Non-Traditional MSA with Continuous Data

by Keith M. Bower, M.S.

Reprinted with permission from the American Society for Quality

As part of their Six Sigma projects, practitioners often must assess the capability of the measurement system used. In the Service Quality/Transactional arena, there are many similarities with standard measurement system analysis (MSA) procedures. Some subtle differences, however, also require consideration.

Several types of data exist for use in MSA studies:

- **Binary:** Only one of two answers can be provided. For many MSA studies, particularly in the Service Quality/Transactional arena, the only type of response that can feasibly be obtained may be of the form success/failure or true/false. Examples include whether a mortgage request should/should not be honored or whether a particular fee should/should not be waived on a credit card bill.

- **Nominal:** More than two answers are possible, and no natural ordering system to these answers exists. An example includes to which department an inbound telephone call should be routed. A well-defined study would identify, for instance, whether particular calls were being routed to technical support when they should be routed to the sales department instead.

- **Ordinal:** Some type of "grading" system is employed for the response. For example, a Quality Assurance specialist may assess an interaction between an employee and a customer by assigning a score. The scoring system may be rated from 1 through 5, with 1 indicating a very poor response, and 5 an excellent response.[1]

The emphasis of this article, however, is on **continuous** data such as cycle times in minutes and financial loans in dollars. In the four sections that follow, we examine continuous data in relation to MSA studies in the Service Quality/Transactional arena, beginning with the theoretical background of the standard (manufacturing) approach and finishing with a practical example:

MSA with Continuous Data–The Standard Approach
A Frequently Encountered Issue for a Service Quality/Transactional MSA
MSA with Continuous Data–Service Quality/Transactional Scenario
A Practical Example

MSA with Continuous Data: The Standard (Manufacturing) Approach

The following widely used model estimates the contribution of the measurement system to the overall variation in a process:

(1) $Y_{ijk} = \mu + Operator_i + Part_j + (Operator*Part)_{ij} + \varepsilon_{k(ij)}$, $i = 1,2,\ldots,a$; $j = 1,2,\ldots,b$; $k = 1,2,\ldots,n$

Typically, *a* operators and *b* parts are randomly selected from wider populations. The experiment is run by allowing each operator to measure each part *n* times. The type of model considered in (1) is a two-factor random effect model, where Operator, Part, Operator * Part, and ε are assumed to be normally distributed with zero mean and variances of $\sigma^2_{\text{Operator}}$, σ^2_{Part}, $\sigma^2_{\text{Operator*Part}}$, and σ^2_{ε}, respectively. Then, the total variation in a measurement (Y) can be decomposed into:

(2) $\sigma^2_{\text{Y}} = \sigma^2_{\text{Operator}} + \sigma^2_{\text{Part}} + \sigma^2_{\text{Operator*Part}} + \sigma^2_{\varepsilon}$

As suggested by (2), the total variation in Y contains two parts: the variation due to the measurement system (gage), σ^2_{Gage}, and the Part-to-Part variation, σ^2_{Part}, where

$$\sigma^2_{\text{Gage}} = \sigma^2_{\text{Operator}} + \sigma^2_{\text{Operator*Part}} + \sigma^2_{\varepsilon}$$

Among the gage variation component, the variation across operators plus the variation for the Operator by Part interaction is called gage reproducibility. The variation due to the random error term ε is called gage repeatability.[2] The ratio of the gage variation and the total variation or functions of that ratio are widely used to assess a measurement system.[3]

A Frequently Encountered Issue for a Service Quality/Transactional MSA

Consider a system that requires analysts to enter information into a computer database and then determine how much money should be lent to a customer. In practice, analysts may have different interpretations of the protocols involved—possibly through variation in the training they received.

Though this may lead to different dollar amounts calculated for a particular loan *between* analysts, the amount "*within*" each analyst typically remains identical. In other words, an analyst would arrive at exactly the same amount *each time* for a particular loan request. The analyst would be conforming to his or her own interpretation of the protocols each time.

Consider a model for this system as:

(3) $Y_{ijk} = \mu + Analyst_i + Loan_j + (Analyst*Loan)_{ij} + \varepsilon_{k(ij)},$ $i = 1,2,\ldots,a$; $j = 1,2,\ldots,b$; $k = 1,2,\ldots,n$

Note that with this scenario, since the analysts always provide the same amounts for the same loan, the variance for repeatability is zero. The measurement system therefore only computes variance for reproducibility (which may still have merit for analysis purposes). However, when analyzing the data using ANOVA functionality in a statistical software package, it may be impossible to compute certain F-tests to assess significance.

Occasionally, practitioners unwisely "fudge" one or more responses to create an artificial error term, simply to obtain statistical output. Clearly, this is wholly inappropriate. In such a situation, it may be more beneficial instead simply to assess the level of agreement:

- Between the analysts, and
- With the "true" values, obtained using widely agreed upon protocols

MSA with Continuous Data: Service Quality/Transactional Scenario

Using the previously discussed scenario and model (3), consider the response variable of interest (Y) to be instead the length of time (in seconds) for a loan request to be assessed and placed into the database. Though the actual dollar amounts reached are obviously of more concern for analysis, useful insights may still be obtained.

