Repeatability & Reproducibility Studies

Introduction

Before we can talk about “gage R&R,” we have to define the word “gage.” When asked to name a gage, people typically think of micrometers, pressure gages, temperature gages, etc. However, the term “gage” actually refers to any device used for making measurements. In this document, the terms “gage” and “device” are used interchangeably and refer to any device or equipment for making a measurement.

Just what is a “gage R&R study, anyway?” Gage R&R is a study of the measurement variation of a gage and the variation of measurements of operators. To understand why this is important, recall that the goal is to constantly improve process control, reducing the variability in the process and product. In order to address actual process variability, the variation due to measurement system must be identified and separated from that process. Studies of measurement variation are a waste of time and money unless they lead to action to reduce process variation and improve process control. Since you cannot address something that cannot be measured precisely, you must start with an assessment of the gage R&R.

Every observation contains both actual process variation and measurement variation (Figure 1). In the case of measurement systems, the sources are:

a) Gage,
b) Operator
c) Variation within the sample.

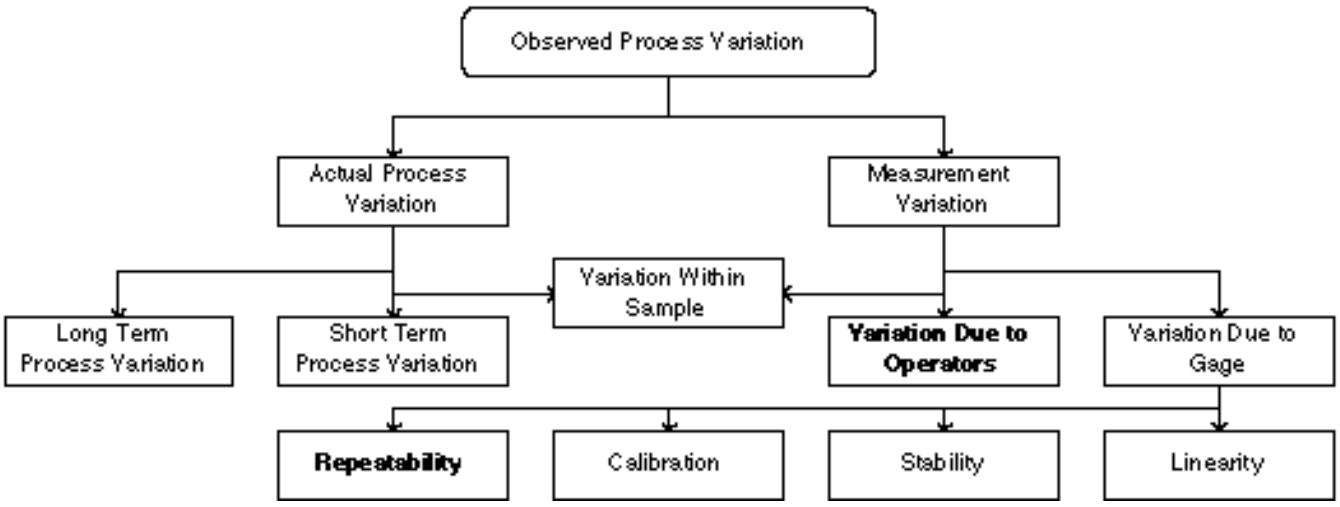

Figure 1

Gage variability can be broken down into additional components, such as:

a) Calibration (is the gage accurate?)
b) Stability (does the gage change over time?)
c) Repeatability (variation of the gage when used by one operator in a brief interval), and
d) Linearity (is the gage more accurate when used at low values than at high values?)

Variation within a sample is included in process variation; yet, it is also often mixed with measurement variation.

Gage Repeatability and Reproducibility (R&R) studies are defined as studies of reproducibility (operator variation) and repeatability. Repeatability is the variation observed when an operator measures the same sample with the same gage several times. Reproducibility is the additional variation observed when different operators use the same gage to measure the same sample. The combination of both sources of variation is referred to as R&R (see Figure 2). The exclusion of other potential sources of measurement variation does not imply that calibration, stability or linearity are unimportant; it is just the impact of that those sources are ordinarily less significant. For that reason, R&R is usually studied and quantified first. Repeatability and reproducibility are actually key output variables (KOV's) of the measurement system. In order to improve them, the key input variables (KIV's) must be addressed via procedures, standards, training and appropriate studies.

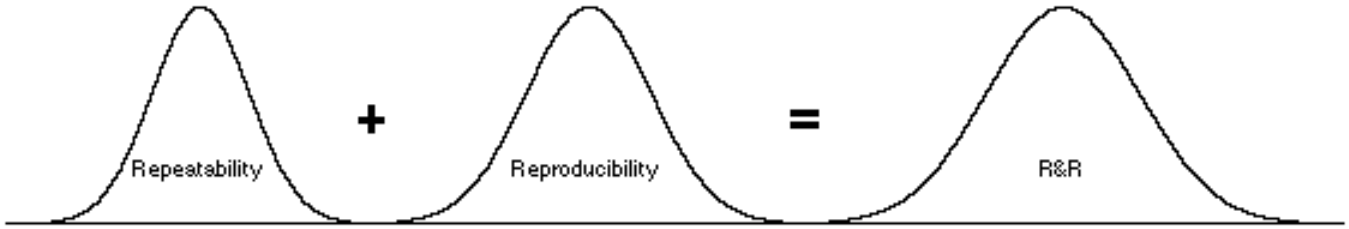

Figure 2

R&R studies are planned and executed in a fashion to avoid confusion with other sources of variation. The other sources of variation cannot be ignored. In particular, actual process variation is the ultimate subject to be addressed. Customers may require both R&R studies and process capability. All significant sources of variation must be addressed in order to bring processes and products under better control. Process capability studies include both process variation and measurement variation. Consequently, R&R studies should be accompanied or quickly followed by evaluations of calibration, variation within sample and, any other relevant source of variability.

Variation within a sample being measured is difficult to exclude from an R&R study. This source is extremely important and should always be pursued with diligence. It not only has relevance to understanding R&R's, but it also provides vital information on how to gain process capability improvement. A specific example of variation within a sample is apparent when measuring surface texture with a profilometer. The test piece itself is sufficiently

variable that if the measurement is made at a random position, the variation within the sample will inflate the estimate of repeatability. It is necessary to identify and measure this variability within the sample; this is not the role of an R&R study alone. This will discussed later in more depth. The key point is to make certain that process variability within the sample does not intrude on an R&R study if it can be avoided. Determination of an unsatisfactory R&R should always lead to an evaluation of whether variation within the sample is part of the problem.

In the ideal case, all variability in measurements will be due to the part-to-part variation, and only a negligible proportion of the variability will be due to operator reproducibility and trial-to-trial repeatability. You can compute the standard indices of repeatability, reproducibility, and part-to-part variation, based either on ranges (as is most common) or from an analysis of variance (ANOVA). The ANOVA table will also contain an F-test (statistical significance test) for the operator-by-part interaction, and report the estimated variances, standard deviations, and confidence intervals for the components of the ANOVA model. Finally, you can compute the respective percentages of total variation, and report so-called percent-of-tolerance statistics.

