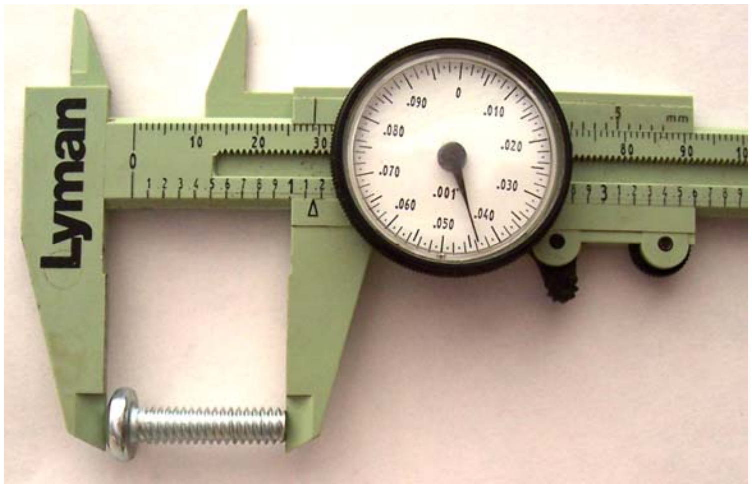
Measurement System Analysis

Gage R&R



When I first got involved with quality, I learned about the "five M's" that constituted most root causes: man, machine, materials, methods, and ***measurement***.

Because I worked in a predominantly service industry, I couldn't quite grasp how measurement could be a *cause* of variation. But, if you work in manufacturing, you know that **gages** can be used in ways that are inexact and thus can be a cause of variation. If you're measuring parts to ensure that they meet customer requirements, but your gage or your measurement process vary too much, you might pass parts that should fail, and fail parts that should pass. To ensure that your customer gets what they want you will want to make sure that your measurements are accurate.

Measurement Systems Analysis (MSA)

A measurement system consists of the processes, standards and gages used to measure a specific feature of a product—height, weight, length, weight, volume, etc. Measurement systems analysis helps determine if the equipment (i.e., the gage) and the measurement process can get the same result consistently.

Measurement System Analysis (MSA) uses many methods to evaluate consistency:

Type of Measurement System	**Methods**
Variable Data	Average and Range, ANOVA, Bias, Linearity
Attribute data	Signal Detection, Hypothesis testing
Destructive testing	Control Charts

MSA is actually quite simple, but even seasoned SPC veterans don't seem to understand it. So I thought I'd simplify it for you.

First, the most common mistake people make when conducting a Gage R&R study is forgetting that it is evaluating their measurement system and NOT their products. Gage R&R does not care about how good your products are. It only cares about how good you measure your products.

Second, when you manufacture products, you want to monitor the output of your machines to make sure that they are producing products that meet the customer's specifications. This means that you have to measure samples coming off the line to determine if they are meeting your customer's requirements.

Third, *Gage R&R* studies are usually performed on variable data - height, length, width, diameter, weight, viscosity, etc. But you can also use pass/fail gages that make for quick analysis, in which case, use the Attribute Gage R&R study.

Fourth, when you measure, there are three sources of variation that come into play:

- **Part variation - PV** (differences between individual pieces) should account for 90+% of all variation.
- **Appraiser variation- AV** (a.k.a., *reproducibility*) - Can two different people get the same measurement using the same *gage*? **Hint:** Different processes yield different results.
- **Equipment variation - EV** (a.k.a., *repeatability*) - Can the same person get the same measurement using the same *gage* on the same part in two or more trials?

You want most of the variation to be caused by *variation among the parts*, and less than 10% of the variation to be caused by the appraisers and equipment. Makes sense, doesn't it? If one appraiser can't get the same measurement twice, or two appraisers can't get the same measurement, then your measurement system is a key source of error.

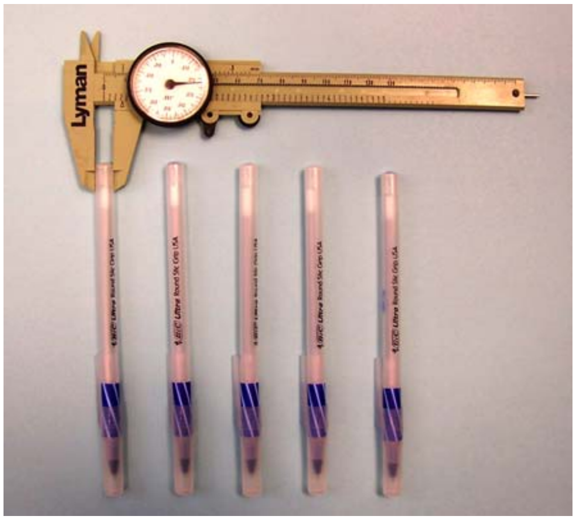
Conducting a Gage R&R Study

To conduct a Gage R&R study, you will need:

- Five to ten parts from one batch or lot (number each part)
- two appraisers (people who measure the parts)
- one measurement tool or *gage*
- and a minimum of two measurement trials, on each part, by each appraiser
- a *Gage R&R* tool like the **Gage R&R template** in the QI Macros.

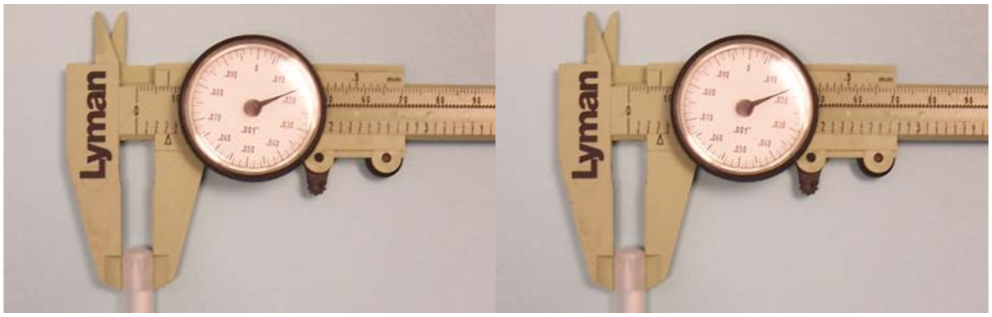
The following examples use the QI Macros Gage R&R template. If you don't have the QI Macros you can download a 30 day evaluation copy at http://www.qimacros.com/freestuff.html or order a copy at https://www.qimacros.com/order230.html.

Step 1. Select 5-10 DIFFERENT parts and at least two "appraisers" and one gage.
Hint: If you don't use enough parts, part variation will be small and magnify the appraiser/equipment variation.



Step 2. RANDOMLY, have appraisers measure each part at least twice.

Hint: If they measure the same part two or three times in a row or see each other's results, they'll start to bend their results to get the desired outcome—consistency.



Step 3. Then enter the results of each measurement in Cells C3:L7, C11:L15, C19:L23

	A	B	C	D	E	F	G
1	**Gage R&R**	http://www.aiag.org/					**Part Nun**
2	**Average & Range Method**		1	2	3	4	5
3	Appraiser 1	Trial 1	0.315	0.314	0.316	0.318	0.316
4	Enter your data here->	Trial2	0.315	0.314	0.315	0.318	0.315
5		Trial3					
6		Trial4					
7		Trial 5					
8		Total	0.63	0.628	0.631	0.636	0.631
9		Average-	0.315	0.314	0.3155	0.318	0.3155
10		Range1	0	0	0.001	0	0.001
11	Appraiser 2	Trial 1	0.315	0.314	0.315	0.317	0.316
12	Enter your data here->	Trial2	0.315	0.314	0.315	0.318	0.316
13		Trial3					
14		Trial4					
15		Trial 5					
16		Total	0.63	0.628	0.63	0.635	0.632
17		Average-	0.315	0.314	0.315	0.3175	0.316
18		Range2	0	0	0	0.001	0
19	Appraiser 3	Trial 1	0.314	0.314	0.315	0.318	0.316
20	Enter your data here->	Trial2	0.314	0.313	0.315	0.318	0.315

Step 4. The Gage R&R template will automatically calculate all of the values based on your data. In this case, EV > AV, indicating the gage may be hard to read accurately. **Note:** Some people expect EV + AV + PV to equal 100%. They will not equal 100% and are not supposed to (AIAG MSA Third Edition, p 116).

