(2nd Printing May 2003)

Introduction

The AIAG MSA reference manual was updated on 10th June 2010, however does not indicate the changes that have been made. This document seeks to highlight the main changes to the document by reference to page numbers and paragraphs in version 3 of the manual. As such, this document needs to be read in conjunction with the MSA Version 3 Manual. Minor amendments, for example document references and renumbering of sections/chapters etc, are not referenced in order to not obscure the changes. In order to expedite publication, implications of the changes are not provided. It is hoped that this document will stimulate the discussion of the implications of the changes in due course. This document has been updated to include the first MSA4 errata sheet.

Methodology

The changes are highlighted with reference to version 3 of the MSA manual. The entire text of the new standard is not included, but it is hoped that this document will enable the user to quickly identify the modified sections of the standard and consider the implications.

The Changes

General:

There are many changes to formatting and method of cross referencing of diagrams etc which are not mentioned here as they do not affect the intent of the document.

Page 8: 'Sensitivity'

OEM has now been clarified to explain 'Original Equipment Manufacturer'

Page 10: 'Traceability'

A new (quite lengthy) section has been added on Calibration Systems. This makes reference to a definition of a calibration system and emphasises the need for traceability to reference standards. It also refers to 'internal calibration laboratories' identifying the need for a laboratory scope defining the calibrations that are capable of being performed by such laboratories. Reference is also made to 'Measurement Assurance Programs' to verify the acceptability of measurement processes used throughout the calibration system.

Page 17: Effect on Product Decisions

Paragraph added under note at foot of page stating 'Risk is the chance of making a decision which will be detrimental to an individual or process'.

Page 45: 'Discrimination'

A paragraph has been added to the end of this page referring to the use of a normal probability plot to identify a lack of discrimination of a gage.

Page 46 'Discrimination'

'Only qualified, technical personnel familiar with the measurement system and process should make and document such decisions (i.e. where resolution issues may be the best that technology allows)' has been deleted.

Page 73: 'Chapter II – Section C 'Preparation for a Measurement System Study'

Additional sub para (c) added under section 2) as 'customer requirements'

Page 77: 'Analysis of the Results.

New section added after the first introductory paragraph entitled 'Assembly or fixture error'. This addresses situations where the gage is improperly designed or assembled to ensure that the problem is corrected prior to running the measurement evaluation.

New paragraph inserted in the section on 'Width Error – Acceptability criteria' after the first paragraph. This states that when evaluating the measurement system it can be useful to set priorities on which to focus. It goes on to say that where SPC is being applied and the process is stable, it can be considered acceptable for use and does not require separate re-

q-das@q-das.de • Tel. +49 (0)6201 3941-0

evaluation. If an out of control condition or non-conformance is found in this situation, the first thing that should be done is to evaluation the measurement system.
A note is also added stating the following: 'If the total process is in statistical control for both the mean and the variation charts then it can be assumed that either:

Both the actual process and the measurement variability are acceptable or,

the measurement variation is not acceptable with respect to the process variation (small) but both are in statistical control

In either case the process is producing acceptable product. In case 2) the existence of non-conforming product or out of control condition could be a false alarm.'
Clarification is also made to the section on 'Acceptability Criteria – Width Error concerning the 'general rule of thumb for measurement system acceptability' as follows:

For situations with under 10% error, it is made clear that this situation is recommended especially when trying to classify parts or when heightened process control is required.

For percentage GRR values of 10 to 30 percent, the statement has been changed from 'may be acceptable based on the importance of application, cost of repair etc' to 'may be acceptable for some applications' and 'Decision should be based on, for example, importance of application measurement, cost of measurement device, cost of rework or repair. Should be approved by the customer' added

Over 30% the sentence 'This condition maybe addressed by the use of an appropriate measurement strategy, for example using the average result of several readings of the same part characteristic in order to reduce the final measurement variation' has been added.

An additional section is added entitled 'Additional Width Metric Error'.
This explains the use of the 'number of distinct categories (ndc) as another statistic of measurement system variation. Although reworded, the intent remains the same, stating that the value of ndc statistic should be greater than or equal to 5.
A 'caution' is added regarding the use of the threshold criteria given above, in that they may drive the wrong behaviour (for example the supplier manipulating results).
Comments on Application and Gage Acceptability are also made which basically state that focus should be made on measurement processes which are important to quality and that those that are not should be treated in a more relaxed manner.

Page 79: Chapter III – Recommended Practices for Simple Measurement Systems

Title has changed replacing 'Simple' by 'Replicable'.