Procedure for Performing and R&R Study

The person conducting the study has several responsibilities to ensure a successful R&R. The first is to obtain proper samples. They should all be of a similar product line. Select actual parts from production. The selected parts should cover the entire tolerance range. Each part should be labeled with a unique identifier to facilitate data collection. The samples should be masked so that the persons involved in the test do not know which sample labels are which. Typically, when an R&R is done, the labels on the samples are changed after each reading or day. If the person conducting the test does not maintain 100% control and traceability of the samples throughout the test, it is invalidated. All persons participating in the test should know that the labels are being changed.

Review the gage instruction method and verify it is the correct one. Prepare the appraisers for data collection. Review the inspection method with the appraisers. Explain the purpose of the study, the method of data collection and, the role of the appraisers.

Measure each part in random order and record the readings on a data sheet. Appraisers should take the normal amount of time to take each measurement. Measure each part in a random order and record the readings on another data sheet. It is important to keep the readings separate so appraisers are not biased by the previous readings.

Repeat taking measurements on each part, one at a time, until the desired number of readings per part are obtained.

Terms Used in R&R

n = Number of Parts [2 to 10]
a = Number of Appraisers
r = Number of Trials
g = r * a
$\bar{R}_A$ = Average for appraiser A, etc.
$\bar{\bar{R}}$ = Average of $\bar{R}_A$, $\bar{R}_B$, etc.
R_p = Range of Average Parts
$\bar{X}_{diff}$ = Difference Between High & Low Appraiser Averages
σ_{ev} = Equipment Variation (Repeatability)
σ_{ov} = Operator Variation (Reproducibility)
σ_s = Sample Variation
σ_p = Part Variation

Data Analysis

An example of data from an R&R study is presented below. Suppose an R&R study is conducted with two persons measuring parts. The specification is 1.000 ± 0.010. Readings are recorded to the nearest 0.001 inch. The measurements obtained are presented in Table 1.

Appraiser 1	Part	1	2	3	4	5	6	7	8	9	10
Trial											
1		1.004	1.005	1.002	1.002	1.004	1.003	1.007	1.000	0.999	0.998
2		1.004	1.005	1.001	1.000	1.004	1.003	1.007	1.000	0.999	0.998
3		1.005	1.005	1.003	1.002	1.004	1.002	1.007	1.001	0.999	0.998
Range		0.001	0.000	0.002	0.002	0.000	0.001	0.000	0.001	0.000	0.000
Appraiser 2											
Trial											
1		1.004	1.005	1.001	1.001	1.004	1.002	1.006	1.000	1.000	0.998
2		1.004	1.006	1.002	1.002	1.004	1.002	1.004	1.000	0.999	0.997
3		1.003	1.005	1.001	1.002	1.004	1.002	1.005	1.000	0.999	0.997
Range		0.001	0.001	0.001	0.001	0.000	0.000	0.002	0.000	0.001	0.001

R_{barA}	0.0007
R_{barB}	0.0008

Table 1[1]

The average ranges for each appraiser are calculated and the upper control limit (UCLR) for the range is calculated $UCLR = \bar{\bar{R}} * D_4$

[1] *Total Quality Management Handbook*, Jack Hradesky, McGraw Hill, 1995

Where the number of trials (r) determines D_4 from table 2.

r	D_4
2	3.267
3	2.574
4	2.282
5	2.114
6	2.004

Table 2

For the example above, $UCRL = 2.574 * 0.00075 = 0.0019$

If any ranges exceed the UCLR, do the following:

Remeasure the part to determine if the out-of-control range was due to recording or measurement error.

If data recording error is suspected, remove the affected measurement from the computations and recalculate the average range and UCLR. When more than one data recording error is found, correct the measurement process and repeat the study.

If there is more than one range exceeding the UCLR due to a measurement error, the method must be revised. No further computations are needed, and the study should be repeated when the method is improved.

When only one range exceeds the UCLR due to measurement error, remove the affected data from the computations and recompute the average range and UCLR.

In the example, three of the ranges (0.002) exceed the UCLR of 0.0019. Due to the measuring equipment increment being 0.001, these are not considered measurement or data recording errors and are not removed.

Repeatability Evaluation

Repeatability is found by computing the Standard Deviation of Repeatability (σ_{ev}). The σ_{ev} is found by $\sigma_{ev} = \frac{1}{d_2} * \bar{\bar{R}}$ where d_2 (found from table 3) with r equal to the number of trials.

r	d_2
2	1.128
3	1.693
4	2.059
5	2.326

Table 3

In the example, $\frac{1}{d_2} = 0.591$ and $\sigma_{ev} = 0.591 * 0.00075 = 0.00044$

The Percent Tolerance Consumed by Repeatability (PTCR) is found by multiplying by 5.15 (based on an interval that contains 99% of the distribution) and dividing by the tolerance spread:

$$PTCR = \left(\frac{5.15 * \sigma_{ev}}{TotalTolerance}\right) * 100$$

$$PTCR = \left(\frac{5.15 * 0.00044}{0.020}\right) * 100 = 11.33\%$$

Reproducibility Evaluation

The reproducibility is found by computing the Standard Deviation of Reproducibility (σ_{ov}). The σ_{ov} is found by $\sigma_{ov} = (\bar{X}_L - \bar{X}_S) * D$

Where D is a factor taken from table 4 and a equal the number of appraisers.

a	D
2	0.709
3	0.524
4	0.446
5	0.403
6	0.375
7	0.353
8	0.338
9	0.325
10	0.314

Table 4

Using the example data $\sigma_{ov} = (1.0024 - 1.0020) * 0.709 = 0.00028$

The Percent Tolerance Consumed by Measurements (PTCM) is found by multiplying by 5.15 (based on an interval that contains 99% of the distribution) and dividing by the tolerance:

$$PTCM = \left(\frac{5.15 * \sigma_{ov}}{TotalTolerance}\right) * 100$$

$$PTCM = \left(\frac{5.15 * 0.00028}{0.020}\right) * 100 = 7.3\%$$

Inspection Capability Evaluation

The Percent Tolerance Consumed by inspection Capability (PTCC) is calculated by:

$$\sigma_c = \sqrt{\sigma_{ev}^2 + \sigma_{ov}^2}$$

$$\sigma_c = \sqrt{0.00044^2 + 0.00028^2} = 0.00052$$

$$PTCC = \frac{5.15 * \sigma_c}{Total\ Tolerance} * 100$$

$$PTCC = \frac{5.15 * 0.00052}{0.020} * 100 = 13.39\%$$

Data Evaluation

Results of the study are evaluated to determine if the measurement method is acceptable, marginal or unacceptable using the following criteria:

PTCC	Study Result
< 10%	Acceptable
10% and 25%	Marginal
> 25%	Unacceptable

Table 5

If the measurement method is marginal or unacceptable, corrective action is required. The study should be repeated when the corrective action is implemented and completed. In the example, the inspection capability is marginal, since the PTCC is between 10% and 25%.