45		% Using
46	AIAG - Automotive Indus	TV
47	EV (Equipment Variatior	0.0003
48	**%EV**	**17.9%**
49	AV: (Appraiser Variation	0.00019
50	**%AV**	**11.4%**
51	R&R (Gage Capability)	0.0003
52	**%R&R**	**21.2%**
53	PV (Part Variation)	0.0016
54	%PV	97.7%
55	TV (Total Variation)	0.0016

Using Spec Tolerance: If you want to evaluate your measurement system based on the tolerance instead of process variation, input your specification tolerance into cell B43. Note: the spec tolerance = USL - LSL

Repeatability(EV)	0.0003		0.577	0.7
Reproducibility(AV)	0.0002		0.4299	0.48
Gage Capability(R&R)	0.0003	Spec Tolerance = USL - LSL		
Spec Tolerance	0.005			
	% Using	%		
AIAG - Automotive Indus	TV	To		
EV (Equipment Variatior	0.0003			
%EV	17.9%	35%	# Parts	#Trial
AV: (Appraiser Variation	0.00019		5	
%AV	11.4%	22%		

Use the calculations in cells C 48, C50 and C52 to evaluate your measurement system:

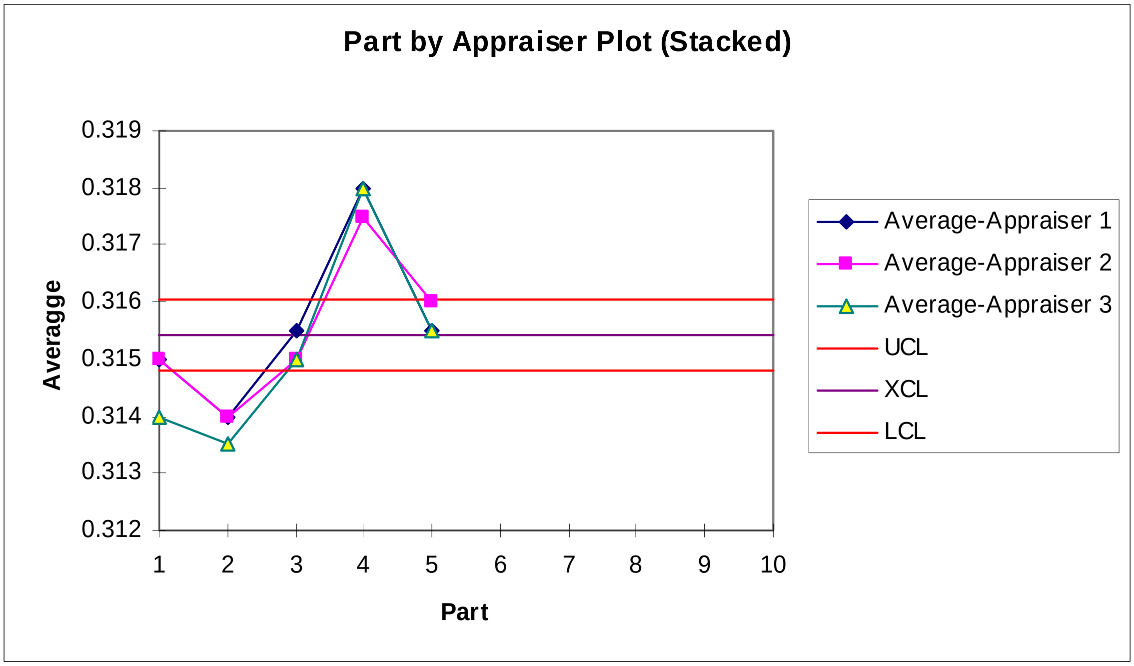
45		% Using	% Using
46	AIAG - Automotive Indus	TV	Tolerance
47	EV (Equipment Variatio	0.0003	
48	**%EV**	**17.9%**	**35%**
49	AV: (Appraiser Variation	0.00019	
50	**%AV**	**11.4%**	**22%**
51	R&R (Gage Capability)	0.0003	
52	**%R&R**	**21.2%**	**42%**
53	PV (Part Variation)	0.0016	
54	%PV	97.7%	193%
55	TV (Total Variation)	0.0016	

Step 5. Evaluate your measurement system based on %R&R and the Average and Range charts. There are two types of charts Average and Range. In the QI Macros Gage R&R template, each is shown stacked and side by side.

- **Average Chart** - Plots the average of all trials for each part for each appraiser and graphs them on a line. Each line shows how consistent each appraiser measures from part to part. When the lines for each appraiser are "stacked" on top of each other or displayed side by side, you can easily see if one appraiser gets more consistent results between parts or if one appraiser tends to measure higher or lower than the others.
- **Range Chart** - Plots the range (Max - Min) of all trials for each part for each appraiser and graphs them on a line. Each line shows how consistent each appraiser measured each part. If the range is zero, then the appraiser measured the part the same in every trial. Like the Average charts, the Range charts are displayed both stacked and side by side.

Step 6. As you can see in the following Stacked Average chart, all three appraisers get close to the same measurement on each part. Appraiser 3 is biased to be lower than the other two.

Half of the parts are above or below the control limits which means, oddly enough, that the measurement system should be adequate to detect part-to-part variation. Less than half outside may imply a lack of precision or a non-representative sample. (AIAG MSA Third Edition, p 102.)



Evaluating the Results - Gage Repeatability and Reproducibility

If ***repeatability*** (i.e., EV - Can the same person using the same gage measure the same thing consistently) is larger than *reproducibility* (EV>AV), reasons include:

- *Gage* needs maintenance (*gages* can get corroded)
- *Gage* needs to be redesigned to be used more accurately
- Clamping of the part or *gage*, or where it's measured needs to be improved (imagine measuring a baseball bat at various places along the tapered contour; you'll get different results.)
- Excessive within-part variation (imagine a steel rod that's bigger at one end than the other. If you measure different ends each time, you'll get widely varying results.)

If ***reproducibility*** (i.e., AV-can two appraisers measure the same thing and get the same result) is larger than *repeatability* (AV>EV), reasons include:

1. Operator needs to be better trained in how to use and read gage

2. Calibrations on gage are not clear
(In the case of the pens, a digital readout might reduce error.)

3. A fixture may be required to help the operator use the gage more consistently

Gage System Acceptability

How acceptable is your measurement system?

