



## Chapter 12

# Measurement Systems Analysis

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## Introduction

- Data integrity assessments are considered a part of Six Sigma measurement systems analysis (MSA) studies.
  - First consider whether we are measuring the right thing.
  - If wrong number is recorded into a database.
- Assessment of any measuring devices.
  - Manufacturing uses many forms of measuring systems when making decisions. However, organizations sometimes do not even consider that their measurement might not be exact.
  - Product/process may appear unsatisfactory because of poor measurement system.

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## Introduction

- Traditionally, the tool to address the appraiser/operator consistency is a gage **repeatability** and **reproducibility** (R&R) study, which is the evaluation of measuring instruments to determine capability to yield a precise response.
- Gage repeatability is the variation in measurements considering one part and one operator.
- Gage reproducibility is the variation between operators measuring one part.
- Nondestructive <> Destructive (nonreplicable) testing



## 12.1 MSA Philosophy

- Moved from focusing on compliance to system understanding and improvement.
- Measurement is a lifelong process, not a single snapshot.
- MSA should cover not only the appraiser/operator and machine in a gage R&R study, but other factors such as temperature, humidity, dirt, training, and other conditions.
- The initial purchase of measurement systems should be addressed as part of an overall Advanced Product Quality Planning (APQP) system.



## 12.2 Variability Sources in a 30,000-ft-level Metric

$$\sigma_T^2 = \sigma_p^2 + \sigma_m^2$$

Total Variance = Process Variance + Measurement Variance

- MSA involves the understanding and quantification of measurement variance.
- Accuracy is the degree of agreement of individual or average measurements with an accepted reference value or level.
- Precision is the degree of mutual agreement among individual measurements made under prescribed like conditions (ASTM 1977).



## 12.2 Variability Sources in a 30,000-ft-level Metric

- MSA assesses the statistical properties of repeatability, reproducibility, bias, stability, and linearity.
- Gage R&R studies address the variability of the measurement system, while bias, stability, and linearity studies address the accuracy of the measurement system.



## 12.3 S<sup>4</sup>/IEE Application Examples: MSA

- Satellite-level metric: Focus was to be given to creating S<sup>4</sup>/IEE projects that improved a company's ROI. As part of a MSA assessment the team decided effort was to be given initially to how the satellite-level metric was calculated. It was thought that there might be some month-to-month inconsistencies in how this metric was being calculated and reported.
- Satellite-level metric: S<sup>4</sup>/IEE projects were to be created that improve the company's customer satisfaction. Focus was given to ensure that the process for measuring customer satisfaction gave an accurate response.



## 12.3 S<sup>4</sup>/IEE Application Examples: MSA

- Transactional 30,000-foot-level metric: DSO reduction was chosen as S<sup>4</sup>/IEE project. Focus was given to ensuring that DSO entries accurately represented what happened within the process.
- Manufacturing 30,000-foot-level metric (KPOV): An S<sup>4</sup>/IEE project was to improve the capability/performance of the diameter for a manufactured product (i.e., reduce the number of parts beyond the specification limits). An MSA was conducted of the measurement gage.



## 12.3 S<sup>4</sup>/IEE Application Examples: MSA

- Transactional and manufacturing 30,000-foot-level cycle time metric (a lean metric): An S<sup>4</sup>/IEE project was to improve the time from order entry to fulfillment was measured. Focus was given to ensure that the cycle time entries accurately represented what happened within the process.
- Transactional and manufacturing 30,000-foot-level inventory metric or satellite-level TOC metric (a lean metric): An S<sup>4</sup>/IEE project was to reduce inventory. Focus was given to ensure that entries accurately represented what happened within the process.



## 12.3 S<sup>4</sup>/IEE Application Examples: MSA

- Manufacturing 30,000-foot-level quality metric: An S<sup>4</sup>/IEE project was to reduce the number of defects in a printed circuit board manufacturing process. An MSA was conducted to determine if defects were both identified and recorded correctly into the company's database.
- Transactional 50-foot-level metric (KPIV): An S<sup>4</sup>/IEE project to improve the 30,000-foot-level metrics for DSOs identified as KPIV to the process. An MSA was conducted to determine the metric is reported accurately.



## 12.3 S<sup>4</sup>/IEE Application Examples: MSA

- Product DFSS: An S<sup>4</sup>/IEE product DFSS project was to reduce the 30,000-foot-level MTBF (mean time between failures) of a product by its vintage (e. g., laptop computer MTBF rate by vintage of the computer). As part of an MSA the development test process was assessed. It was discovered that much of the test process activities was not aligned with the types of problems typically experienced by customers.