The Range Method

The range method is a modified variable gage study which provides a quick approximation of measurement variability. This method will only provide the overall picture of the measurement system. It does not decompose the variability into repeatability and reproducibility.

The range method uses two operators and five parts for the study. In this study, both operators measure each part once. The range for each part is the absolute value of the difference between the measurements obtained by operator A by operator B. The sum of the ranges is found and the average range ($\bar{R}$) is calculated. The total measurement variability is found by multiplying $\bar{R}$ by 4.33. It is interesting to know the percentage of the process variation (or tolerance) that measurement variation consumes. To convert the R&R into a

percentage, multiply by 100 and divide by the process variation (or tolerance). In the example below, the tolerance for this characteristic is ± 0.20.

Parts	Operator A	Operator B	Range (A–B)
1	0.85	0.80	0.05
2	0.75	0.70	0.05
3	1.00	0.95	0.05
4	0.45	0.55	0.10
5	0.50	0.60	0.10

Table 6

$$\bar{R} = \frac{\sum R_i}{5} = \frac{0.35}{5} = 0.07$$

$$R\&R = \bar{R} * 4.33 = (0.07)(4.33) = 0.303$$

$$Process\ Variation = 0.40\ or\ Tolerance$$

$$\%R\&R = 100 * \frac{R\&R}{Process\ Variation} = 100 * \frac{0.303}{0.400} = 75.5\%$$

Now that the %R&R for the measurement system is determined, an interpretation of the results should be made. Using the guidelines above, this measurement system is in need of improvement.

Average and Range Method

The average and range method is a mathematical method which determines both repeatability and reproducibility for a measurement system. Unlike the range method, this allows the measurement system to be decomposed into separate components, repeatability and reproducibility. Using the example data first presented above, a presentation of the Average and Range method worksheet is presented in figure 3 below.

	1	2	3	4	5	6	7	8	9	10	11	12
	Appraiser 1				Appraiser 2				Appraiser 3			
Part	1st Test	2nd Test	3rd Test	Range	1st Test	2nd Test	3rd Test	Range	1st Test	2nd Test	3rd Test	Range
1	1.004	1.005	1.005	0.001	1.004	1.004	1.003	0.001				
2	1.005	1.005	1.005	0.000	1.005	1.006	1.005	0.001				
3	1.002	1.001	1.003	0.002	1.001	1.002	1.001	0.001				
4	1.002	1.000	1.002	0.002	1.001	1.002	1.002	0.001				
5	1.004	1.004	1.004	0.000	1.004	1.004	1.004	0.000				
6	1.003	1.003	1.002	0.001	1.002	1.002	1.002	0.000				
7	1.007	1.007	1.007	0.000	1.006	1.004	1.005	0.002				
8	1.000	1.000	1.001	0.001	1.000	1.000	1.000	0.000				
9	0.999	0.999	0.999	0.000	1.000	0.999	0.999	0.001				
10	0.998	0.998	0.998	0.000	0.998	0.997	0.997	0.001				
Totals	10.024	10.022	10.026		10.021	10.020	10.018					
		20.046		0.0007		20.041		0.0008				
		10.026		R–Bar 1		30.059		R–Bar 2				R–Bar 3
	Sum	30.072			Sum	30.059			Sum			
	X–Bar 1	1.0024			X–Bar 2	1.0020			X–Bar 3			

R–Bar 1	0.00070
R–Bar 2	0.00080
R–Bar 3	
Sum	0.00150
R–Bar All	0.00075

# Trials	D 4
2	3.27
3	2.58

R Bar All x D 4 = UCL r

0.00075 x 2.58 = 0.00194

Max X–Bar	1.0024
Min X–Bar	1.0020
X–Bar Diff	0.0004

Measurement Unit Analysis **Percent Tolerance Analysis**

Tolerance = 0.010

Repeatability - Equipment Variation (EV)

%EV = 100(EV/Tolerance)

EV = Rdbar x K1

Trials	2	3	4
K1	4.56	3.05	2.50

%EV = 11.44%

EV = 0.00229

K1 = 3.05

σEV = EV/5.15 = 0.00044

Reproducibility - Operator Variation (OV)

%OV = 100(OV/Tolerance)

OV = SQRT((X-Bar Diff x K2)^2 - [(EV)^2/(n*r)])

Operators	2	3	4
K2	3.65	2.70	2.30

%OV = 7.63%

OV = 0.00153

K2 = 3.65

n = 10

r = 3

σOV = OV/5.15 = 0.00030

Repeatability and Reproducibility (R&R)

%R&R = 100(R&R/Tolerance)

R&R = SQRT(EV^2 + OV^2)

%R&R = 13.75%

R&R = 0.00275

σR&R = R&R/5.15 = 0.00053

Figure 3

Prepared By:

There are minor differences in the results using the Average and Range method. They are due to various errors introduced in the derivations of the different factors involved. When there is an R&R study that yields unsatisfactory results, advanced methods can be employed to determine exactly which area to improve upon. Tests such as Analysis of Variance (ANOVA), t-tests and others will help to isolate the problem. The ANOVA test is particularly useful when a single factor does not indicate where improvement is needed, but can be used to determine if there are interactions between factors.

Derivation of the K_1 Factor

To understand the methodology for determining the proper K_1 factors, we first need to understand what is going on. On the R&R worksheet presented in figure 3, the range of the trials for each operator–sample is calculated. These are averaged to get the average range for each operator ($\bar{R}_A$,$\bar{R}_B$, etc.). The grand average, ($\bar{\bar{R}}$), is the average for all operations.

These ranges of measurements on the same sample by the same operator are used to ascertain the standard deviation of repeatability (σ_{ev}). From quality control statistics, recall that ranges and standard deviations are related by the formula:

$$\sigma_{ev} = \frac{\bar{\bar{R}}}{d_2}$$

Where d_2 varies with the sample size (number of trials) used to calculate a single range. Repeatability (ev) as defined above as 99% of the spread of 5.15 σ_{ev}.

$$Repeatability = 5.15\sigma_{ev} = 5.15\left(\frac{\bar{\bar{R}}}{d_2}\right) = K_1 * \bar{\bar{R}} \quad where \; K_1 = \frac{5.15}{d_2}$$

For example, for a study with three trials, the sample size for each range is three. The tabular value for d_2 for a sample size of three is $d_2 = 1.69$.

$$Repeatability = 5.15\sigma_{ev} = 5.15\left(\frac{\bar{\bar{R}}}{d_2}\right) = 5.15\left(\frac{\bar{\bar{R}}}{1.69}\right) = 3.05 * \bar{\bar{R}}$$

This is the source of the K_1 factors presented in figure 3 for a given number of trials.

It has been shown that his estimation procedure must be corrected when there are a small number of ranges. Since there is a range calculated for each operator–sample combination, the product of (# operators) x (# samples) is the number of ranges. If there are more than 15 such ranges, the usual estimation procedure using d_2 is appropriate. In such a case, the values of K_1 from the form in figure 3 are correct.