%R&R<10%	Gage System Okay (Most variation caused by parts, not people or equipment)
%R&R<30%	May be acceptable based on importance of application and cost of gage or repair.
%R&R>30%	Gage system needs improvement (people and equipment cause over 1/3 of variation)

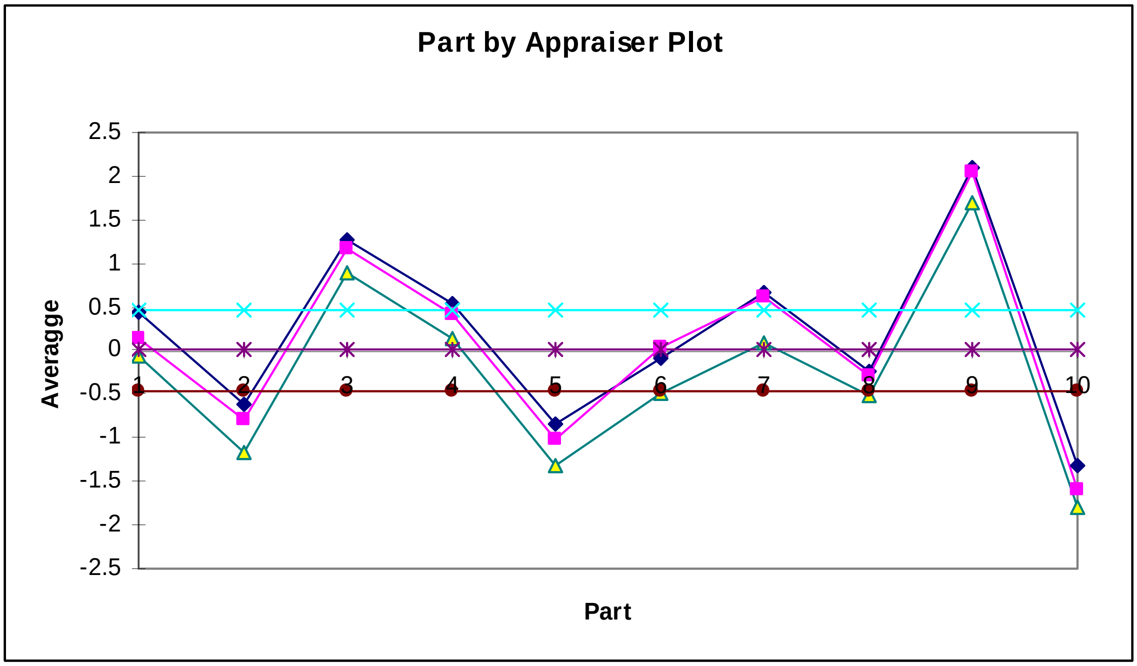
Here are samples of the QI Macros **Gage R&R template** input sheet and results sections using sample data from the AIAG Measurement Systems Analysis Third Edition:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**Gage R&R**	http://www.aiag.org/					**Part Number**		http://www.qimacros.com/free-lean-six-sigma-tips/aiag-ms					
2	**Average & Range Method**		1	2	3	4	5	6	7	8	9	10	Sum	
3	Appraiser 1	Trial 1	0.29	-0.56	1.34	0.47	-0.8	0.02	0.59	-0.31	2.26	-1.36	5.710	
4	Enter your data here->	Trial2	0.41	-0.68	1.17	0.5	-0.92	-0.11	0.75	-0.2	1.99	-1.25		
5		Trial3	0.64	-0.58	1.27	0.64	-0.84	-0.21	0.66	-0.17	2.01	-1.31	2.110	
6		Trial4												Xbar1
7		Trial 5												0.190333
8		Total	1.34	-1.82	3.78	1.61	-2.56	-0.3	2	-0.68	6.26	-3.92		
9		Average-	0.4467	-0.607	1.26	0.5367	-0.853	-0.1	0.6667	-0.227	2.0867	-1.307		Rbar1
10		Range1	0.35	0.12	0.17	0.17	0.12	0.23	0.16	0.14	0.27	0.11		0.184
11	Appraiser 2	Trial 1	0.08	-0.47	1.19	0.01	-0.56	-0.2	0.47	-0.63	1.8	-1.68	2.050	
12	Enter your data here->	Trial2	0.25	-1.22	0.94	1.03	-1.2	0.22	0.55	0.08	2.12	-1.62		
13		Trial3	0.07	-0.68	1.34	0.2	-1.28	0.06	0.83	-0.34	2.19	-1.5	0.890	
14		Trial4												Xbar2
15		Trial 5												0.068333
16		Total	0.4	-2.37	3.47	1.24	-3.04	0.08	1.85	-0.89	6.11	-4.8		
17		Average-	0.1333	-0.79	1.1567	0.4133	-1.013	0.0267	0.6167	-0.297	2.0367	-1.6		Rbar2
18		Range2	0.18	0.75	0.4	1.02	0.72	0.42	0.36	0.71	0.39	0.18		0.513
19	Appraiser	Trial 1	0.04	-1.38	0.88	0.14	-1.46	-0.29	0.02	-0.46	1.77	-1.49	-7.630	
20	Enter your data here->	Trial2	-0.11	-1.13	1.09	0.2	-1.07	-0.67	0.01	-0.56	1.45	-1.77		
21		Trial3	-0.15	-0.96	0.67	0.11	-1.45	-0.49	0.21	-0.49	1.87	-2.16	-2.840	
22		Trial4												Xbar3
23		Trial 5												-0.25433
24		Total	-0.22	-3.47	2.64	0.45	-3.98	-1.45	0.24	-1.51	5.09	-5.42		
25		Average-	-0.0733	-1.157	0.88	0.15	-1.327	-0.483	0.08	-0.503	1.6967	-1.807		Rbar3
26		Range3	0.19	0.42	0.42	0.09	0.39	0.38	0.2	0.1	0.42	0.67		0.328

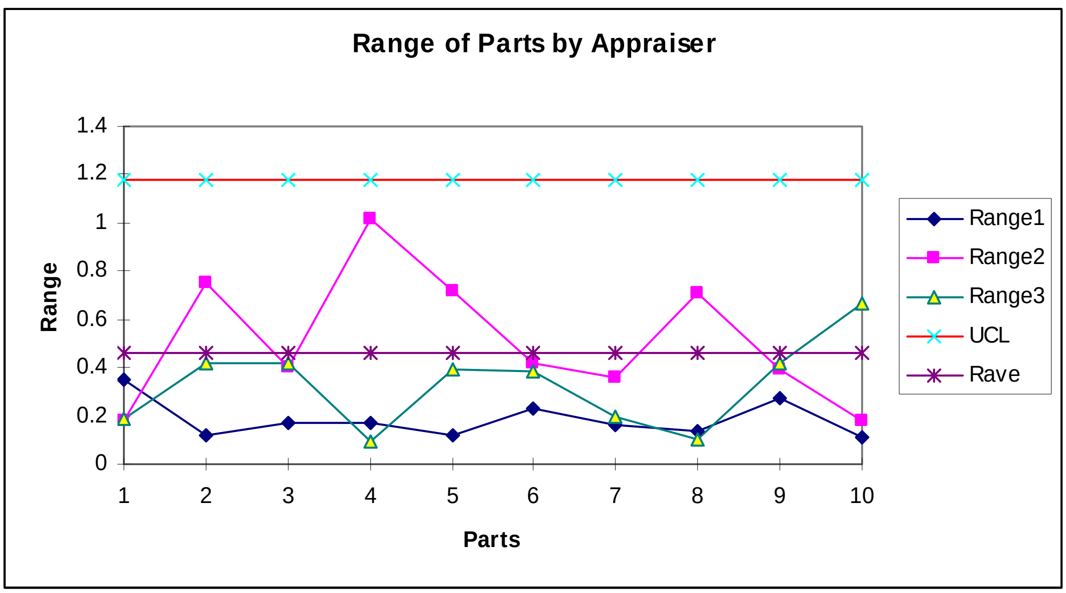
GAGER&R / Range Method / Bias / Linearity / Attribute Gage Worksheet / Analytic Attribute M