## 12.4 Terminology

- Accuracy is the closeness of agreement between an observed value and the accepted reference value.
- Precision is the net effect of discrimination, sensitivity, and repeatability over the operating range (size, range, and time) of the measurement system.
- Part variation (PV), as related to measurement systems analysis. Represents the expected part-to-part and time-to-time variation for a stable process.
- Measurement system error is the combined variation due to gage bias, repeatability, reproducibility, stability, and linearity.



## 12.4 Terminology

- Bias is the difference between the observed average of measurements (trials under repeatability conditions) and a reference value; historically referred to as accuracy. Bias is evaluated and expressed at a single point with the operating range of the measurement system.
- Repeatability is the variability resulting from successive trials under defined conditions of measurement. It is often referred to as equipment variation (EV), which can be a misleading term. The best term for repeatability is within-system variation, when the conditions of measurement are fixed and defined (i.e., fixed part, instrument, standard, method, operator, environment, and assumptions).



## 12.4 Terminology

- Reproducibility is the variation in the average of measurements caused by a normal condition(s) of change in the measurement process. Typically, it has been defined as the variation in average measurements of the same part (measurand) between different appraisers (operators) using the same measurement instrument and method in a stable environment. This is often true for manual instruments influenced by the skill of the operator. It is not true, however, for measurement processes (i.e., automated systems) where the operator is not a major source of variation. For this reason, reproducibility is referred to as the average variation between-systems or between-conditions of measurement.



## 12.4 Terminology

- Appraiser variation (AV) is the average measurements of the same part between different appraisers using the same measuring instrument and method in a stable environment. AV is one of the common sources of measurement system variation that results from difference in operator skill or technique using the same measurement system.
- Stability refers to both statistical stability of measurement process and measurement stability over time. Both are vital for a measurement system to be adequate for its intended purpose. Statistical stability implies a predictable, underlying measurement process operating within common cause variation. Measurement drift addresses the necessary conformance to the measurement standard or reference over the operating life (time) of the measurement system.



## 12.4 Terminology

Minitab definition:

- Repeatability is the variation due to the measuring device. It is the variation observed when the same operator measures the same part repeatedly with the same device.
- Reproducibility is the variation due to the measurement system. It is the variation observed when different operators measure the same parts using the same device.





## 12.5 Gage R&R Considerations

- Measurement must be in statistical control (statistical stability).
- Variability of the measurement system must be small compared with both the manufacturing process and specification limits.
- Increment of measurement must be small relative to both process variability and specification limits. (A common rule of thumb is that the increments should be no greater than 1/10 of the smaller of the process variability and specification limits.)



## 12.5 Gage R&R Considerations

A measurement is characterized by location and spread, which are impacted by the following metrics:

- **Location:** Bias, stability, and linearity
- **Spread:** repeatability and reproducibility



## 12.5 Gage R&R Considerations

**Bias** assessments need an accepted reference value for a part, which can be done with tool room or layout inspection equipment.

- Measure one part in a tool room.
- Instruct one appraiser to measure the same part 10 times, using the gage being evaluated.
- The difference between the reference and the observed average is the measurement system bias.
- Express percent of process variation for bias.
- Express percent of tolerance for bias.



## 12.5 Gage R&R Considerations

Measurement system **stability** is the amount of total variation in system's bias over time on a given part or master part.

- One method of study is to plot the average and range of repeated master part readings on a regular basis.

**Linearity** graphs are a plot of bias values throughout the expected operating range of the gage.



## 12.5 Gage R&R Considerations

Expressions of measurement system spread:

- Standard deviation from gage R&R study multiplied by 5.15 (99% of normal distribution)
- Percent of tolerance
- Percent of process variation
- Number of distinct data categories (Discrimination or resolution)
  - Recommended discrimination is at most 1/10 of process capability ( $6\sigma$ )
  - Unacceptable discrimination symptoms can appear in a range chart (less than 4 possible values, or  $\frac{1}{4}$  of the ranges are zero.)