If (# operators) x (# samples) is less than 16, then the estimation procedure should be adjusted. This is done by using a parameter d_2* rather than d_2 and:

$$K_1 = \frac{5.15}{d_2 *}$$

The values of K_1 using the factor d_2* are presented in table 7. The particular value depends on the number of ranges calculated (# operators) x (# samples) and the number of trials. For example, if three operators are used with four samples for three trials, then (# operators) x (# samples) =12. Since the three trials provide a subgroup size of three, the referenced table for (# operators) x (# samples) =12 and r = 3 provides K_1 of 3.01.

K_1 Factors When (# operators) x (# samples) is ≤ 15

(# operators) x (# samples)

# of Trials	3	4	5	6	7	8	9	10	11	12	13	14	15	≥ 16
2	4.19	4.26	4.33	4.36	4.40	4.40	4.44	4.44	4.44	4.48	4.48	4.48	4.48	4.56
3	2.91	2.94	2.96	2.98	2.98	2.99	2.99	2.99	3.01	3.01	3.01	3.01	3.01	3.05
4	2.43	2.44	2.45	2.46	2.46	2.48	2.48	2.48	2.48	2.49	2.49	2.49	2.49	2.50
5	2.16	2.17	2.18	2.19	2.19	2.19	2.20	2.20	2.20	2.20	2.20	2.20	2.20	2.21
6	2.00	2.00	2.01	2.01	2.02	2.02	2.02	2.02	2.02	2.02	2.02	2.03	2.03	2.04

Notes:

1. If possible, select the number of operators and samples for (# operators) x (# samples) to exceed 15.
2. If (# operators) x (# samples) is ≤ 15, enough trials should be run to avoid the shaded area of the table. (Otherwise, be aware the estimates may be imprecise)

Table 7

Table of d_2* For Computing K_2

g

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1.41	1.28	1.23	1.21	1.19	1.18	1.17	1.17	1.16	1.16	1.16	1.15	1.15	1.15	1.15
3	1.91	1.81	1.77	1.75	1.74	1.73	1.73	1.72	1.72	1.72	1.71	1.71	1.71	1.71	1.71
4	2.24	2.15	2.12	2.11	2.10	2.09	2.09	2.08	2.08	2.08	2.08	2.07	2.07	2.07	2.07
5	2.48	2.40	2.38	2.37	2.36	2.35	2.35	2.35	2.34	2.34	2.34	2.34	2.34	2.34	2.34
6	2.67	2.60	2.58	2.57	2.56	2.56	2.55	2.55	2.55	2.55	2.55	2.55	2.55	2.54	2.54

Table 8[2]

The table of adjusted K_1 factors contains a recommendation for the minimum number of trials in an R&R study. A study is not limited to that minimum. However, the recommended number of trials for a given number of ranges (# operators) x (# samples) assures that the

[2] *Quality Control and Industrial Statistics*, A. J. Duncan, Richard D. Irwin, Inc., 4th Ed., 1974

estimates of σ_{ev} are all based on at least 14 degrees of freedom, the minimum number for the case where (# operators) x (# samples) is greater than 15 and there are two trials.

When the combinations of number of operators, trials and samples fall into the shaded area of the table, the level of confidence in the resulting estimates becomes poor. These conditions should be avoided when possible.

This procedure is also valid when there is a single operator (if there is no operator, assume a single operator). In that case, the rules concerning (# operators) x (# samples) become simply (# samples) greater than 15. If the number of samples is 15 or less, then the number of trials must be increased accordingly.

Previous discussions indicated that the parts used in the study should cover the range of measurements expected. However, there is also a requirement that the range of measurements used in the study be defined such that the average range on each part is expected to estimate a common σ_{ev}. If this "homogeneity" is not met, the results become erratic.

Derivation of the K_2 Factor

The same rationale that was applied to obtain the K_1 factors is also applied to obtain the K_2 factors on the form. The K_2 factor is used to ascertain the standard deviation of the operator's averages, leading to the estimation of reproducibility. The ranges of the operator's averages is labeled $\bar{X}_{diff}$. Again:

$$\sigma_{ov} = \frac{R}{d_2} = \frac{\bar{X}_{diff}}{d_2}$$

Such a small number of ranges (one!) forces the use of d_2^* instead of d_2. Then:

$$\sigma_{ov} = \frac{\bar{X}_{diff}}{d_2^*}$$

Table 8 is used to find the proper value for d_2^* for g = 1 and m = # of operators. Reproducibility is defined as a 5.15σ range. Then, if the same basis is used for the σ of operator's averages, then:

$$5.15\sigma_{ov} = \frac{5.15\bar{X}_{diff}}{d_2^*} = K_2 * \bar{X}_{diff} \text{ where } K_2 = \frac{5.15}{d_2^*}$$

For example, for four operators, table 8 provides g=1 and m = 4 a value of 2.24. Then:

$$K_2 = \frac{5.15}{d_2^*} = \frac{5.15}{2.24} = 2.30$$

The other values of K_2 are estimated similarly.

This procedure provides degrees of freedom = (# operators - 1). Thus, if there are two operators, there is only one degree of freedom for the estimate. This is why it is recommended to use three or four operators, if practical.

Note if there is a single operator, it is not possible to calculate the K_2 factor.

Derivation of the Calculation of Reproducibility

The derivation of the calculation of reproducibility was covered earlier with the origins of the K2 factor used. The origin of the formula of reproducibility (ov) is presented here. Since reproducibility may also be referred to as operator variation, the notation “ov” is used.

The entire procedure of an R&R study is based on Analysis of Variance (ANOVA). It helps for this section if a knowledge of ANOVA is known. If not, it is recommended that the concept be studied further for a better understanding.

AN R&R study is a designed experiment where the factors are: 1) samples; 2) operators. The trials are nested within operators and samples. It is assumed that:

1. The samples are random representatives from a large population of possible samples.
2. The operators are random samples of a population of possible operators. (This assumption is debatable if there are only a few operators. This along with the low degrees of freedom, is why reproducibility is estimated with much less confidence than repeatability).
3. The interaction of operators and samples does not exist. The calculation of the Upper Control Limit of the ranges for the R&R is a weak attempt to confirm this assumption. Interaction could also be a result from too broad a range for the sample.
4. The study has been performed in such a manner, using random order of tests, that the pooled variation of the trials for each operator–sample combination is an estimate of the error mean square in the ANOVA.