	A	B	C	D	E	F	G	H	I	J	K	L
44	Acceptability(%)	0.0692		Gage system may be acceptable based on importance of application and cost								
45		% Using	% Using		Operator may need to be better trained or gage is hard to read							
46	AIAG - Automotive Indu	TV	Tolerance									
47	EV (Equipment Variatio	0.2019					Equipment Variation (EV)					
48	%EV	17.6%	27%	# Parts	#Trials	#Ops		% of Total Variation (TV)				
49	AV: (Appraiser Variatio	0.22967		10	3	3	Appraiser Variation(AV)					
50	%AV	20.0%	31%					% of Total Variation (TV)				
51	R&R (Gage Capability)	0.3058					Repeatability and Reproducibility (R&R)					
52	%R&R	26.7%	42%					% of Total Variation (TV)				
53	PV (Part Variation)	1.1046					Part Variation (PV)					
54	%PV	96.4%	150%					% of Total Variation (TV)				
55	TV (Total Variation)	1.1461					Total Variation (TV)					
56	**Calculate GageR&R using Anova**				With Interaction		Without Interaction					
57	**Anova Source**	**df**	**SS**	**MS**	**F**	**P**	**F**	**P**				
58	**Appraiser**	2	3.1673	1.5836	34.44	0.0000	39.617	0.0000				
59	**Parts**	9	88.3619	9.818	213.52	0.0000	245.61	0.0000				
60	**Appraiser x Part**	18	0.3590	0.0199	0.4337	0.9741						
61	**Gage w AP Interactio**	60	2.7589	0.0460								
62	**Gage w/o AP Interact**	78	3.1179	0.0400								
63	**Total**	89	94.647									
64												
65	Without Interaction	Estimate of Variance	Std. Dev		Total Variation	% Contribution	With Interaction	Estimate of Variance	Std. Dev		Total Variation	% Contribution
66	**Repeatability**	0.03997	0.1999	EV	18%	3%	**Repeata**	0.0460	0.2144	EV	20%	4%
67	**Appraiser**	0.0515	0.2268	AV	21%	4%	**Apprais**	0.0521	0.2283	AV	21%	4%
68	**AppraiserxPart**	0	0	INT	0%	0%	**Apprais**	0.0087	0.0932	INT	9%	1%
69	**R&R**	0.09143	0.3024	R&R	28%	8%	**R&R**	0.1068	0.3268	R&R	30%	9%
70	**Part**	1.08645	1.0423	PV	96%	92%	**Part**	1.0887	1.0434	PV	96%	92%
71				TV	5.5893					TV	5.6309	

In this example, %R&R is 26.7% and the Anova %R&R is 28%, but is it acceptable? It depends. Let's look at the charts. In the Stacked Average Chart, measurements are similar, but one appraiser is biased consistently lower than the other.



If we look at the Stacked Range Chart, we can see that the measurements are in control and that appraiser 1 has less variation when measuring each part than the other two.



What To Look For If %R&R Is Not Acceptable

Repeatability: Percent Equipment Variation (%EV). If you simply look at the measurements, can each appraiser get the same result on the same part consistently, or is there too much variation?

<u>Example</u> looking at measurements from one appraiser only:

Trial 1	0.315	0.314	0.315	0.317	0.316
Trial2	0.315	0.314	0.315	0.318	0.316

No Equipment Variation: (Parts 1, 2, 3, 5)

Equipment Variation: (Part 4 : 0.317 and 0.318)

Reproducibility: Percent Appraiser Variation (% AV-can two appraisers measure the same thing and get the same answer?)

<u>Example</u> looking at *measurements* of the same part by two appraisers:

Appraiser 1	Trial 1	0.315	Appraiser 2	Trial 1	0.315
Enter your data here->	Trial2	0.315	Enter your data here->	Trial2	0.315
	Trial3			Trial3	
	Trial4			Trial4	
	Trial 5			Trial 5	
	Total	0.63		Total	0.63
	Average-	0.315		Average-	0.315
	Range1	0		Range2	0
Appraiser 2	Trial 1	0.315	Appraiser 3	Trial 1	0.314
Enter your data here->	Trial2	0.315	Enter your data here->	Trial2	0.314

No Appraiser Variation:
Appraiser 1, Part 1: 0.315;
Appraiser 2, Part 1: 0.315

Appraiser Variation:
Appraiser 2, Part 1: 0.315
Appraiser 3, Part 1: 0.314

If you look at the line graph of appraiser performance, you'll be able to tell if one person consistently over reads or under reads the measurement.

Mistakes People Make

Many people call us because they don't like the answer they get using the QI Macros **Gage R&R** template. Most of the time, it's because they didn't follow the instructions for conducting the study or they're confused about what Measurement System Analysis does. Here are some of the common mistakes I've seen:

- Using fake data. Most people try to test the Gage R&R by using data they've dreamed up rather than the AIAG SPC data in c:\qimacros\testdata.
- Using only one part. If you only use one part, THERE CAN'T BE ANY PART VARIATION, so people and equipment will be the ONLY source of variation.
- Using the one part measurement for all 10 parts (again, there won't be any part variation, so it all falls on the people and equipment).
- Using too many trials (if you use five trials, you have more opportunity for equipment variation).
- Using too many appraisers (if you use all three, you have more opportunity for appraiser variation).
- Using a gage that measures in too much detail. If your part is 74mm +/- 0.05, then you don't need a gage that measures to a thousandth of an inch (0.001) you only need one that measures to the hundredth of an inch (0.01).
- Confusing what the study says about your measurement system with what it says about your products. Gage R&R evaluates your measurement system NOT your products.

Challenges You Will Face

One customer faced an unusual challenge: they were producing parts so precisely that there was little or no part variation even when measured down to 1/10,000th of an inch. Their existing *gage*s had ceased to detect any variation from part to part.

As your process improves and your product approaches the ideal target measurement, you'll have less part variation and more chance for your equipment or people to become the major source of variation. As your product and your process improve, your measurement system will need to improve as well.

Gage R&R Example

In the following example, the supplier is producing sleeves that can be no larger than 0.6710 (USL). The supplier is producing sleeves that measure 0.6705 to 0.6707 that fit well, but they wondered why their gage system needs improvement.

Anytime %R&R is greater than 10%, you *may* need to improve the measurement system. When it's greater than 30% (i.e., equipment and appraiser variation is more than 30% of total variation), you *need to* improve the measurement system: In this example, the [QI Macros]() Gage R&R template gives these values for %EV, %AV, and %R&R:

	A	B
45		% Using
46	AIAG - Automotive Indus	TV
47	EV (Equipment Variatior	0.00004
48	**%EV**	**75.1%**
49	AV: (Appraiser Variation	0.00002
50	**%AV**	**42.2%**
51	R&R (Gage Capability)	0.00005
52	**%R&R**	**86.1%**
53	PV (Part Variation)	0.00003
54	%PV	50.8%
55	TV (Total Variation)	0.00006

Here's what most people don't understand about Gage R&R studies:

1. **Gage R&R doesn't care how good your parts are** (i.e., the parts all measured between .6705 and .6707 and were below the USL).
2. **All that matters is how accurately and consistently you can measure the parts** (i.e., if a part is .6705, every appraiser that measures that part should get .6705, not .6704 or .6706 or .6707).

	A	B	C	D	E	F	G
1	**Gage R&R**						**Part Num**
2	**Average & Range Method**		1	2	3	4	5
3	Appraiser 1	Trial 1	0.6705	0.6706	[illegible]	[illegible]	[illegible]
4	Enter your data here->	Trial2	0.6705	0.6705	[illegible]	[illegible]	[illegible]
5		Trial3	0.6705	0.6705	[illegible]	[illegible]	[illegible]
6		Trial4					
7		Trial 5					
8		Total	2.0115	2.0116	[illegible]	[illegible]	[illegible]
9		Average-Ap	0.6705	0.6705	[illegible]	[illegible]	[illegible]
10		Range1	0.0000	0.0001	[illegible]	[illegible]	[illegible]
11	Appraiser 2	Trial 1	0.6706	0.6706	[illegible]	[illegible]	[illegible]
12	Enter your data here->	Trial2	0.6705	0.6705	0.6706	0.6706	0.6706
13		Trial3	0.6706	0.6706	0.6705	0.6706	0.6707

Appraiser can't get the same answer on the same part with the same gage. (EV)

And neither can any other appraiser

3. In this case, each appraiser *cannot* get the same answer on the same part (see above), which means that there is too much equipment variation (EV>75%).