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## 12.5 Gage R&R Considerations



- Can be used for control only if the process variation is small or the loss function is flat; and the main source(s) of variation causes mean shift
- Unacceptable for estimating process parameters.
- Can be used with semi-variable control techniques
- Can produce insensitive control charts
- Only provides rough estimates
- Can be used with variable control charts
- Recommended for analysis

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## 12.6 Gage R&R Relationships

- A measurement process is said to be consistent when the results for the operators are repeatable, and the results between operators are reproducible.
- A gage is able to detect part-to-part variation whenever the variability of operator measurements is small relative to process variability.

$$\sigma_m = \sqrt{\sigma_e^2 + \sigma_o^2}$$

where  $\sigma_m$  = Measurement system standard deviation =  $\sqrt{GRR}$

$\sigma_e$  = Gage standard deviation =  $\sqrt{\text{Equipment Variation}} = \sqrt{EV}$

$\sigma_o$  = Appraiser std deviation =  $\sqrt{\text{Appraiser Variation}} = \sqrt{AV}$



## 12.6 Gage R&R Relationships

$$\sigma_T^2 = \sigma_p^2 + \sigma_m^2 \text{ or } TV = \sqrt{GRR + PV}$$

where  $\sigma_T^2$  = Total Variance (TV)

$\sigma_p^2$  = Process Variance (PV)

$\sigma_m^2$  = Measurement Variance = GRR

- The percent of process variation is estimated by

$$\%R\&R = \frac{\sigma_m}{\sigma_T} \times 100 \text{ or } \%GRR = 100 \sqrt{\frac{GRR}{TV}}$$

- The percent of tolerance is estimated by

$$\%Tolerance = \frac{5.15 \sigma_m}{tolerance} \times 100$$



## 12.6 Gage R&R Relationships

- The component of total process variation contributed by the measurement system for repeatability and reproducibility:

$$\frac{\sigma_m^2}{\sigma_T^2} = \frac{GRR}{TV}$$

- The components of total process variation contributed by the equipment, appraiser, and process are:

$$\frac{\sigma_e^2}{\sigma_T^2} = \frac{EV}{TV}, \frac{\sigma_o^2}{\sigma_T^2} = \frac{AV}{TV}, \frac{\sigma_p^2}{\sigma_T^2} = \frac{PV}{TV}$$



## 12.6 Gage R&R Relationships

- The number of distinct categories (ndc) is

$$\frac{\sigma_p}{\sigma_m} \times 1.41 = 1.41 \sqrt{\frac{PV}{GRR}}$$

- The number of distinct categories must be at least 5 for the measurement system to be acceptable.
- One generally recognized industry practice suggests a short method of evaluation using 5 samples, 2 operators, and no replication. A gage is considered acceptable if the gage error is  $\leq 20\%$  of the specification tolerance.



## 12.6 Gage R&R Relationships

- The output from a gage R&R analysis typically includes  $\bar{x}$  and  $R$  charts. The horizontal axis is segmented into regions for the various operators.
- The control limits are
 
$$UCL_{\bar{x}} = \bar{\bar{x}} + A_2\bar{R}; LCL_{\bar{x}} = \bar{\bar{x}} - A_2\bar{R}$$
 $\bar{\bar{x}}$  is the overall average (between and within operator),  $\bar{R}$  is an estimate of within operator variability.
- Out-of-control conditions in an  $\bar{x}$  chart indicate that part variability is high compared to R&R (desirable).
- The inconsistencies of appraisers appear as out-of-control (unpredictable process) conditions in the  $R$  chart.



## 12.8 Preparation for a MSA

1. Plan the approach. For instance, determine if there is appraiser influence in calibrating or using the instrument.
2. Select number of appraisers, number of sample of parts, and number of repeat reading. Consider using at least 2 operators and 10 samples, each operator measuring each sample at least twice (all using the same device). Select appraisers who normally operate the instruments.
3. Select sample parts from the process that represent its entire operating range. Number each part.
4. Ensure that the instrument has a discrimination that is at least one-tenth of the expected process variation of the characteristic to be read.



## 12.8 Preparation for a MSA

Other considerations:

1. Execute measurements in random order to ensure that drift or changes that occur will be spread randomly throughout the study.
2. Record readings to the nearest number obtained. When possible, make readings to nearest one-half of the smallest graduation (e.g., 0.00005 for 0.0001 graduations).
3. Use an observer who recognizes the importance of using caution when conducting the study.
4. Ensure that each appraiser uses the same procedure when taking measurements.