Page 81: Example Test Procedures – Introduction

The last bullet point listed has been changed from 'The parts do not change dimensionally during the study' to 'the parts do not change *functionally or* dimensionally during the study, *ie are replicable.*

Page 85: Guidelines for Determining Bias – Independent Sample Method, Conducting the Study

An introductory paragraph has been added explaining the methodology behind the hypothesis test used to evaluate measurement system bias.

Page 85: Analysis of Results – Graphical

An additional step has been added '3)' stating that the bias of each reading should be determined (i.e. better explaining the process).

Page 85: Analysis of Results – Numerical

Clarification is made that the average bias of the readings should be computed (as opposed to the average value of the individual readings). The formula given has been amended to show this.

The method for calculation of the repeatability standard deviation has been changed to calculate the actual (sample) standard deviation from all of the data as opposed to the Range/d^*_2 method as previously specified.

Page 86:

An additional section '7)' has been added after the first paragraph on this page.
This specifies that the repeatability should be evaluated as to its acceptability by calculating the %EV by using the formula 100[EV/TV] or 100[$\sigma_{repeatability}$/TV] where the total variation is based on the expected process variation (preferred) or the specification (tolerance) range divided by 6. Reference is made to the evaluation criteria given in Chapter II section D.
Interestingly an editorial note has been left in stating 'what specifically are we supposed to look at in Section D that links to EV?' It is hoped that some explanation will be forthcoming from AIAG in due course!

Page 86: Determine the t statistic for the bias

The sentence 'bias – observed average measurement – reference value' has been deleted along with the formula t = bias/σ_b. The following formula has been added:
$t_{statistic} = t_{bias} =$ ave bias/σ_b
Page 86 – 'Bias is acceptable at the α level if zero falls within the 1 - α confidence bounds around the bias value' has been reworded to include an additional criterion, ie that the 'p-value associated with t_{bias} is more* than α. No further explanation is given.
**Note: this was changed from 'less' in the MSA subsequently issued errata sheet.*
The value of υ (representing the number of degrees of freedom) is now determined as n-1 (as opposed to looking up g and m values from Appendix C as previously specified)
Page 86 Note 36 at the foot of the page has been deleted (as this refers to the use of the average range method for estimation of standard deviation).

Page 87: Following 'Figure 10: Bias Study – Histogram of Bias Study

A new histogram pictorial has been added to show the expected variation in the average values, and additional explanation added. A numeric illustration has also been added to compare %EV with expected process variation.

Page 87: Final paragraph

The values assigned to the confidence interval of the bias has been changed to (-0.1107, 0.1349) reflecting the change in method of calculation.

Page 88: Table 4 – Bias Study

Values of Standard Deviation and Standard Error of Mean have been changed to 0.2120 and 0.0547 respectively.
$t_{statistic}$, df, significant t value, lower and upper values changed to 0.12, 14, 2.14479, -0.1107 and 0.1241 respectively.
A note has been added to indicate that relating to the number of decimal places indicated in the table.

Page 88: Guidelines for determining bias – Control chart method

Conducting the study
Additional item '2)' added after original '1)' requiring conduct of a stability study with 20 or more subgroups.

Page 88: Analysis of results – graphical

Additional item '3)' added stating that if the control chart reveals that the process is stable and m = 1, then the analysis described for the independent sample method should be used.
Original item '2)' modified to commence with 'if m ≥ 2'.

Page 89: Analysis of results – numerical

Additional step '8)' added after original item '5)' to determine whether the repeatability is acceptable by evaluating %EV by using the formula 100[EV/TV] or 100[$\sigma_{repeatability}$/TV] where the total variation is based on the expected process variation (preferred) or the specification (tolerance) range divided by 6.

Page 89: Original item '6)'
Formula t – bias/σ_b has been replaced with $t_{statistic} = t_{bias}$ = ave bias/σ_b
Page 93: Guidelines for determining linearity
Additional step '7' added after existing step '6)' to determine whether the repeatability is acceptable by evaluating %EV by using the formula 100[EV/TV] or 100[$\sigma_{repeatability}$/TV] where the total variation is based on the expected process variation (preferred) or the specification (tolerance) range divided by 6.
Page 97: Guidelines for determining repeatability and reproducibility
Additional sentence added after 'Except for the Range method, the study design is very similar for each of the methods', stating that the ANOVA method is preferred because it measures the operator to part interaction gauge error whereas the range and the average range method does not include this variation.
At the end of the first paragraph, additional information is provided concerning the ANOVA approach ability to identify additional sources of variation.
Page 99: Average and Range Method
The words 'but not their interaction' have been deleted and replaced with 'However, variation between the appraiser and the part/gage is not accounted for in the analysis'
Page 107: Whiskers Chart
The three existing Figure 19 whiskers charts have now been combined into one.
Page 113: Gage Repeatability and Reproducibility Data Collection Sheet
Values in the average column of the sheet were originally truncated and have now been expanded to three/four decimal places.