4. Since EV (75%) is greater than AV (42%), the problem is most likely that:

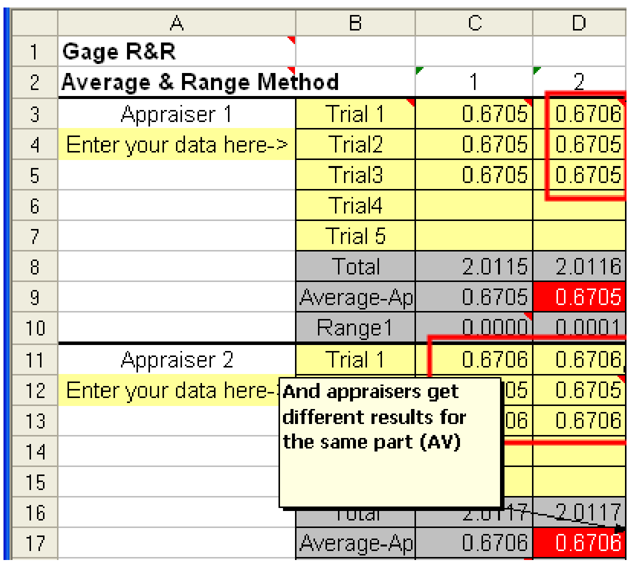
a. Gage instrument needs maintenance

b. Gage needs to be redesigned

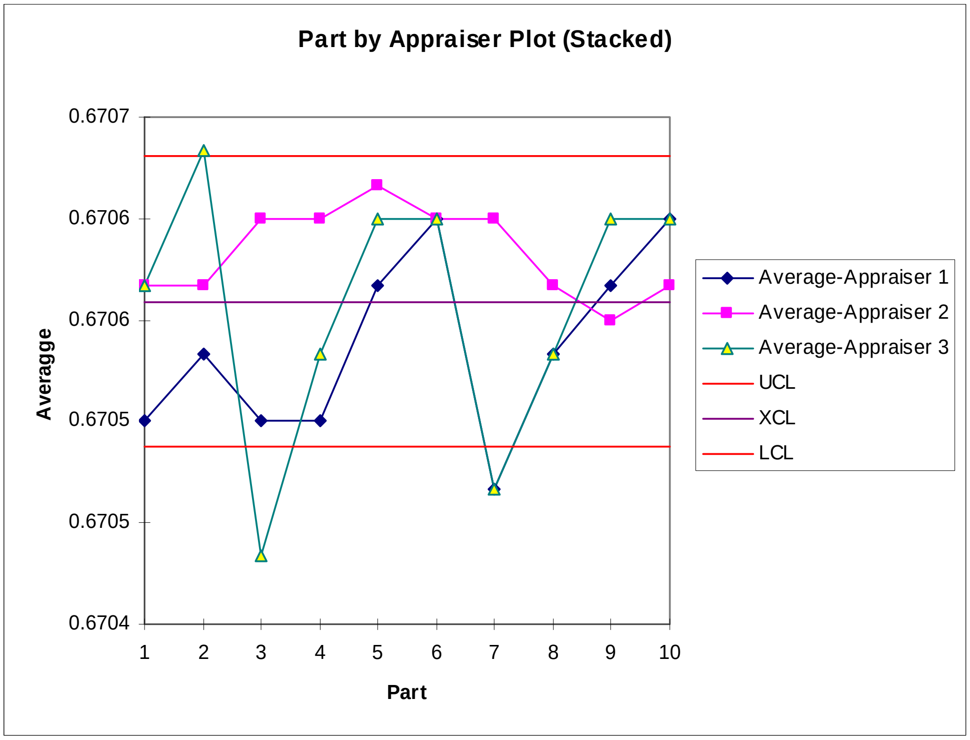
c. Clamping or location needs to be improved

d. Excessive within-part variation

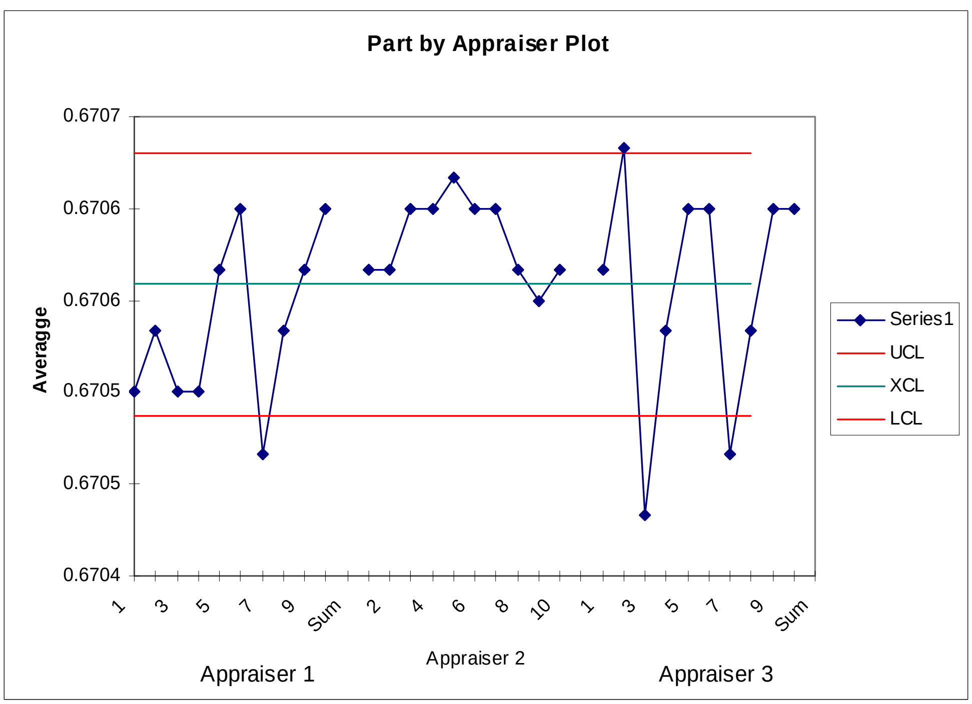
5. Additionally, AV=42 % indicates the appraisers *cannot* get the same answer on the same part, which means that there is too much variation between appraisers.



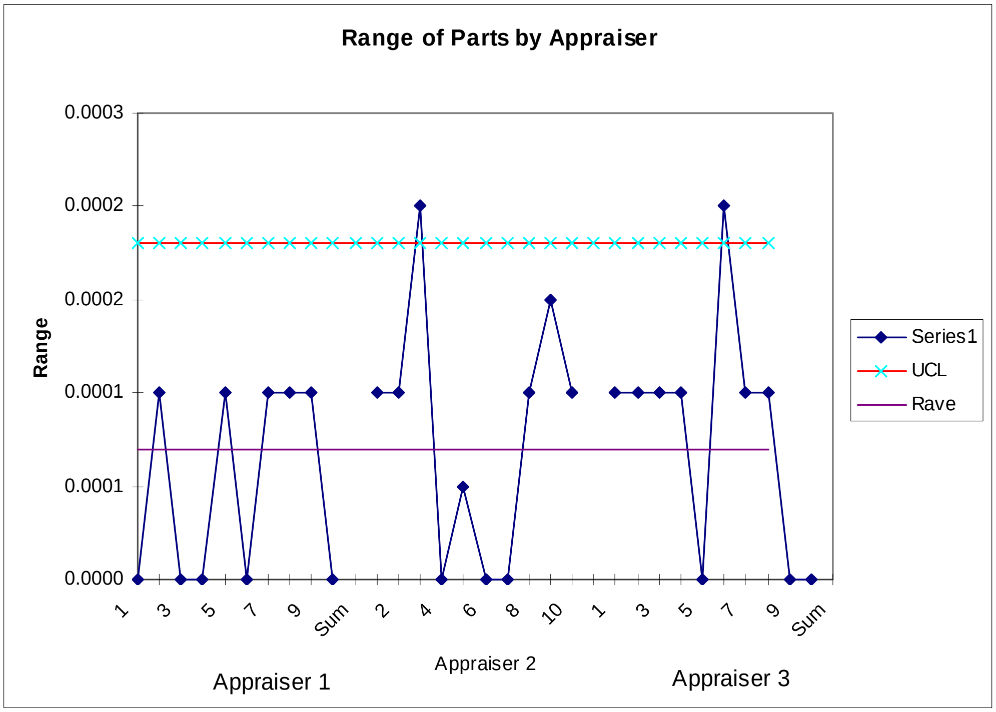
You can see this on the Stacked Average Part by Appraiser plot.