## 12.9 Example 12.1 Gage R&R

S<sup>4</sup>/IEE Application Example

- Manufacturing 30,000-foot-level metric (KPOV): An S<sup>4</sup>/IEE project was to improve the capability/performance of the diameter for a manufactured product. An MSA was conducted of the measurement gage.

5 samples, 2 appraisers. Each part is measured 3 times by each appraiser.



# 12.9 Example 12.1 Gage R&R

Appraiser 1						
Trial	Part 1	Part 2	Part 3	Part 4	Part 5	
1	217	220	217	214	216	
2	216	216	216	212	219	
3	216	218	216	212	220	
Avg.	216.3	218.0	216.3	212.7	218.3	216.3
Range	1.0	4.0	1.0	2.0	4.0	2.4

Appraiser 2						
Trial	Part 1	Part 2	Part 3	Part 4	Part 5	
1	216	216	216	216	220	
2	219	216	215	212	220	
3	220	220	216	212	220	
Avg.	218.3	217.3	215.7	213.3	220.0	216.9
Range	4.0	4.0	1.0	4.0	0.0	2.6

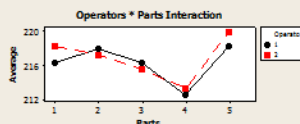
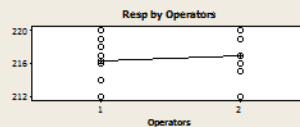
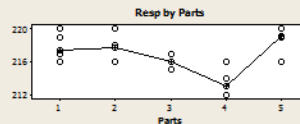
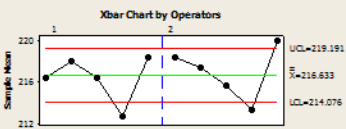
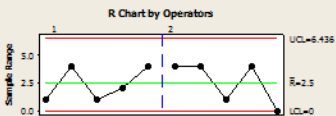
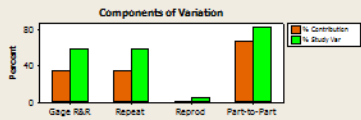


# 12.9 Example 12.1 Gage R&R

## Gage R&R (ANOVA) for Resp

Gage name:  
Date of study:

Reported by:  
Tolerance:  
Misc:



Minitab:  
Stat  
Quality Tools  
Gage Study  
Gage R&R Study  
(crossed)  
ANOVA method





## 12.9 Example 12.1 Gage R&R

### Gage R&R Study - ANOVA Method

#### Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Parts	4	129.467	32.3667	13.6761	0.013
Operators	1	2.700	2.7000	1.1408	0.346
Parts * Operators	4	9.467	2.3667	0.9221	0.471
Repeatability	20	51.333	2.5667		
Total	29	192.967			

#### Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Parts	4	129.467	32.3667	12.7763	0.000
Operators	1	2.700	2.7000	1.0658	0.312
Repeatability	24	60.800	2.5333		
Total	29	192.967			



## 12.9 Example 12.1 Gage R&R

### Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	2.54444	33.85
Repeatability	2.53333	33.70
Reproducibility	0.01111	0.15
Operators	0.01111	0.15
Part-To-Part	4.97222	66.15
Total Variation	7.51667	100.00

$$\frac{\sigma_m^2}{\sigma_T^2} = \frac{GRR}{TV} = \frac{2.54444}{7.51667}$$

$$\frac{\sigma_p^2}{\sigma_T^2} = \frac{PV}{TV} = \frac{4.97222}{7.51667}$$



## 12.9 Example 12.1 Gage R&R

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)
Total Gage R&R	1.59513	9.5708	58.18
Repeatability	1.59164	9.5499	58.05
Reproducibility	0.10541	0.6325	3.84
Operators	0.10541	0.6325	3.84
Part-To-Part	2.22985	13.3791	81.33
Total Variation	2.74165	16.4499	100.00

Number of Distinct Categories = 1

$$\%R\&R = \frac{\sigma_m}{\sigma_T} \times 100 = \frac{1.59513}{2.74165}$$



## 12.10 Linearity

- Linearity is the difference in the bias values through the expected operating range of the gage.
- For a linearity evaluation, one or more operators measure parts selected throughout the operating range of the gage.
- For each chosen parts, the average difference between the reference value and the observed average measurement is the estimated bias.
- If a graph between bias and reference follows a straight line throughout the operating range, a regression line is formed.
- The slope value is then multiplied by the process variation (tolerance) to determine an index of linearity of the gage.