Page 115: Analysis of Results – Numerical
A section has been added at the bottom of the page indicating that there are generally four different approaches to determine the process variation which is used to analyse the acceptability of the measurement variation. These are:
1). Using process variation
Process variation, taken from the parts in the GRR study itself (Use when the selected sample represents the expected process variation – preferred option)
2). Surrogate process variation
Use when sufficient samples to represent the process are not available, but an existing process with similar process variation is available.
3). Pp (or Ppk) target value
Use when sufficient samples to represent the process are not available and an existing process with similar process variation is not available or the new process is expected to have less variability than an existing process
4). Specification tolerance
When the measurement system is to be used to sort the process and the process has a Pp 〈 1.0
Page 116:
A new heading has been introduced before the second paragraph entitled 'Using Historical Information'. This states that in order to use this process, the information must be from a process that is in statistical control.
The sentence 'Both of these values (TV and PV) would replace those previously calculated' has been removed and a new heading 'Using a Pp (or Ppk) target value. This states that in order to use the Pp option, use the following TV in the GRR analysis. Since Pp = (USL – LSL)/6σ_p = (USL – LSL)/6s = (USL – LSL)/6TV
Then, TV = (USL – LSL)/6Pp and PV = $\sqrt{[(TV)^2 - (GRR)^2]}$
A new heading has been added 'Using a tolerance (specification) range'
This states that when comparing measurement error from a GRR study to a tolerance, this is the same as comparing to a production process with Pp = 1.0. OEM customers will rarely

q-das@q-das.de • Tel. +49 (0)6201 3941-0

expect to have as low a Ppk as 1.0, nor do they accept a process at that low of a performance level. It may make more sense to compare the measurement variation to a target performance tool which meets the customer requirement. To use this option, use the following GRR analysis: TV = (USL – LSL)/6Pp and PV = $\sqrt{[(TV)^2 – (GRR)^2]}$

Page 117: End of section on Analysis of Results - Numerical

The sentence 'In addition, the ndc is truncated to the integer and ought to be greater than or equal to 5' has been changed to 'For analysis, the ndc is the maximum of 1 or the calculated value truncated to the integer. The result should be greater than or equal to 5'

Some text has been added commenting on the fact that some computer programs may round the truncated results resulting in differences in final reports. The basis for calculation of ndc when using the Pp approach is also given with $PV^2 = (TV)^2 – (GRR)^2$ and ndc = 1.41 (PV/GRR) giving ndc = $[(TV)^2 – (GRR)^2]$/GRR.

Page 121:

The ndc value is now rounded up to 5, instead of being rounded down to 4.

Page 126: Chapter III – Section C Attribute Measurement Systems Study

Possible approaches.

The Upper Specification (USL) limit in Figure 28 has been changed from 0.545 to 0.55. A sentence has been added after the first paragraph stating that 'As for all gages, this attribute study will have 'gray' areas where wrong decisions can be made'.

Page 126:

The paragraph after Figure 29 has been modified stating that 'since this has not been documented by the team, it needs to study the measurement system. However, to address the areas of risk around the specification limits, the team chose approximately 25% of the parts or close to the lower specification limit and 25% of the parts at the upper specification limit. In some cases where it is difficult to make such parts, the team may decide to use a lower percentage recognising that this may increase the variability of the results. If it is not possible to make parts close to the specification limits, the team should reconsider the use of attribute gaging for this process. As appropriate for each characteristic, the parts should be independently measured with a variable gage with acceptable variation (e.g. a CMM). When measuring a true attribute that cannot be measured with a variable gage use other means such as experts to determine which samples are good or defective'.

Page 128: Hypothesis Test Analysis – Cross-Tab Method

A fairly lengthy introduction to the principles by which the cross tabulation method is calculated has been added after the first paragraph. This states 'The cross-tabulation process analyses distribution data for two or more categorical values. The results, presented in a matrix format form a contingency table that illustrates the interdependence between variables'. The section goes on to explain the steps undertaken comparing the results of pairs of evaluators for each part.

A note has been added at the foot of the page stating that cross tabulation functions are available in many statistical analysis programs and in spreadsheet pivot table functions.

The analysis of the process continues with estimation of the expected data distribution. The following has been added: 'What is the probability that an observer pair will agree or disagree on an observation purely by chance? In 150 observations, Observer A rejected the part 50 times and observer B rejected the part 47 times.