First, note certain important features of the designed experiment:

1. A high number of randomly selected loans would better represent what the process encounters. Choosing more loans would incorporate both "ends of the tail"—the easier and more difficult scenarios, as well as those "in the middle."
2. Randomly selecting the analysts for the experiment does not incur any bias; using only the fastest typists would be a misrepresentation.
3. Having the analysts enter each loan request several times provides an estimate of "repeatability."
4. The order in which the loan requests are assessed is randomized in each sequence for the analysts (to "wash out" effects of fatigue, etc.).
5. A time delay between the cycles of loan request entries may be appropriate to avoid memory recall.

Though the experiment itself does not directly address the resulting loan values themselves (those responses could be assessed using an attribute agreement analysis study), it may still provide useful process information. For example, with regard to the processing times, the study may exhibit:

a. Variation between loans
b. Variation between analysts
c. Interaction effects (i.e., certain analysts having markedly different times for specific loans)
d. "Repeatability" (i.e., the amount of variation within an analyst for processing the same loan)

A Practical Example

Three analysts—Jack, Renee, and Ross—randomly selected for the study each enter data from twenty randomly selected loans. They complete this task twice, the loans being assessed in separate random sequences. The time taken between the cycles of data entry is considered large enough to negate any memory recall effects.

Figure 1 shows the ANOVA output from the experimental results using the two-way random effects model in (3).

Figure 1

Gage R&R Study - ANOVA Method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Loan	19	134774	7093.36	122.236	0.000
Analyst	2	282	140.93	2.429	0.102
Loan * Analyst	38	2205	58.03	8.597	0.000
Repeatability	60	405	6.75		
Total	119	137666			

Clearly, most of the variation in processing times is attributable to variation between loans. Importantly, there is a statistically significant interaction effect (i.e., reject H_0: $\sigma^2_{Loan*Analyst} = 0$ in favor of H_1: $\sigma^2_{Loan*Analyst} > 0$). The presence of an interaction effect can seriously impact conclusions drawn from solely assessing the main effects.

With regard to the variance components, similar to (2), we may consider:

(4) $\sigma^2_Y = \sigma^2_{Loan} + \sigma^2_{Analyst} + \sigma^2_{Loan*Analyst} + \sigma^2_\varepsilon$

Regarding certain ratios of these variance components, as shown in Figure 2, only 2.86% of the variation of entry times in the study can be attributed to the measuring system. Of this small percentage, it is the Analyst*Loan interaction effect that has the most importance (2.12%).

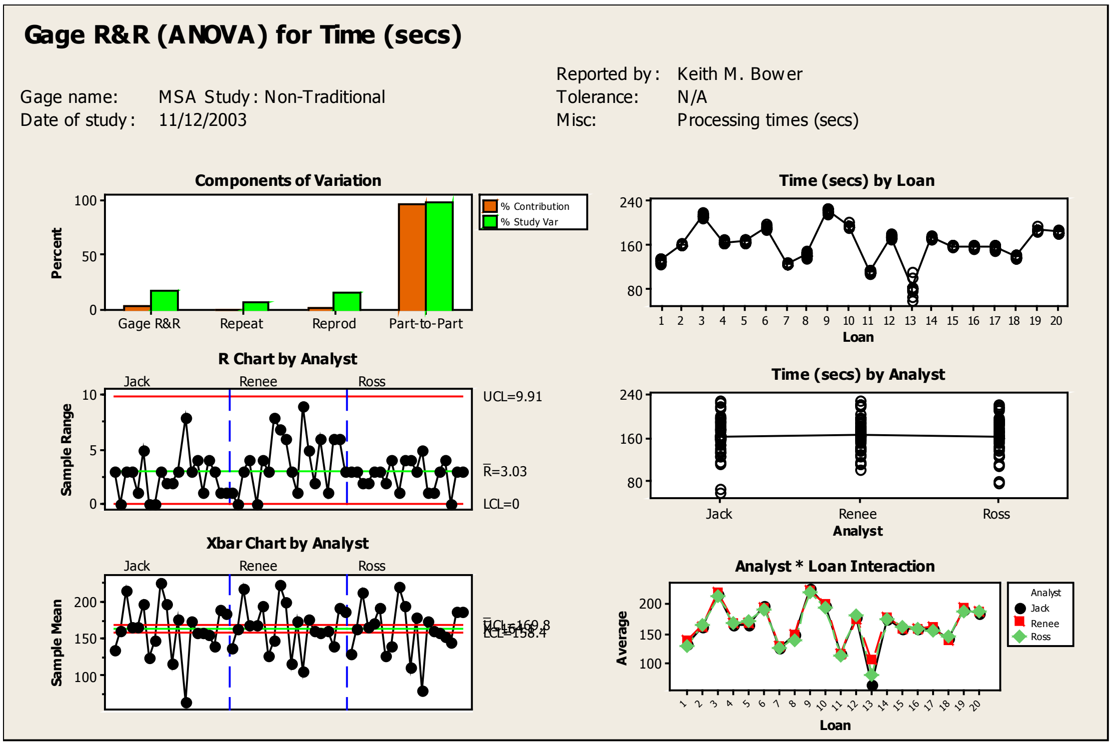
Figure 2

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	34.46	2.86
Repeatability	6.75	0.56
Reproducibility	27.71	2.30
Analyst	2.07	0.17
Analyst*Loan	25.64	2.12
Part-To-Part	1172.56	97.14
Total Variation	1207.02	100.00