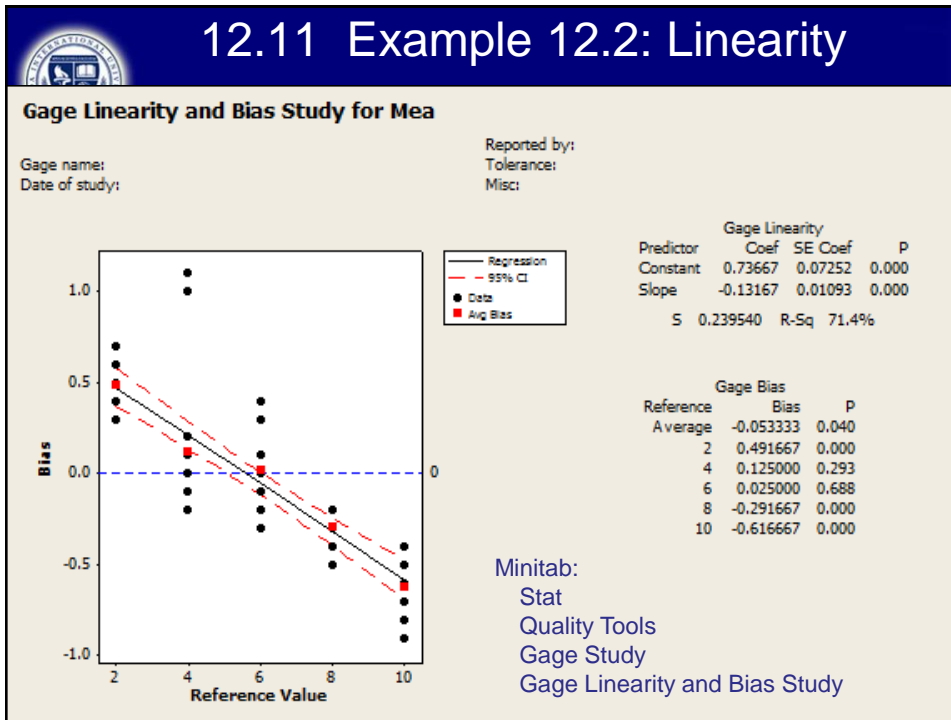
## 12.11 Example 12.2: Linearity

- 5 parts selected to represent the operating range of the gage.
- Layout inspection determined the part reference values.
- Appraisers measured each part 12 times in random.



## 12.11 Example 12.2: Linearity

Part	1	2	3	4	5
Ref.	2.00	4.00	6.00	8.00	10.00
1	2.70	5.10	5.80	7.60	9.10
2	2.50	3.90	5.70	7.70	9.30
3	2.40	4.20	5.90	7.80	9.50
4	2.50	5.00	5.90	7.70	9.30
5	2.70	3.80	6.00	7.80	9.40
6	2.30	3.90	6.10	7.80	9.50
7	2.50	3.90	6.00	7.80	9.50
8	2.50	3.90	6.10	7.70	9.50
9	2.40	3.90	6.40	7.80	9.60
10	2.40	4.00	6.30	7.50	9.20
11	2.60	4.10	6.00	7.60	9.30
12	2.40	3.80	6.10	7.70	9.40
Avg.	2.492	4.125	6.025	7.708	9.383
Range	0.40	1.30	0.70	0.30	0.50
Bias	0.492	0.125	0.025	-0.292	-0.617



## 12.12 Attribute Gage Study

- An attribute gage either accepts or rejects a part after comparison to a set of limits.
- Select 20 parts (some parts are slightly below and some above specification limits).
- Use 2 appraisers and conduct the study in a manner to prevent appraiser bias. Appraisers inspect each part twice, deciding whether the part is acceptable or not.
- If all measurements agree, the gage is accepted.
- Gage needs improvement or reevaluation if measurement decisions do not agree.

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## 12.13 Example 12.3: Attribute Gage Study

- A company is training 5 new appraisers for the written portion of a standardized essay test. The ability of the appraiser to rate essays relative to standards needs to be assessed.
- 15 essays were rated by each appraiser on a five-point scale (-2, -1, 0, 1, 2).