ANOVA Table

Source	Expected Mean Square
Samples	$\sigma_{ev}^2 + (o*r)\sigma_s^2$
Operators	$\sigma_{ev}^2 + (r*n)\sigma_{av}^2$
Error	σ_{ev}^2

Where n = # samples; o = # operators; r = # trials

Table 9

There should be no interest in the variability due to samples; in fact, samples should be selected to cover the range of possible materials (covered earlier).

$$\sigma_{ev}^2 = \left(\frac{\bar{R}}{d_2}\right)^2$$

The variance in operator's averages, however, is given by

$$\left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2$$

Since each operator's average is the average of (r x n) measurements, then using ANOVA tables expected mean square.

$$\sigma^2_{OperatorsAverages} = \left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2 = \frac{\sigma_{ev}^2 + (r * n * \sigma_{ov}^2)}{r * n}$$

Although this all looks formidable, it is a restatement of the rule: if the variance of one observation is σ^2, the variance of an average of *m* observations is σ^2/m. Thus:

$$\sigma^2_{OperatorsAverages} = \sigma_{ov}^2 + \frac{\sigma_{ev}^2}{r * n} = \left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2$$

and

$$\sigma_{ov}^2 = \left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2 - \frac{\sigma_{ev}^2}{r * n} = \left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2 - \frac{\left(\frac{\bar{R}}{d_2^*}\right)^2}{r * n}$$

This is what is happening in the calculations of the R&R form. The only difference is that the form is working with 5.15 x sigma, thus using K_1 and K_2. That results in:

$$ev = 5.15\sigma_{ev} = K_1 R \qquad ov^2 = 5.15^2\sigma_{ov}^2 = 5.15^2\left(\frac{\bar{X}_{diff}}{d_2^*}\right)^2 - \frac{\sigma_{ev}^2}{r * n}$$

$$ev^2 = 5.15^2\sigma_{ev}^2 \qquad ov^2 = \left(K_2 * \bar{X}_{diff}\right)^2 - \frac{ev^2}{r * n}$$

Then:

$$ov = \sqrt{\left(K_2 \bar{X}_{diff}\right)^2 - \frac{ev^2}{r * n}}$$

For a discussion of the respective roles and advantages of ANOVA versus $\bar{X}$ –R charts, refer to pages 683–684 of Duncan's[3] text.

How to Address Variation within A Sample

There are some measurements where variation within the sample cannot be prevented from having an effect on the R&R study. For example, the measurement of "out-of-roundness" of a bar. There is variability in this measurement depending upon where on the sample bar the measurement is made. Another example is the surface roughness of a sample that may vary significantly across the sample. Still another is a destructive measurement where the same test cannot be reused.

The variation within the sample is extremely important and can be a major component of process capability. Neither the measurement variation nor the process variation can be understood without defining the variation within the sample. The issue is the order of attack.

In order of priorities:

1. There must be a calibration procedure in place.
2. The measurement capability (R&R) must be defined, excluding variation within the sample.
3. The variation within the sample must be defined.
4. Move on to process control.

There are two situations that must be considered:

Case A: Either what to expect is not known or it is logical to assume no significant variation within the sample.

In such instances, perform an R&R in the routine fashion, making no effort to avoid the effect of variation within the sample. This would mean that the operator is not instructed to test or retest at the same location.

1. If the R&R calculations for repeatability are satisfactory, any variation within the sample is either insignificant or not a pressing concern.
2. If the result of the calculations of repeatability is poor or marginal, the existence of significant variation within the sample must be addressed and defined. The R&R study will have provided an estimate of σ_{ev}. If there is significant variation within the sample, this estimate is actually the estimate of:

$$\sigma_{ev} = \sqrt{\sigma_e^2 + \sigma_{pv}^2}$$

[3] *Quality Control and Industrial Statistics*, A. J. Duncan, Richard D. Irwin, Inc., 4th Ed., 1974

Where σ_e estimates the actual sigma of repeatability and σ_{pv} estimates the sigma of variation within the sample. There is a need to separate and estimate each component. It cannot be done as the study was performed. A second study must be performed to estimate σ_e allowing an estimation of σ_{pv}.

How is an estimate of σ_e done without the effect of σ_{pv}? Repeating an R&R study either with special samples known to contain small within sample variation or specify exactly where on the sample the repeated measures are to be made does this. Examples would be using standard test blocks in testing a hardness tester or specifying the exact point on a bar to test for out-of-roundness. This time $\sigma_e = \sigma_{ev}$ of the second study

Using this value, the prior equation can be used to solve for σ_{pv}. Assume in an R&R study of a profilometer, $\sigma_{ev} = 2.33$, an undesirable result. The design permitted the variation within the sample to be mixed with that of repeatability. Then $\sigma_{ev} = 2.33 = \sqrt{\sigma_e^2 + \sigma_{pv}^2}$

Assume the study is rerun, this time the point of measurement is controlled. The new estimate is $\sigma_e = 1.21 = \sigma_{ev}$ of the second study. Now:

$$2.33 = \sqrt{\sigma_e^2 + \sigma_{pv}^2} = \sqrt{1.21^2 + \sigma_{pv}^2}$$

$$\sigma_{pv}^2 = 3.97$$

$$\sigma_p = 1.99$$

An estimate of σ_{ev} has now been found and does not include σ_{pv}. An estimate of σ_{pv} has also been obtained. We could continue to measure the R&R results based on the clean estimate of σ_{ev}. This indicates the ability to measure. Note the actual, routine measurements will be more variable than indicated by the R&R if (the variation within the sample) is a significant contributor to the variation. If measurements are made at a randomly selected location on the samples, that measurement variation contains the effect of both within sample variation and repeatability of the measurement device.

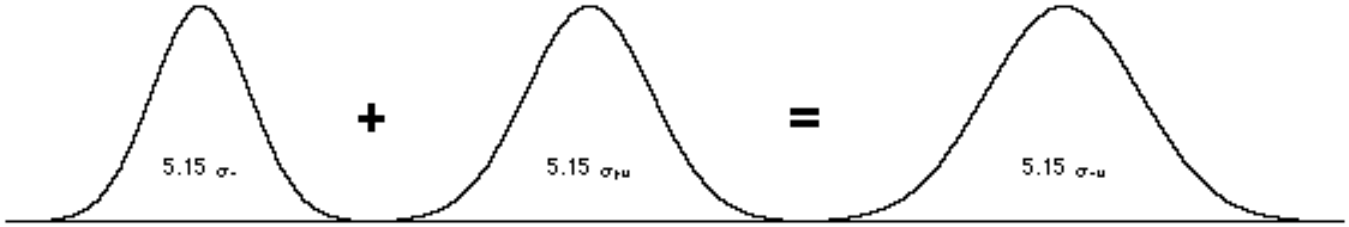

Figure 4

If the measurements are now made at the same location on each sample, then the measurement variation does not contain the effect of variation within the sample. In that case, the $\sigma_{ev} = \sigma_{ev}$ and σ_{pv} is not measured at all. Note in this example, the estimation of repeatability is seriously influenced by whether variation within the sample is present.

Results of variation within the sample are important, providing vital information on the process control characteristics to be addressed. What is to be done about these results? The fact that a clean estimate of R&R is now available does not mean the variation within the sample can be ignored. That information must be brought forward to stimulate process improvement. There are ways to indicate the significance of variation within the sample.

1. Calculate percent variation within sample $\dfrac{5.15\sigma_{pv}}{tolerance}$

2. Calculate percent R&R and variation within the sample $\dfrac{5.15\sqrt{\sigma_e^2 + \sigma_{pv}^2 + \sigma_{ov}^2}}{tolerance}$

This allows the relative importance of the variation within the sample to be described and compared with PTCR, PTCM and PTCC.