6. Look at the Average Part by Appraiser plot side by side. Appraiser 2 is the most consistent; 1 and 3 vary wildly. Since over 50% of the averages are within the LCL and UCL, the measurement system is not adequate to detect part variation.



7. If you look at the side by side range chart of the three measurements, you'll see that Appraiser 1 has less variation and five parts (1,3,4,7,and 10) measured exactly the same each time.



8. If you enter a reference value into the template, you'll discover that Appraiser 1 is biased *low* compared to the other two appraisers. Note: The reference value is either the actual value of the part or the target value if you don't have the actual measurement.

	Trial4	Xbar1	Reference
	Trial 5	0.6705	0.6707
	Total		Bias
	Average-Ap	Rbar1	-0.0002
	Range1	0.0000	
Appraiser 2	Trial 1		
Enter your data here-	And appraiser		
	different resul		
	the same part	Xbar2	Reference
		0.6706	0.6707
	Total		Bias
	Average-Ap	Rbar2	-0.0001
	Range2	0.0001	
Appraiser 3	Trial 1		
Enter your data here->	Trial2		
	Trial3		
	Trial4	Xbar3	Reference
	Trial 5	0.6706	0.6707
	Total		Bias
	Average-Ap	Rbar3	-0.0001

9. Based on these results, figure out how to copy appraiser 1’s method to measure each part consistently and 2's method to achieve consistency between parts to reduce AV. The most common problems related to appraiser variation are:

a. Operator needs to be better trained in how to use and read gage

b. Calibrations on gage are not clear

c. A fixture is required to help the operator use the gage more consistently

I’d guess that they are not all measuring the part at the same location in the same way. Solve that problem and your Gage R&R results will improve.

Bias and Linearity

Two other factors affect the accuracy of your measurement system: Bias and Linearity.

Bias asks: Does your *gage* tend to over- or under-read the same size part? (Imagine measuring the length or diameter of a steel rod with known dimensions.)

Linearity: Does your *gage* tend to over-read or under-read across a range of different sized parts? (Imagine using the *gage* on tin cans of various diameters, from small, 6 oz. juice cans to 64 oz., family-sized cans.)

If you want to know the "bias" of your *gage*, simply input the "target" or "reference" value for the parts being measured into the **Gage R&R template**, and the template will calculate the bias of the *gage* (plus or minus). Reference values are determined by using a calibrated gage that is highly accurate.

If you want to know the "bias and linearity" of your *gage*, switch to the linearity worksheet in the QI Macros **Gage R&R template** and conduct a linearity study.

Linearity Study

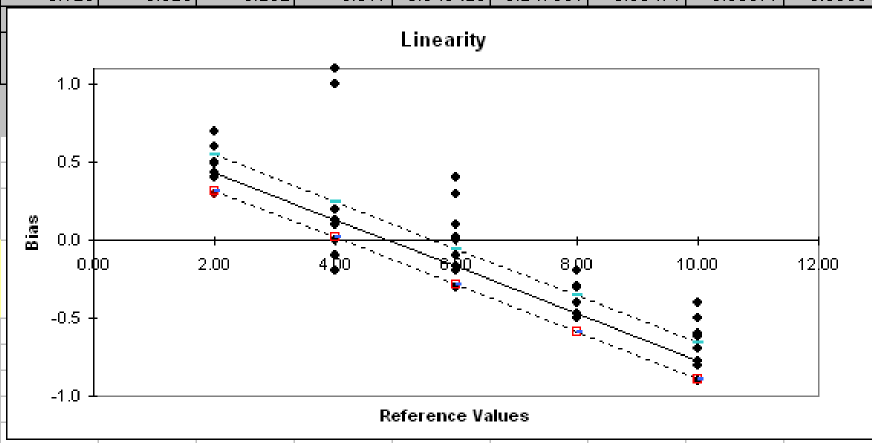
To conduct a linearity study, you will need five parts of DIFFERENT sizes that have been accurately measured to provide a reference value.

1. Have each of the five parts measured 12 times in random order.
2. Input the data into the **Gage R&R template's** Linearity worksheet
3. Input the accurate measurements for each part as a reference
4. Analyze the linearity using the line graph on the worksheet.

Ideally, there shouldn't be any change in bias from small to large. If you look at the line graph, it should be a horizontal line. More often, however, a *gage* may over read the small and under read the large. If there is too much slope to the line (too much bias), you may want to use the *gage* in its optimal range and find other *gages* to measure where this gage's bias is too large.

Here's an example of a linearity study using AIAG MSA 3rd Edition data.

	A	B	C	D	E	F	G	H	I	J	K
1				Part					Bias		
2	Trials	1	2	3	4	5	2.00	4.00	6.00	8.00	10.00
3	1	2.7	5.1	5.8	7.6	9.1	0.7	1.1	-0.2	-0.4	-0.9
4	2	2.5	3.9	5.7	7.7	9.3	0.5	-0.1	-0.3	-0.3	-0.7
5	3	2.4	4.2	5.9	7.8	9.5	0.4	0.2	-0.1	-0.2	-0.5
6	4	2.5	5	5.9	7.7	9.3	0.5	1.0	-0.1	-0.3	-0.7
7	5	2.7	3.8	6	7.8	9.4	0.7	-0.2	0.0	-0.2	-0.6
8	6	2.3	3.9	6.1	7.8	9.5	0.3	-0.1	0.1	-0.2	-0.5
9	7	2.5	3.9	6	7.8	9.5	0.5	-0.1	0.0	-0.2	-0.5
10	8	2.5	3.9	6.1	7.7	9.5	0.5	-0.1	0.1	-0.3	-0.5
11	9	2.4	3.9	6.4	7.8	9.6	0.4	-0.1	0.4	-0.2	-0.4
12	10	2.4	4	6.3	7.5	9.2	0.4	0.0	0.3	-0.5	-0.8
13	11	2.6	4.1	6	7.6	9.3	0.6	0.1	0.0	-0.4	-0.7
14	12	2.4	3.8	6.1	7.7	9.4	0.4	-0.2	0.1	-0.3	-0.6
15	Average	2.49	4.13	6.03	7.71	9.38	0.492	0.125	0.025	-0.292	-0.617
16	Reference Values	2.00	4.00	6.00	8.00	10.00	0.319771	0.017703	-0.28436	-0.58643	-0.8885
17	Bias	0.492	0.125	0.025	-0.292	-0.617	0.549428	0.247361	-0.05471	-0.35677	-0.65884
18	Range	0.40									
19	bias= slope*ref+b	0.4346									
20											
21	n=	60									
22	Slope=	-0.15103									
23	b=	0.736667									
24	Goodness of fit	0.977907									
25	Insert Process Variation	2.488105									
26	Linearity	0.375787									
27	%Linearity	15%									
28											
29											
30											