## 12.13 Example 12.3: Attribute Gage Study

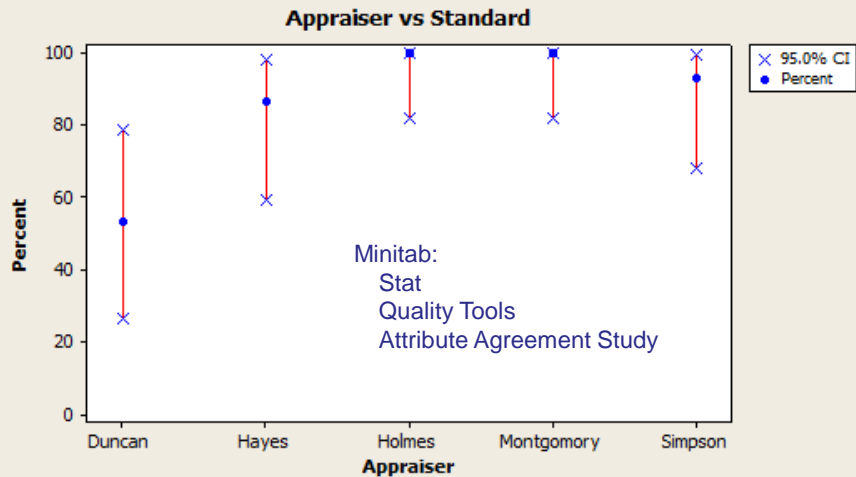
Appraiser	Sample	Rating	Attribute	Appraiser	Sample	Rating	Attribute	Appraiser	Sample	Rating	Attribute
Simpson	1	2	2	Simpson	6	1	1	Simpson	11	-2	-2
Montgomery	1	2	2	Montgomery	6	1	1	Montgomery	11	-2	-2
Holmes	1	2	2	Holmes	6	1	1	Holmes	11	-2	-2
Duncan	1	1	2	Duncan	6	1	1	Duncan	11	-2	-2
Hayes	1	2	2	Hayes	6	1	1	Hayes	11	-1	-2
Simpson	2	-1	-1	Simpson	7	2	2	Simpson	12	0	0
Montgomery	2	-1	-1	Montgomery	7	2	2	Montgomery	12	0	0
Holmes	2	-1	-1	Holmes	7	2	2	Holmes	12	0	0
Duncan	2	-2	-1	Duncan	7	1	2	Duncan	12	-1	0
Hayes	2	-1	-1	Hayes	7	2	2	Hayes	12	0	0
Simpson	3	1	0	Simpson	8	0	0	Simpson	13	2	2
Montgomery	3	0	0	Montgomery	8	0	0	Montgomery	13	2	2
Holmes	3	0	0	Holmes	8	0	0	Holmes	13	2	2
Duncan	3	0	0	Duncan	8	0	0	Duncan	13	2	2
Hayes	3	0	0	Hayes	8	0	0	Hayes	13	2	2
Simpson	4	-2	-2	Simpson	9	-1	-1	Simpson	14	-1	-1
Montgomery	4	-2	-2	Montgomery	9	-1	-1	Montgomery	14	-1	-1
Holmes	4	-2	-2	Holmes	9	-1	-1	Holmes	14	-1	-1
Duncan	4	-2	-2	Duncan	9	-2	-1	Duncan	14	-1	-1
Hayes	4	-2	-2	Hayes	9	-1	-1	Hayes	14	-1	-1
Simpson	5	0	0	Simpson	10	1	1	Simpson	15	1	1
Montgomery	5	0	0	Montgomery	10	1	1	Montgomery	15	1	1
Holmes	5	0	0	Holmes	10	1	1	Holmes	15	1	1
Duncan	5	-1	0	Duncan	10	0	1	Duncan	15	1	1
Hayes	5	0	0	Hayes	10	2	1	Hayes	15	1	1



## 12.13 Example 12.3: Attribute Gage Study

### Assessment Agreement

Date of study:  
Reported by:  
Name of product:  
Misc:



## 12.13 Example 12.3: Attribute Gage Study

### Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Duncan	15	8	53.33	(26.59, 78.73)
Hayes	15	13	86.67	(59.54, 98.34)
Holmes	15	15	100.00	(81.90, 100.00)
Montgomery	15	15	100.00	(81.90, 100.00)
Simpson	15	14	93.33	(68.05, 99.83)

# Matched: Appraiser's assessment across trials agrees with the known standard.



## 12.14 Gage Study of Destructive Testing

- Destructive tests cannot test the same unit repeatedly to obtain an estimate for pure measurement error.
- An upper bound on measurement error for destructive tests is determinable using the control chart technique (Wheeler 1990).
- It is often possible to minimize the product variation between pairs of measurements through the careful selection of the material to be measured.
- Through repeated duplicate measurements on material that is thought to minimize product variation, an upper bound is obtainable for the variation due to the measurement process.

