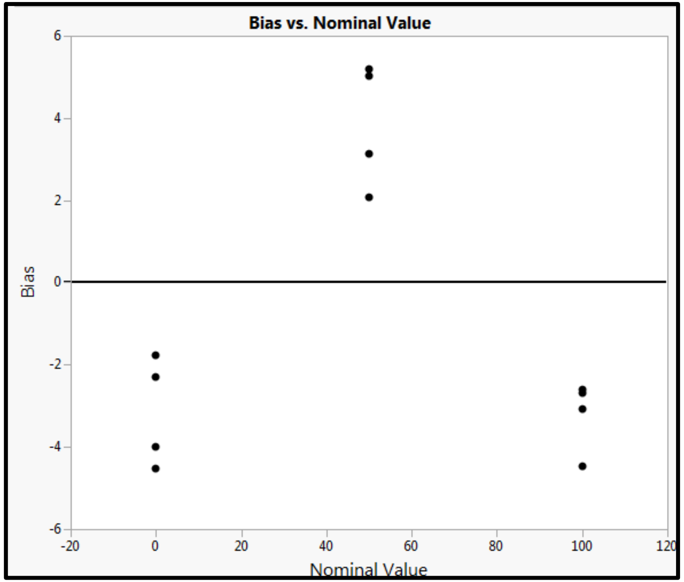

Bias vs. Nominal Value
Bias
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
-5.0
-20
0
20
40
60
80
100
120
Nominal Value

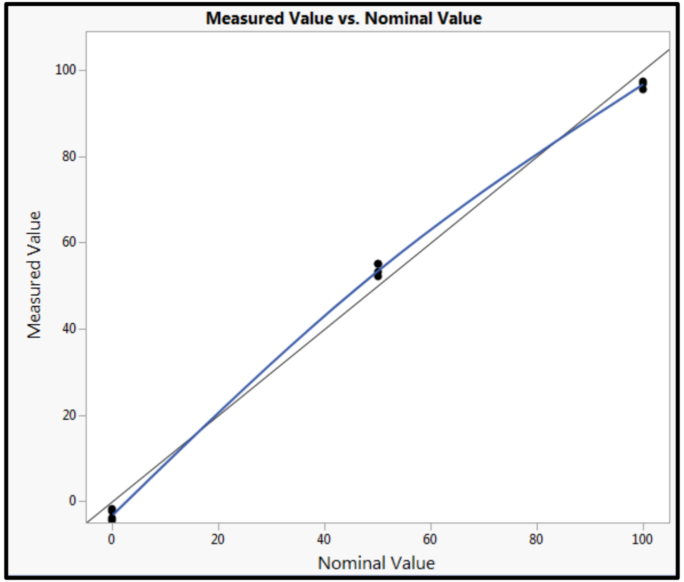

Measured Value vs. Nominal Value
Measured Value
100
80
60
40
20
0
0
20
40
60
80
100
Nominal Value

Nonlinear Bias



Nonlinear Bias



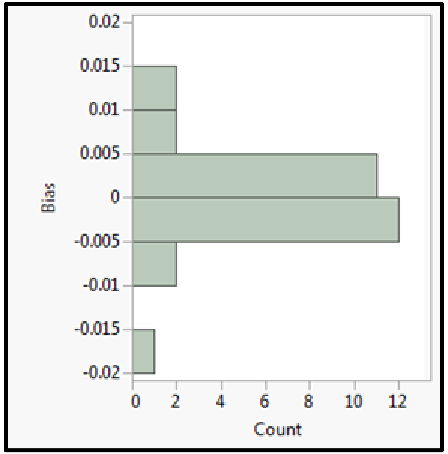


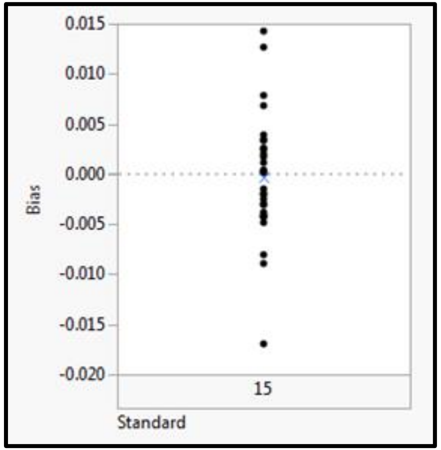
Simple Bias Study

- Obtain a standard with known or assumed value.
- Measure the standard 30 times on the gauge.
- Do not remove the standard from the gauge between readings.
- The estimate of bias is the sample mean minus the known value of the standard.

Graphical Analysis

Histogram and dot plot of measurements show bias and repeatability.





Statistical Analysis

Confidence interval and *t* test on the mean test if the bias is zero.

Average Bias	-0.00028
Std Error	0.00112
t Ratio	-0.25286
Prob> |t|	0.80216
Lower 95%	-0.00258
Upper 95%	0.00201

Linearity

Linearity is the absence of changes in bias over a range of measurement values.

Methodology

- Obtain standards
- Set up the gauge
- Calibrate, if necessary
- Determine sample size
- Collect data
- Analyze data

Obtain Standards

- Choose at least three nominal values in the range of measurements.
- Obtain at least one standard at each of the nominal values.
  - Standards should be either certified reference material or have known or assumed values.
  - Standards should be similar to usual parts or values should be transferable to usual parts.

Set Up Gauge

- Implement the appropriate gauge setup.
- The gauge should be set up as in normal production.

Calibration

- Determine whether the gauge needs to be calibrated.
- Calibrate, if needed, following the manufacturer’s instructions.
- Do not recalibrate during the study, unless calibration is required during normal operating procedure.

Sample Size

To determine the required minimum sample size to estimate bias, the following quantities need to be specified:

- σ, estimate of reproducibility over the time period of the study
- δ, magnitude of bias to be detected
- α, largest acceptable probability of calibrating when unnecessary
- β, largest acceptable probability of not calibrating when necessary

Stability

- Stability is the absence of changes in the bias due to taking data over time.
- If stability information is desired, perform the study over multiple days.
    - make at least two measurements per day
    - over at least three days