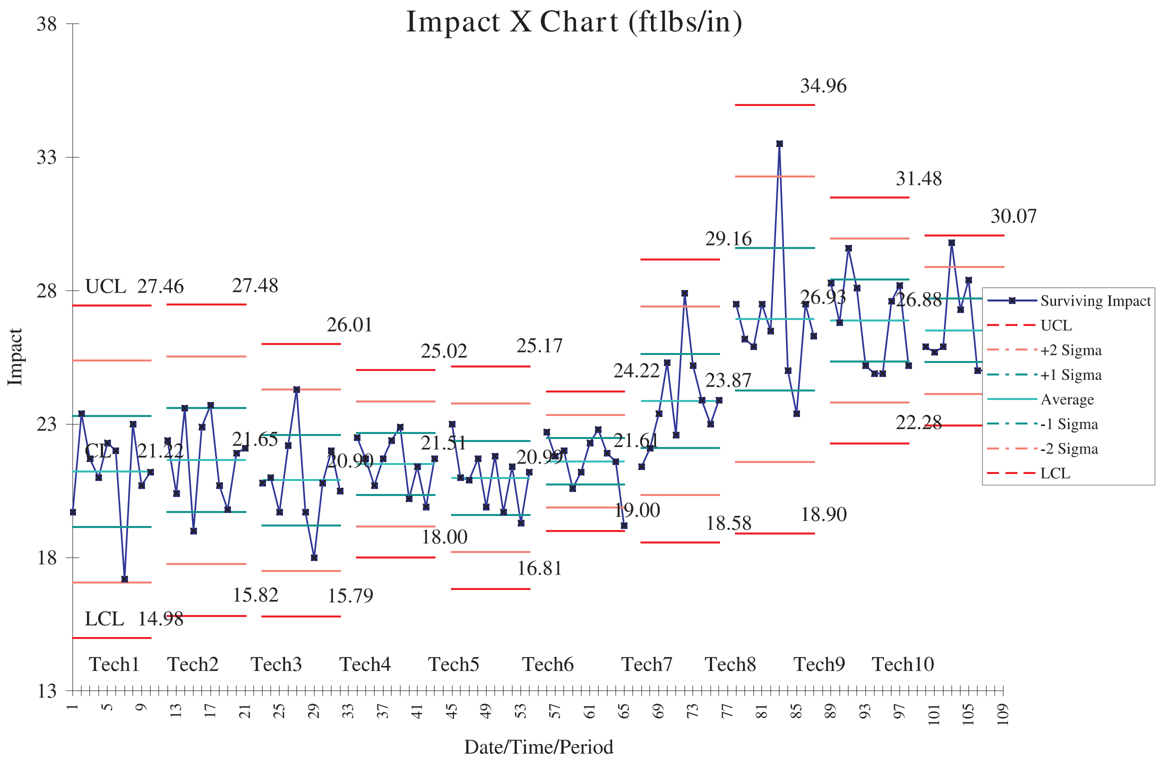
Notice how the actual values differ from the *reference values* in B15:F15. Looking at the linearity chart, you can see that the gage tends to over-read smaller measurements (2.00) and under-read the larger ones (8.00 and 10.00). This gage may only be optimal for the 4.0-to-6.0 range.

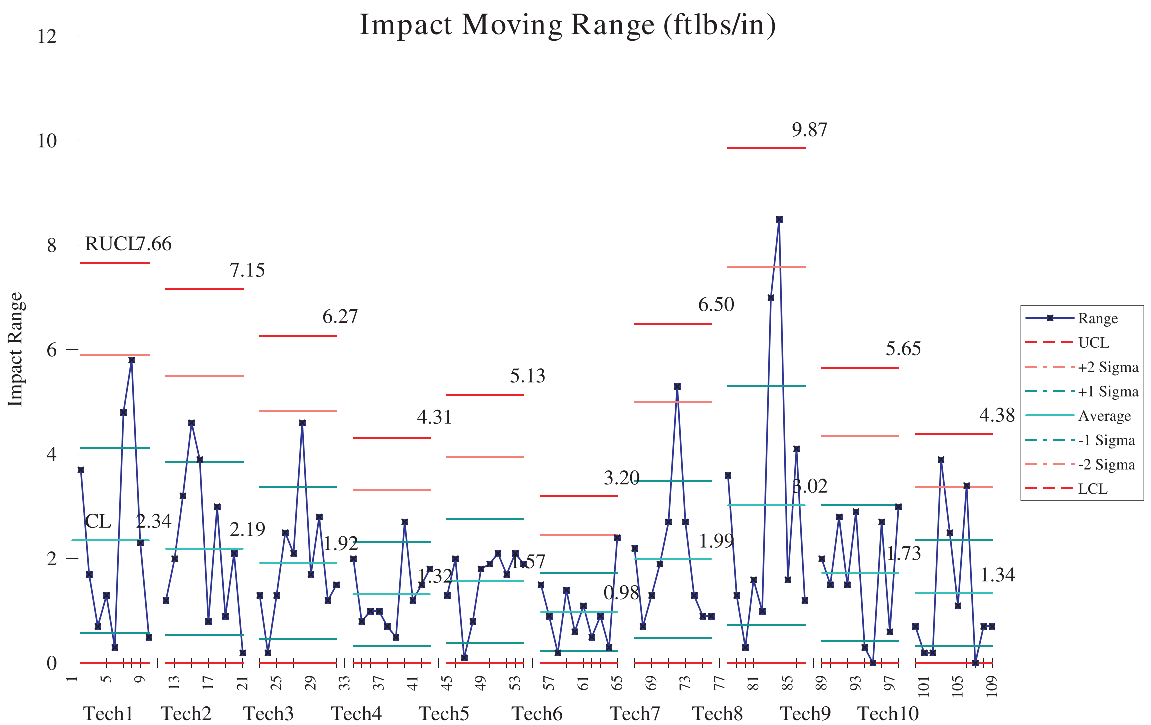
Destructive Testing

What can you do when each piece is destroyed when measured? What if you conduct an impact or other test that makes it impossible to measure repeatedly? Let's look at using an XmR chart.

In this example, there were 10 testers destructively testing 10 samples each from the same production lot. The lower specification limit is 16.7 foot pounds/sq in.

If we put a blank row between each of the 10 appraisers' data and run an XmR chart in the QI Macros we get an average and range that look like the following:





If all 10 appraisers are using the same process to measure samples from a well-known and established process, then we'd expect the averages to be the same and the variation represented by the UCL and LCL to be equal.

If the averages or standard deviations are different from appraiser to appraiser (i.e., tech to tech), then the measurement system needs improvement.

Look at the X chart. Six of the ten share roughly the same average and the last three share a different average. If they are all measuring samples from the same lot, then they must be using two different measurement processes or the production process changed. Tech 7 is between the other two groups. Which of the two main groups are measuring correctly? How can you train the others to match?

Look at the X chart and the R chart to evaluate the variation (another key factor in Gage R&R).

Technician four, six, and ten have the least variation. What are they doing that produces more consistent results? How can you train the others to match the consistency?

Ideally, all of the appraisers should get close to the same average and standard deviation. Adjust the measurement process until you get the UCL, LCL, and CL to line up more closely.

Conducting a Destructive Test Gage R&R Study

In brief, you should select your parts as follows:

The samples for Part 1 (that will be repeated by a single operator, and those to be reproduced by another operator) should come from ***as nearly a homogeneous sample/batch as possible***.

The samples used as Part 2 **should be from a different batch, different production run, or have sufficient time to allow long term variation to occur**.

Repeat for Part 3, 4, 5, etc. (see below).