Case B: It is known that significant variation exists within the sample.

The only difference between this situation and the previous case is the order of the successive R&R studies.

1. If it is already known there is significant variation within the sample, first perform the R&R study using whatever procedure will eliminate or avoid variation within the sample. This could mean using standards, special samples and/or carefully identifying and controlling the point of measurement. The resulting R&R study will estimate σ_{ev}, an estimate that is free of within sample variation. This study will provide the official R&R results.

2. Next, rerun an R&R on production samples or in a fashion that will produce the combined estimate of:

$$\sigma_{ev} = \sqrt{\sigma_e^2 + \sigma_{pv}^2}$$

Use the two sets of results to solve as before for σ_{pv}. This second study will provide the estimate for variation within the sample.

These final comments are offered:

1. R&R studies are limited, by definition, to the effects of operators and measurement device.

2. Indications or expectations of variation within the sample must be followed up with an appropriate study. Otherwise, knowledge of the process is incomplete and may mislead efforts to achieve improvement. To understand the measurement system, this component of variation must be measured and studied.

3. This discussion has touched briefly on the mechanics, striving to convey the logic and philosophy for dealing with variation within a sample. In confusing or difficult cases, more advanced statistical techniques may be required.

4. An R&R study should not be run once and forgotten. Such studies should be performed on a regular basis and records kept to monitor performance over time.

Relationship of Process Capability and R&R

For this section, the following relationships apply:

σ_o The sigma of the observed process as determined from a capability study. This sigma should preferably come from the $\frac{\bar{R}}{d_2}$ of a control chart, it is sigma of the process based on at least 100 observations. This sigma includes the variation of the actual process and measurement.

σ_a The sigma of the actual process without the measurement variation. It is not directly measurable.

$\sigma_{R\&R}$ The sigma from an R&R study. This sigma indicates the variation due to measurement.

$$\sigma_o^2 = \sigma_a^2 + \sigma_{R\&R}^2, \text{ and}$$

$$C_p = \frac{tolerance}{6\sigma_o} = observed\ C_p, \text{ then}$$

$$\sigma_o = \frac{tolerance}{6C_p}$$

Let %R&R = X%,

$$\%R\ \&\ R = \left(\frac{5.15\ \sigma_{R\&R}}{tolerance}\right) * 100 = X\%, \text{ then}$$

$$X = \frac{5.15\ \sigma_{R\&R}}{tolerance}, \text{ thus}$$

$$\sigma_{R\&R} = \frac{X\ (tolerance)}{5.15}$$

Using $\sigma_o^2 = \sigma_a^2 + \sigma_{R\&R}^2$ and the formula for σ_o and $\sigma_{R\&R}$,

$$\sigma_a = \sqrt{\sigma_o^2 - \sigma_{R\&R}^2}$$

$$\sigma_a = \sqrt{\left(\frac{tolerance}{6\ C_p}\right)^2 - \left(\frac{X)\ (tolerance)}{5.15}\right)^2}$$

$$\sigma_a = tolerance\sqrt{\left(\frac{1}{6\ C_p}\right)^2 - \left(\frac{X}{5.15}\right)^2}, \text{ then}$$

$$C_{pA} = Actual\ C_p = \frac{tolerance}{6\ \sigma_a}, \text{ and}$$

$$C_{pA} = \frac{tolerance}{6\left[tolerance\sqrt{\left(\frac{1}{6\ C_p}\right)^2 - \left(\frac{X}{5.15}\right)^2}\right]}, =$$

$$C_{pA} = \frac{1}{6\sqrt{\left(\frac{1}{6\ C_p}\right)^2 - \left(\frac{X}{5.15}\right)^2}}$$

C_{pA}, the actual C_p, can be calculated for each combination of the observed C_p and X (the proportion of R&R to the tolerance). Note that some combinations of X and observed C_p are impossible.

These results are shown in table 10.

Table of Actual C_p for Combination of Observed C_p and % R&R

Observed C_p	%R&R 0%	10%	20%	30%	40%	50%	60%	70%
0.50	0.50	0.50	0.50	0.51	0.51	0.52	0.53	0.55
0.60	0.60	0.60	0.61	0.61	0.62	0.64	0.66	0.69
0.70	0.70	0.70	0.71	0.72	0.74	0.77	0.80	0.85
0.80	0.80	0.80	0.81	0.83	0.86	0.90	0.96	1.06
0.90	0.90	0.90	0.92	0.95	0.99	1.06	1.16	1.33
1.00	1.00	1.01	1.03	1.07	1.13	1.23	1.40	1.73
1.10	1.10	1.11	1.14	1.19	1.28	1.43	1.72	2.49
1.20	1.20	1.21	1.25	1.32	1.45	1.68	2.20	5.83
1.30	1.30	1.32	1.36	1.46	1.63	1.99	3.11	
1.40	1.40	1.42	1.48	1.61	1.85	2.42	6.81	
1.50	1.50	1.52	1.60	1.76	2.10	3.08		
1.60	1.60	1.63	1.72	1.93	2.40	4.41		
1.70	1.70	1.73	1.85	2.11	2.79			
1.80	1.80	1.84	1.98	2.32	3.31			
1.90	1.90	1.95	2.12	2.54	4.09			
2.00	2.00	2.06	2.26	2.80	5.52			

Table 10

Attribute Data

The concept of R&R studies are the same for attribute data as for variable data, but the measurement of these is entirely different. The emphasis is on how capable or effective the appraiser is in detecting conforming or nonconforming parts repeatedly and how biased the appraiser is toward rejecting conforming parts or accepting nonconforming parts. The effectiveness of different appraisers can be compared when assessing reproducibility.

The measures used in the inspection capability study for attribute data are defined as follows:

Effectiveness (E): The ability to accurately detect conforming and nonconforming parts. This is expressed as a number between 0 and 1, where 1 is perfect, and is computed by:

$$E = \frac{Number\ of\ Parts\ Correctly\ Identified}{Total\ Opportunities\ to\ be\ Correct}$$

The total opportunities to be correct are a function of the number of parts used and how many times each part is inspected. If 10 parts are selected and each is inspected three times, there are a total of $3 \times 10 = 30$ opportunities to be correct.

Probability of a Miss (P_{miss}): The probability of a miss is the chance of not rejecting a nonconforming part. This is a serious type of error since a nonconforming part is accepted. The probability of a miss is computed by the following formula:

$$P_{miss} = \frac{number\ of\ misses}{number\ of\ opportunities\ for\ a\ miss}$$

The number of opportunities for a miss is a function of the number of nonconforming parts used in the study and the number of times each part is inspected. If five nonconforming parts are used and each part is inspected three times, there are $3 \times 5 = 15$ opportunities for a miss.

