Ten Ways to Improve Your Measurement Systems Assessments

Steven Wachs
Principal Statistician
Integral Concepts, Inc.

Why Measurement Systems Assessment (MSA)?

Effective use of data to drive decision making requires adequate measurement systems. For example, when implementing statistical process control charts, we assume that a signal represents a significant change in the *process* and we react as such. However, inadequate measurement systems may result in inappropriate signals or even worse, charts that fail to detect important process changes. Thus, it is incumbent upon us to ensure that measurement systems are adequate for their intended use via proper assessments *prior to their use*. Only capable measurement systems should be utilized in data based methods such as Statistical Process Control, Design of Experiments, Inspection activities, etc.

While most companies perform some aspects of MSA, such as Gage Repeatability & Reproducibility studies, we often observe inadequate assessments of measurement systems. This article identifies ten improvements that most companies can make to their measurement systems assessments.

1. **Understand and Consider *All* Types of Measurement Error**. Gage R&R studies focus on the *precision* of the measurement system at a point in time. That is, the variability observed when the same part is measured by a single operator (repeatability) and by multiple operators (reproducibility). However, measurement systems may exhibit other types of errors and appropriate assessments should be performed to ensure adequacy. These errors include:
    a. **Bias** – difference between the average of repeated measurements and the true/reference value.
    b. **Non-Linearity with Respect to Accuracy** – the bias *changes* across the range of part sizes or materials measured with the gage.
    c. **Non-Linearity with Respect to Precision** – the precision (variation) *changes* across the range of part sizes or materials measured with the gage.
    d. **Instability Over Time** – just like any other process, the measurement system may change over time. Significant changes in bias or precision must be detected in a timely fashion.

2. **Select Specimens Wisely for Gage R&R Studies.** Most Gage R&R studies involve measuring and re-measuring about 10 parts. Often we encounter practitioners who put little or no thought into specimen selection. Ideally, the specimens should represent as much variability as we'd ever be likely to see from the process over time, including from unexpected process changes. The purpose of the study is to assess the measurement system and we can learn the most about the measurement system by including highly variable parts. For example, one operator may have difficulty repeating measurements on "small" parts but be adept at repeatable measurements for "large" parts. If we only included parts close to nominal in the

study, we wouldn't learn about any differences in measurement error over the range of parts measured. Parts should be selected from different batches, time periods, or cycles, and if necessary, parts may be purposely produced with excessive variation for the purpose of the study.

3. **Ensure Adequate Gage Discrimination.** We have seen many studies where the gage segregates all of the measured values into just a few categories. For example, the gage reads to the nearest one thousandth and the parts all range between 2.499 and 2.501. While the resulting data is numerical, it cannot be considered continuous and attempts to use methods for continuous data - such as SPC - will be fruitless. A common guideline is to ensure that the discrimination is at least 1/10th of the actual process variation (e.g. the range describing what the process uses 99% of the time). So if the process range is .003, the measurement system should be able to read to the nearest 0.0003. Note that in cases where the process variation is very small and the process capability is very high, the costs associated with improving discrimination must be weighed against the benefits of doing so.

4. **Understand, Calculate, and Interpret Gage R&R Metrics Correctly.** Common metrics for assessing the results of the gage R&R include:

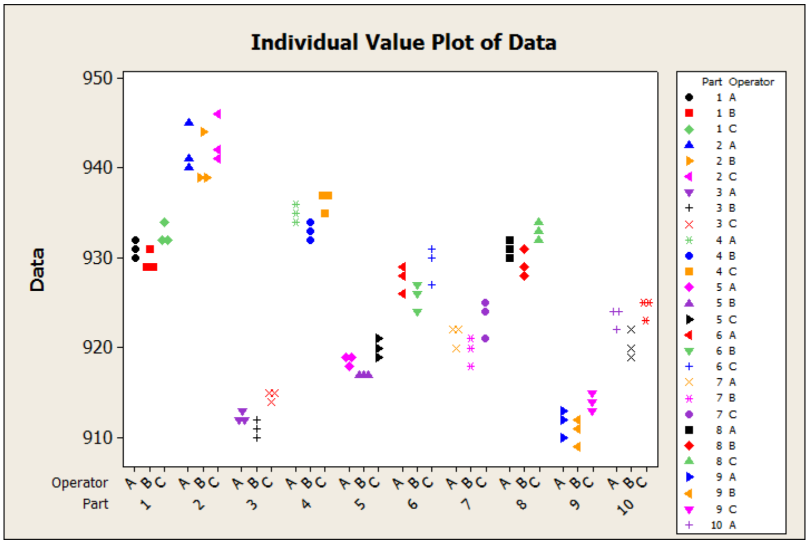
$$\%\ Tolerance\ RR = \frac{5.15\sigma_{RR}}{USL - LSL}(100)$$

$$\%\ Process\ RR = \frac{5.15\sigma_{RR}}{5.15\sigma_{Process}}(100)$$

σ_{RR} = Standard Deviation of Repeatability & Reproducibility Error
$\sigma_{process}$ = Standard Deviation of the Production Process
USL = Upper Specification Limit
LSL = Lower Specification Limit