	A	B	C	D	E	F	G	H	I
1	**Gage R&R**	http://www.aiag.org/					**Part Number**		http://\
2	**Average & Range Method**		1	2	3	4	5	6	7
3	Appraiser 1	Trial 1	2.31	2.12					
4	Enter your data here->	Trial2	2.29	2.15					
5		Trial3	2.34	2.16					
6		Trial4	2.27	2.14					
7		Trial 5	2.3	2.15					
8		Total	11.51	10.72					
9		Average-	2.302	2.144					
10		Range1	0.07	0.04					
11	Appraiser 2	Trial 1	2.29	2.11					
12	Enter your data here->	Trial2	2.31	2.17					
13		Trial3	2.32	2.14					
14		Trial4	2.28	2.13					
15		Trial 5	2.27	2.13					
16		Total	11.47	10.68					
17		Average-	2.294	2.136					
18		Range2	0.05	0.06					

The results (see below) will show some variation in equipment because of the variation in destructive tests.

	A	B
46	AIAG - Automotive Ind	TV
47	EV (Equipment Variati	0.0236
48	%EV	20.7%
49	AV: (Appraiser Variatic	0.00000
50	%AV	0.0%
51	R&R (Gage Capability)	0.0236
52	%R&R	20.7%
53	PV (Part Variation)	0.1117
54	%PV	97.8%
55	TV (Total Variation)	0.1142

The Gage R&R template will show that 20.7% for the %R&R means that the gage system may be acceptable (given destructive testing has occurred and we can't use the same part twice).

Attribute Gage R&R Study

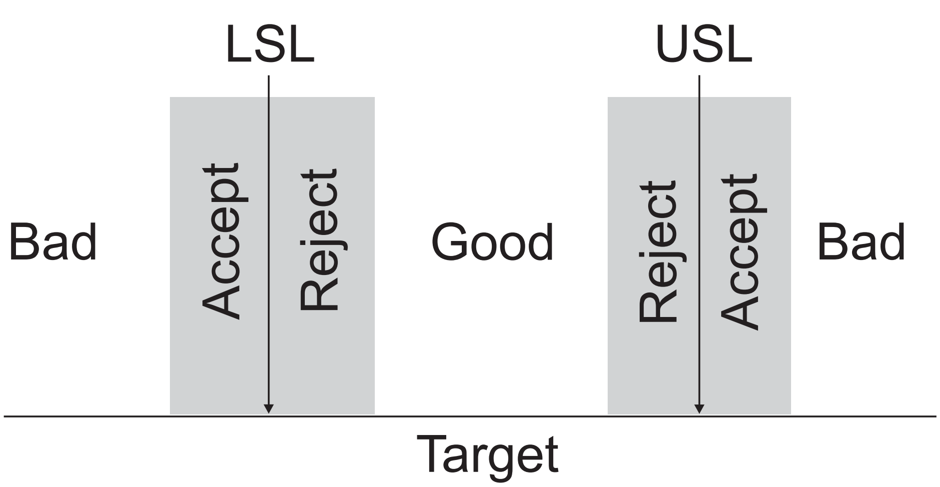
Another form of Gage R&R study is *attribute* Gage R&R. Many gages are designed for operators to quickly assess whether the part passes or fails, not the actual dimensions of the part. Imagine a hunk of metal with two slots in it: one that will tolerate the part if it's too big and one that will tolerate the part if it's too small. Operators simply take the part and slip it into the slots. If it fits either one, it's out of spec.

To conduct an attribute gage study, you need at least 10 parts. Measure these accurately with a good gage to determine the reference value. Using the QI Macros Gage R&R template (see below), input the upper and lower specification limits into cells (N3:O3) and input the reference values for each part into cells (L2:L50). Then have appraisers measure the parts randomly using the pass/fail gage. Enter a "1" for *pass* or "0" for *fail* based on each appraiser's evaluation (B3:J50).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		Appraiser A			Appraiser B			Appraiser C			Reference Results						
2	Part #	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Reference	Ref Value	Code	LSL	USL	%R&R	Gray zone
3	1	1	1	1	1	1	1	1	1	1	1	0.476901	+	0.45	0.55	25%	0.02500
4	2	0	0	0	0	0	0	0	0	0	0	0.576459	-	INSTRUCTIONS			
5	3	1	1	0	1	1	0	1	0	0	1	0.544951	X	1. Enter Spec Limits Above			
6	4										0		-	2. Have 3 appraisers accept/reject 10+ pa			
7	5										0		-	3. Enter pass/fail in columns B-J			
8	6										0		-	4. Enter Reference value in column L			
9	7										0		-	5. Evaluate effectiveness columns AJ-AV			

Bias / Linearity / **Attribute Gage Worksheet** / Analytic Attribute Method

Do they fail good parts? Do they pass bad ones?



Effectiveness and Miss Rate - Evaluate the resulting analysis (cells AT5 to AB15, see below) to find out how your gage and appraisers function. In this case, they score an 85% effectiveness and a miss rate of 15% (i.e., 15% of the time they pass a part that should fail or fail a part that should pass).

Comparison of Appraisers to Reference Value (0=fail and 1= pass) - If you look at the cross tabulations for appraisers vs. the REF value (middle tables in columns AO to AR), you'll see that A, B and C all pass parts that should fail (see red cell: AP9, AP15, AP21), but never fail parts that should pass (cells AQ8, AQ14, AQ20). In essence, the pass/fail gage is passing too high a proportion of the products compared to the reference value.

Comparison of Appraisers to Each Other - The cross tabulations between appraisers are in columns AJ to AM. For instance, cell AL14 tells you that appraiser B failed one part that appraiser C passed. These columns may provide interesting information but are not as valuable as the comparisons to the reference values in the middle columns.

	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV
5	Crosstabulation					Crosstabulation							Miss
6		B		Total			REF		Total		Effectiveness		Rate
7	A	0	1			A	0	1			A	89%	11%
8	0	4	0	4		0	3	0	3		B	89%	11%
9	1	0	5	5		1	1	5	5		C	78%	22%
10	Total	4	5	9		Total	4	5	9		System	85%	15%
11												85%	
12		C		Total			REF		Total		Effectiveness		
13	B	0	1			B	0	1			Acceptable		>= 90%
14	0	4	1	5		0	3	0	3		Marginal		>=80%
15	1	0	4	4		1	1	5	6		Unacceptable		<80%
16	Total	4	5	9		Total	4	5	9				
17													
18		A		Total			REF		Total				
19	C	0	1			C	0	1					
20	0	4	0	4		0	3	0	3				
21	1	1	4	5		1	2	4	6				
22	Total	5	4	9		Total	5	4	9				

Bias / Linearity / Attribute Gage Worksheet / Analytic Attribute Method

Since, the pass/fail gage is passing too high a proportion of the products compared to the reference value, ask yourself: Do you need to repair, replace or upgrade the gage? Do you need some additional training for the appraisers? Make the changes and repeat the study until you get the level of quality you desire.

Conclusion

So, there, in a nutshell, is Gage Repeatability, Reproducibility, Bias, and Linearity.

Your goal is to minimize the amount of variation and error introduced by measurement, so that you can focus on part variation. This, of course, leads you back into the other root causes of variation: process, machines, and materials.

If you manufacture anything, measurement system analysis can help you improve the quality of your products, get more business from big customers, and baffle your competition. Enjoy.

Resources

For more detailed information on the Gage R&R, consider the AIAG Measurement System Analysis book. You can find one at http://www.aiag.org.

The QI Macros Gage R&R template uses formulas from the most recent AIAG MSA edition. The Gage R&R template is made up of several different templates including Average and Range Method, Anova method, Bias, Linearity, and Attribute Method.

The QI Macros also draws line, pie, bar, pareto, histogram, scatter and control charts. It contains over 80 fill in the blank templates such as the Ishikawa diagram, QFD, DOE, FMEA, EMEA, and AIAG PPAP templates. You can order a copy of the QI Macros at http://www.qimacros.com/products.php