Probability of a False Alarm (P_{fa}): The probability of a false alarm is the chance of rejecting a conforming part. This type of error is not as serious as a miss, since a conforming part is rejected. However, rejecting a conforming part causes rework and re-inspection to be performed when it is not necessary. If the P_{fa} gets too large, large sums of money are wasted on rework and re-inspection. The probability of a false alarm is computed by the following formula:

$$P_{fa} = \frac{number\ of\ false\ alarms}{number\ of\ opportunities\ for\ false\ alarm}$$

The number of opportunities for a false alarm is a function of the number of conforming parts used in the study and the number of times each part is inspected. If six conforming parts are used and each part is inspected three times, there are $3 \times 6 = 18$ opportunities for a false alarm.

Bias (B): Bias is a measure of the tendency to classify an item as conforming or nonconforming. Bias is a function of P_{miss} and P_{fa}. Bias values are equal to or greater than zero and have the following interpretation:

B = 1 implies no bias.

B > 1 implies bias towards rejecting parts.

B < 1 implies bias towards accepting parts.

The value of bias is computed by:

$$B = \frac{P_{fa}}{P_{miss}}$$

Data Collection

The collection of samples for evaluating an inspection capability with attribute data is quite different from collecting samples for variable data.

The parts are not selected at random. Parts are selected by appropriate personnel and must be determined as conforming or nonconforming. The number of parts to be selected is shown in table 11. The parts are selected so there will be one-third conforming, one-third nonconforming and one-third marginal. Marginal parts are further divided so they are one-half marginally conforming and one-half marginally nonconforming. This results in the total sample being one-half conforming and one-half nonconforming.

Quantity of Appraisers	Quantity of Gages	Minimum Number of Parts	Minimum Number of Measurements per Part
1	0	24	5
1	1		
2	0	18	4
2 or More	1		
2	2 or More		
3 or More	0	12	3
2 or More	2 or More		

Table 11

Once the parts are selected, they are inspected once in random order by each inspector and the results are recorded on data sheets. Each inspector repeats an inspection and the results are recorded on separate data sheets to eliminate unintentional bias. This is repeated until the required number of inspections are completed. Inspectors should take a normal amount of time for each inspection.

Data Analysis

Analysis of the data is performed using the appropriate worksheets to compute P_{miss}, P_{fa}, E and B. The analysis procedure is illustrated by an example[4].

The example is concerned with a plating operation on a printer part. The visual inspection detects stains and deposits on the part after plating. Three persons are involved in the study: the plating operator, inspector and lead inspector. Seventeen parts are selected initially and after evaluation of the samples by the quality engineer, manufacturing engineer and inspection supervisor, 14 parts (8 conforming and 6 nonconforming) were actually used in the study. Each part was inspected three times. The data obtained is shown in table 12.

The column marked A/R contains the true condition of the part, where A is acceptable and R is reject.

The analysis consists mainly of counting and division. The details of the computations are shown in tables 13 and 14.

Attribute Example Data

Assembly	A/R	A			B			C		
		1	2	3	1	2	3	1	2	3
1	A	A	A	A	A	A	A	A	A	A
2	R	R	R	R	R	R	R	R	R	R
3	A	A	A	A	A	A	A	A	A	A
4	R	R	R	R	R	R	R	R	R	R
5	R	R	R	R	R	A	R	R	R	R
6	A	R	R	R	A	A	A	A	A	A
7	A	R	A	R	A	A	A	A	R	A
8	A	A	A	A	A	A	A	A	A	A
9	R	R	R	R	A	A	A	A	A	A
10	A	A	A	A	A	A	A	A	A	A
11	A	A	A	A	A	A	A	A	A	A
12	R	R	R	R	R	R	R	R	R	R
13	A	A	A	A	A	A	A	A	A	A
14	R	R	R	R	R	R	R	R	R	R

Table 12

[4] *Total Quality Management Handbook*, Jack Hradesky, McGraw Hill, 1995

Inspection Results

Appraiser	Number Good Correct	Number Bad Correct	Number Correct	Number False Alarms	Number Miss	Number Total
A	19	18	37	5	0	42
B	24	15	38	0	4	42
C	23	15	38	1	3	42

Table 13

Calculations

Appraiser	E	P_{fa}	P_{miss}
A	37/42=0.88	5/24=0.21	0/18=0
B	38/42=0.90	0/24=0	4/18=0.22
C	38/42=0.90	1/24=0.04	3/18=0.17

Table 14

The inspection capability study is evaluated using table 15 containing criteria for the parameters. For any marginally acceptable or unacceptable gages or appraisers, corrective action is required and when corrective action is completed, the inspection capability study must be redone.

Attribute Data Criteria

Parameter	Acceptable	Marginal	Unacceptable
E	> 0.90	0.80 to 0.90	< 0.80
P_{fa}	< 0.05	0.05 to 0.10	> 0.10
P_{miss}	< 0.02	0.02 to 0.05	> 0.05

Measurement Error Analysis Guide

Measurement Error Analysis (MEA) is a tool for determining consistency and repeatability of a measurement system. This tool considers operator technique and the discriminatory power of the system.

Definition of Terms for This Section:

N_s Number of samples to be run or compared. Samples should be of a common characteristic.

N_o Number of operators in the study.

N_r Number of readings, days, or times to repeat the study.

S_e Standard Deviation of Error. A statistical measurement of inconsistency in the measurement system. This number reflects the consistency of the measurement system.

S_s Standard Deviation of Samples. A statistical measurement of the difference between samples. This number reflects the amount of difference between the samples in the study.

S_o Standard Deviation of Operators. The statistical measurement of the difference between individual operators. Is there a difference in how operators consistently read each sample?

GCR Gage Classification Ratio. This ratio quantifies the measurement system's ability to discriminate between samples. The higher the ratio, the more capable the measurement system is of telling the difference between samples.

F Represents a ratio found in the F table. This number is used in a formula to help determine if a significant technique difference is present between operators in a measurement system. The F table is a commonly used statistical tool. There are different levels or types of F tables. In an MEA, the 99% F table should always be used. To find the appropriate ratio in the F table it is necessary to determine both horizontal and vertical degrees of freedom (df).

Collecting the Data:

A commonly run and fairly easy MEA would involve a **N_s**, **N_o**, and **N_r** of five (5).

The person conducting the study has several responsibilities to ensure a successful MEA. The first is to obtain proper samples. They should all be of a similar product line. For example, if a facility wants to study micrometers, the person conducting the study could pick a gauge of wire to use for samples (such as 13 Ga.). He would not want to get five samples of wire from the same carrier since there probably would not be a lot of variation within the carrier. Instead he would want to get a sample from five different carriers that were drawn on different days, by different operators, on different machines and maybe even different heats of raw material. He would not want to use two samples of 13 Ga., 2 samples of 13 Ga., and one sample of 12 Ga. The important thing to remember is that you do not want samples exactly alike or samples that are very different.

The second responsibility of the person conducting the MEA is to keep track of the samples during the test. The samples should be masked so that the persons involved in the test do not know which samples are which. Typically, when a MEA is done, the labels on the samples are changed after each reading or day. If the person conducting the test does not maintain 100% control and traceability of the samples throughout the test, it is invalidated. All persons participating in the test should know that the labels are being changed.