The *%Tolerance RR* compares the measurement error to the tolerance and primarily assesses the ability to distinguish conforming parts from non-conforming parts. The *%Process RR* compares the measurement error to the process variation and assesses the ability to distinguish parts from *each other*. In other words, %Tolerance RR measures our ability to control the *product* (e.g. inspection) and % Process RR measures our ability to control the *process* (e.g. process control). A frequent mistake is to base the %Process RR on the parts used in the Gage R&R study. These parts often have much more variation than we typically see in the process, especially if the above guideline is followed, and distort the %Process RR statistic. The %Process RR statistic should be based on process variation that is estimated independently from the Gage R&R study – using a sufficient sample from the typical (stable) process.

5. **Look Beyond the "Pass" or "Fail" Outcomes in a Gage R&R study.** The Gage R&R study provides a lot of data but often only one result gets reported (such as *"32% of Process – FAIL"*). The data gets aggregated into one or two metrics that don't tell the whole story. Usually, a deeper dive into the data provides greater insight. In addition to the standard graphical output from common software packages, a simple Individual value plot of each measurement broken down by part number and operator is very useful. An example is shown below.



Sometimes a bad measurement, transcription error, or special cause of variation has an unduly large influence and contaminates the results. Usually, these are easily spotted by graphing the data as above. If repeatability is an issue, do all operators contribute to the problem or does 1 operator account for the bulk of it? Is measurement error on some parts significantly different than others and can this be explained? A careful review of the data graphically helps us understand the contributors to the sources of variation comprising the measurement error.

6. **Use ANOVA for Gage R&R Studies**. The early methods for analyzing Gage R&R data (Xbar and R) are obsolete given the readily available software today. ANOVA has many advantages over the Xbar and R methods including:
   a. More accurate and precise estimates of variation.
   b. The ability to include more factors in the study (besides part and operator).
   c. The ability to quantify interactions between experimental factors (e.g. part/operator interaction).
   d. Hypothesis tests for significance of study factors.
   e. The ability to compute confidence intervals on the key metrics.

7. **Expand Gage R&R Studies to Include Potential Sources of Variation.** Standard Gage R&R studies only include different parts and operators. However other factors often influence the measured values and are simply attributed to measurement error in the basic studies. For example, variation due to part non-uniformity may be accounted for by including measurement location as a factor in the Gage R&R study. Alternatively the exact location to be measured may be controlled in the study. Other factors that may be considered for their impact on measurement error include fixture type, measurement location, temperature or other environmental factors. Any factors which are thought to potentially impact the measurement system should be considered for inclusion in the study.

8. **Apply Methods for Non-Replicable Systems as Necessary.** Standard MSA techniques assume that parts may be re-measured repeatedly. However situations arise where the testing is destructive (e.g. a burst test) and the same parts cannot be re-measured. Other non-replicable situations occur when the process conditions, which affect measurement, cannot be controlled or when the part/sample properties change quickly over time preventing replicate

measurements. Various techniques may be used when the measurement systems are non-replicable although the techniques may require specific assumptions and alternate statistical methods (such as nested ANOVA).

9. **Use Control Charts to Assess the Stability of the Measurement Process.** Just because a Gage R&R and Accuracy study passes with flying colors does not ensure the measurement system will remain adequate over time. Many companies will repeat Gage R&R studies on a periodic basis. However, this approach is risky since the measurement system can change significantly long before the next assessment is performed. Furthermore, time consuming studies may not be necessary. A proactive approach is to regularly monitor both the accuracy and precision of measurement systems using standard control charts. This typically requires that specific parts are retained and measured at periodic intervals (such as weekly). For example, an x-bar chart may be used to monitor the accuracy and the s (standard deviation) chart may be used to monitor precision over time. The same design issues inherent in regular process control charts, such as selecting appropriate sample sizes to detect significant process changes, apply to charts for measurement systems.

10. **Attribute Gage Studies Are Important Too.** Just as with variable gages, attribute gages (such as go/no-go) must be assessed both for accuracy and precision. Special attention should be paid to testing units in which conformance to specification is borderline. Where variable measurements may be taken with an accurate and precise variable measurement system, the risks of accepting non-conforming parts and rejecting conforming parts should be estimated so that the risks are understood and improvements can be made if necessary.