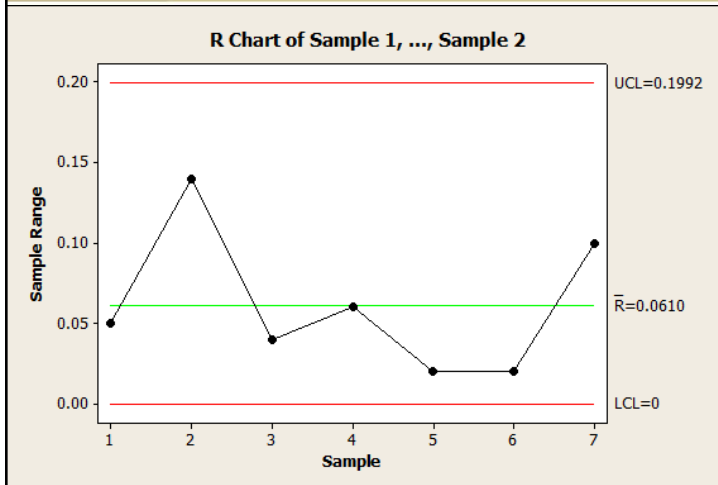


## 12.15 Example 12.4: Gage Study of Destructive Testing

Lot	1	2	3	4	5	6	7
Sample1	20.48	19.37	20.35	19.87	20.36	19.32	20.58
Sample2	20.43	19.23	20.39	19.93	20.34	19.30	20.68
Average	20.46	19.30	20.37	19.90	20.35	19.31	20.63
Range	0.05	0.14	0.04	0.06	0.02	0.02	0.10



## 12.15 Example 12.4: Gage Study of Destructive Testing



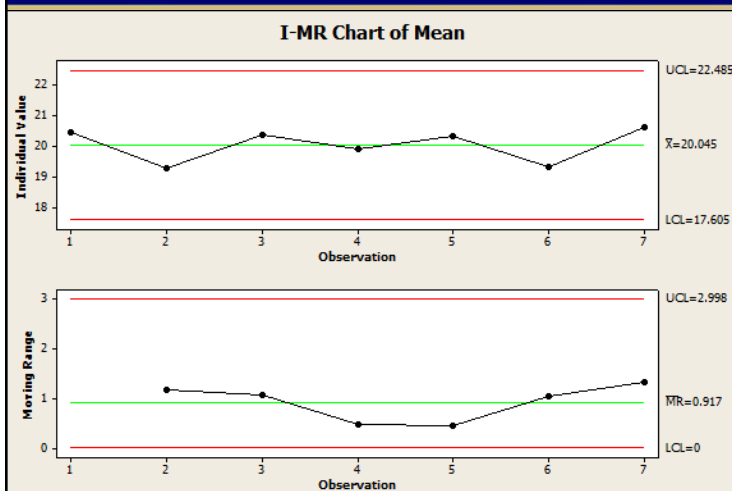
$$\sigma_m = \frac{\bar{R}}{d_2} = \frac{0.0614}{1.128}$$

$$= 0.054$$

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING



## 12.15 Example 12.4: Gage Study of Destructive Testing



$$\sigma_p = \frac{\bar{R}}{d_2} = \frac{0.9175}{1.128}$$

$$= 0.813$$

$$ndc = \frac{\sigma_p}{\sigma_m} \times 1.41$$

$$= 21.2$$

FLORIDA INTERNATIONAL UNIVERSITY  
COLLEGE OF ENGINEERING





## 12.16 A 5-step Measurement Improvement Process

- Machine Variation
- Fixture Study
- Accuracy (Linearity)
- Repeatability and Reproducibility
- Long-term Stability
  - Source of variation
  - How to conduct the test
  - Acceptance criteria
  - Comments



## 12.16 A 5-step Measurement Improvement Process

### Terminology

- $NPV$  = Normal Process Variation
- $T$  = Tolerance
- $P$  = Precision
- $S_{MS}$  = Std. deviation of measurement system
- $S_{Total}$  = Std. deviation of total variability of measurements over time
- $P/T = (5.15 \times S_{MS})/Tolerance$
- $P/NPV = (5.15 \times S_{MS})/(5.15 \times S_{Total}) = S_{MS}/S_{Total}$