After the data is collected, the data from each operator should be placed in a form similar to the one below. Use a similar form to compile the data only. The operators involved in the study should not see or compare their measurements from day to day or between operators in the study. We are trying to determine the amount of variation in the measurement system. If the measurements are all the same, it may invalidate the test.

Operator 1

	Part 1	Part 2	Part 3	Part 4	Part 5	
Day 1						
Day 2						
Day 3						
Day 4						
Day 5						**Grand**
Average						
Range						

Table 15

This is the basic information block that will help organize the data from each of the operators involved in the test. Once the data is in a structured format it is possible to determine $\mathbf{S}_e$, $\mathbf{S}_s$, and $\mathbf{S}_o$.

Key Values:

$\mathbf{S}_e$ Average the average ranges of each operator to determine the grand average of the ranges. Divide the grand range by the d_2 value using $\mathbf{N}_r$ for the subgroup sample size. This is the Standard Deviation of Error. The $\mathbf{S}_e$ quantifies the total amount of variation demonstrated by the measurement system. In other words, if the operators in the study measure a sample that was exactly X, you could expect the operators to measure it between ±3 $\mathbf{S}_e$.

$\mathbf{S}_s$ Average all of the operators **Part 1** readings. Repeat this for all of the parts. When finished, subtract the smallest part average from the largest part average to determine the range. Divide the range by a d_2 value using the $\mathbf{N}_s$ for the sample size. This is the Standard Deviation of the Samples. The $\mathbf{S}_s$ quantifies the total amount of variation demonstrated between the samples. Six (6) times $\mathbf{S}_s$ tells us the total statistical "spread" of the samples.

$\mathbf{S}_o$ List the grand averages for each operator. Subtract the smallest average from the largest average to determine the range. Divide the range by a d_2 value using the $\mathbf{N}_o$ for the sample size. This is the Standard Deviation of the Operators. The $\mathbf{S}_o$ can be used to statistically quantify the difference in how the operators consistently see and record the samples.

With these three values it is now possible to determine the GCR and if there is a significant difference in operator technique.

Gage Classification Ratio (GCR):

The GCR can be calculated by using the following formula:

$$\left(\frac{S_s}{S_e}\right) * 1.414$$

Rules for GCR:

1. If the value is below one (1), it is not possible to tell the difference between samples. The measurement system can not discriminate between samples.
2. If the value is above one (1), it is possible to tell the difference between samples. The measurement system can discriminate.

The higher the value is above one (1), the higher the measurement system's ability to discriminate between samples. One way to look at it is to take the range of the measured product and divide it by the GCR. These units are what the measurement system can discriminate between.

Example: The range for 13 Ga. Wire is .090 - .092. The total range is .002. The GCR is 4. The measurement system can discriminate between 4 different sizes of 13 Ga:

0.0900 - 0.0905
0.0905 - 0.0910
0.0910 - 0.0915
0.0915 - 0.0920

Calculation for Significant Difference in Operator Technique:

This calculation can tell you if there is a statistical difference between operators and how they read a measurement. If they were using the exact same technique, there would be little or no difference between operators. The formula is as follows:

$$\left(S_o * \sqrt{\frac{N_R * N_S}{F}}\right) < S_e$$

If the formula in parentheses is greater than the $\mathbf{S_e}$ value, there is a significant difference in operator technique. Steps must be taken to ensure operator technique differences are reduced before proceeding with the MEA.

Graphical Analysis:

A lot of information can be gained by putting the measurement error data into graphical form. Put the data into an $\bar{X}$ & R Chart.

Grand Average: Use the Grand average of all operators.

LCL and UCL: Average range of all operators multiplied by $\mathbf{A_2}$ and added or subtracted from the grand average using the $\mathbf{N_r}$ for the subgroup sample size.

$$LCL = \bar{\bar{X}} - \left(A_2 * \bar{R}\right)$$

$$UCL = \bar{\bar{X}} + \left(A_2 * \bar{R}\right)$$

URL: Average range of all operators multiplied by $\mathbf{D_4}$.

$$URL = \left(A_2 * \bar{R}\right)$$

LRL: Average range of all operators multiplied by $\mathbf{D_3}$.

$$LRL = \left(D_3 * \bar{R}\right)$$

When the chart limits are established, chart the sample averages from operator 1, operator 2, etc.

Rules for Graphical Analysis of Measurement System:

Control limits are established using the range of readings by the operators. The less variation in the readings, the tighter the control limits.

Look for the following:

1. Low URL--shows very little variation in the readings.
2. Narrow UCL and LCL--caused by low average range. An ideal situation would be one in which the UCL and LCL would very close, if not right on the grand average.
3. On the $\bar{X}$ chart, very few points, if any, should be within the control limits. Unlike normal $\bar{X} - R$ charts, the more points outside the control limits, the better. This is for two reasons. The first is because of the narrow control limits described above. The second reason is also based, in part, on the control limits. In theory, the variation between sample size should be greater than variation within readings. There should be enough difference between samples to distinguish them beyond the variation of measurement error.

Extracting Actual Sigma from Observed Sigma:

Observed sigma = taken from the $\bar{X}$ chart.

$$Actual\ Sigma = \sqrt{\left(Observed\ Sigma\right)^2 - S_e^{\,2}}$$

Factors Table

Sample Size	D_2	A_2	D_3	D_4
2	1.128	1.88	0.00	3.27
3	1.693	1.02	0.00	2.58
4	2.059	0.73	0.00	2.58
5	2.326	0.58	0.00	2.11
6	2.534	0.48	0.00	2.00
7	2.704	0.42	0.08	1.92
8	2.847	0.37	0.14	1.86
9	2.970	0.34	0.18	1.82
10	3.078	0.31	0.22	1.78

Bibliography

Total Quality Management Handbook, Jack Hradesky, McGraw-Hill, Inc., 1995

Concepts for R&R Studies, L. B. Barrentine, LTV Steel Company, Unknown Date

Measurement Systems Analysis, Cayman Systems, http://qs9000.com/pdf_files/MSA_J.pdf, 1998,

Process Quality Control, Ellis R. Ott, Edward G. Shilling, McGraw-Hill, Inc., 1990

Gage Repeatability and Reproducibility, Statsoft, Inc., http://www.statsoft.com/textbook/stathome.html

Measurement Error Analysis, http://deming.eng.clemson.edu/pub/tqmbbs/tools-techs/mea.zip

Simplified Method for Assessing Uncertainties in a Commercial Production Environment, Ian Instone, http://www-uktm.external.hp.com/mikehut/ieepaper/iee2.html

Calculating the Uncertainty of a Single Measurement, Ian Instone, http://www-uktm.external.hp.com/mikehut/one_meas/unc.html

Essentials of expressing measurement uncertainty, http://physics.nist.gov/cuu/Uncertainty/basic.html

Process Capability Studies, Don Winton, http://qs9000.com/pdf_files/CPK.pdf, 1999

The NIST Reference on Constants, Units and Uncertainty, http://physics.nist.gov/cuu/uncertainty/basic.html

A very special thanks to Kelly Speiser for the invaluable assistance, edits and contributions.