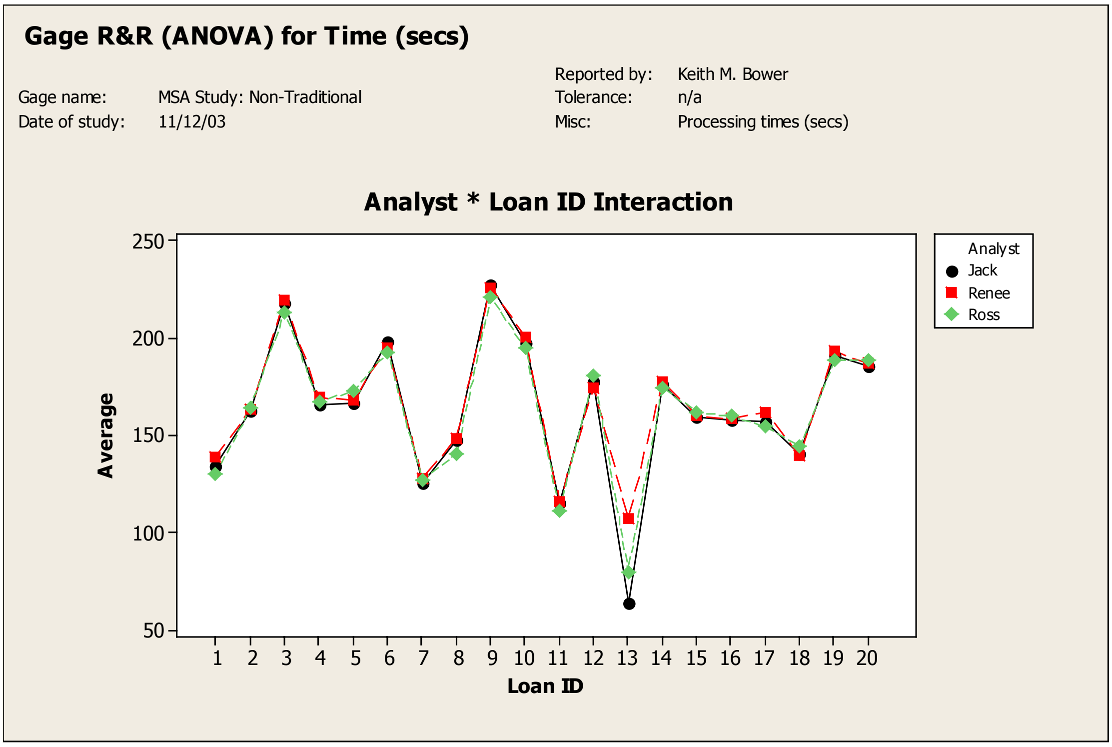
Considering the results in Figure 3, we see that the “repeatability” (i.e., the differences in the two times recorded for each analyst/loan combination) is relatively consistent with an average range of 3.03 seconds. Renee appears to exhibit slightly more variation in the recorded times than Jack and Ross.

Figure 3



As the interaction plot in Figure 4 (reproduced from Figure 3) clearly indicates, a marked difference occurs in the mean entry times between the operators for loan 13. Renee takes longer than Ross, and especially more time than Jack. A continuous quality improvement initiative would address this issue by investigating why the difference occurred for loan 13, and not to the same extent for the others.

Figure 4



Summary

This article has sought to discuss certain nuances when Six Sigma practitioners address MSA procedures in the Service Quality/Transactional arena. Admittedly, the practical example illustrated cannot be considered a traditional MSA per se. The cycle time may be considered a "by-product" of the final recorded value (the loan amount in dollars). That is not to diminish, however, the practical importance of a well-defined and executed study.

About the Author

Keith M. Bower is a technical training specialist with Minitab Inc. He received a bachelor's degree in mathematics with economics from Strathclyde University in Great Britain and a master's degree in quality management and productivity from the University of Iowa in Iowa City, USA. Bower has conducted training courses for firms including GE Capital, American Express, and Motorola. He is a member of ASQ and the Six Sigma Forum.

References

1. For more information on the use and implementation of MSA studies using these data types, see Keith M. Bower, "Measurement System Analysis with Attribute Data" *KeepingTAB* 35, (2002): 10-11, available via http://www.minitab.com/resources/Articles/StatisticsArticles.aspx

2. For a detailed discussion on the use and implementation of this approach to MSA, see Douglas C. Montgomery and George C. Runger, "Gauge Capability and Designed Experiments: Part I: Basic Methods," *Quality Engineering* 6, no.1 (1993): 115-135.

3. See the Automotive Industry Action Group manual, Automotive Industry Action Group, *Measurement System Analysis*, (2002).