$P_{A0} = 47/150 = 0.313$

$P_{B0} = 50/150 = 0.333$

Since the two observers are independent, the probability that they will agree that the part is bad is given by $P(A0 \cap B0) = P_{A0} P_{B0} = 0.104$

The expected number of times for observer A and observer B agree the part is estimated by multiplying the combined probability by the number of observations

$150 \times P_{A0} P_{B0} = 150 \times (47/150) \times (50/150) = 25.7$

The team made similar estimations of each category pair for each observer pair to complete the following table..' The table is provided as per the original document.

Page 129:

The first sentence of the first paragraph has been deleted.

Page 132:

New section entitled 'Sample Size' added before 'concerns' at the foot at the page.
This section addresses the selection of sample size in the study. This states that 'a sufficient number of samples should be selected to cover the expected operating range. With attribute measurement systems the area(s) of interest are the Type II areas. If the process capability is good, then as the process capability improves, the required random sample for the attribute study should become larger. In the example below, the indices were Pp, Ppk = 0.5 (ie an expected process performance of approximately 13% non-conformance), the sample selected was 50%. An alternate approach to large samples is a 'salted sample' where parts are selected specifically from the Type II areas to augment a random sample to ensure that the effect of appraiser variability is seen'.

Page 134: Signal Detection Approach

The entire section has been re-written using the same example data set. Additional explanation is provided in the method of calculation of GRR by this method. The end result has been corrected (from 29%) to 24%.

Page 136:

The formula given at the bottom of the page has been changed to
$T = 6.078 \times |\text{bias}| / \sigma_{repeatability}$

Page 138:

After the calculation of R (= 0.0073), an additional formula has been added to calculate
$\sigma_{repeatability} = R/5.15 = 0.00142$
And the associated GRR range = 0.0085
The formulae given after 'to determine if the bias is significantly different from zero, calculate' have been changed as follows
$t = 6.078 \times |\text{bias}| / \sigma_{repeatability}$
$t = 6.078 \times 0.0023/0.00142$
$t = 9.84$

Page 140: Attribute Gage Performance Curve

GRR value on graph has been changed from 0.0079 to 0.0085

Page 143: Practices for Complex or Non-Replicable Measurement Systems

The table (13) has been deleted and replaced with text which effectively re-iterates what was given in the table as follows:
'Destructive Measurement Systems
When the part (characteristic) being measured is destroyed by the act of measuring the process is known as destructive measurement. This includes the whole class of measurement systems known as 'destructive measurement systems', for example destructive weld testing, destructive plating testing, salt spray/humidity booth testing, impact testing (gravelometer) or mass spectrometry and other material characteristic testing processes. These are 'classic' examples of non-replicable measurement systems since repeated measurements cannot be taken on any single part.
Systems where the part changes on use/test
However, there are other measurement systems which are non-replicable where the part itself, is not harmed by the measurement systems but the characteristic being measured will change. Examples of these are: leak tests with qualitative data, testing using engine test stands, transmission test stands, vehicle dynamometers etc.
Analysis of these systems will depend on:

a homogenous set of parts (small between part variation) can be found to represent a single part

The shelf life of the characteristic (property) is known and extends beyond the expected duration of the study – i.e. the measured characteristic does not change over the expected period of use.

The dynamic (changing) properties can be stabilized'

Page 144: Table – Methods Based on Type of Measurement System
The word 'other' added to 'Non-replicable measurement systems and the item 'Test stands' changed to Measurement systems with dynamic characteristics – e.g. test stands.
Page 161: Average and Range Method – Additional Treatment
Reference to Wheeler and Lyday publication deleted.

Page 161: Procedural Steps
Item 5, first bullet point, has 'if one appraiser is out of control, his method differs from the others' added.
Page 171: Gage Performance Curve Example
In 5th Paragraph, 5.15 has been changed to 6 and P_a values changed to 0.99865 (z=3) and 0.00135 (z=3) respectively.
In 6th paragraph, LL has been changed to LSl and UL has been changed to USL
The note at the bottom of this section has been deleted.
Page 172: Gage Performance Curve Without Error
'Low Limit' is changed to 'LSL' and 'High Limit' to 'USL'
Page 173: Gage Performance Curve - Example
'Low Limit' is changed to 'LSL' and 'High Limit' to 'USL'
GRR Range changed from 0.24 to 0.28
Page 174: Gage Performance Curve Plotted on Normal Probability Paper
GRR Range changed from 0.24 to 0.28
Page 180: Xbar/s Chart
Mean chart UCL has been changed to 0.374 and mean to 0.001. LCL changed to 0.372. S chart UCL changed to 0.4901 and sbar to 0.1908.
Page 188: 5.15 Sigma Spread (Now changed to 6 Sigma Spread)
5.15 values in this section have all been changed to 6.
Page 190: Table 19b Tabulated ANOVA results
5.15 (σ) changed to 6
EV changed to 1.199598
AV changed to 1.361028
GRR changed to 1.814238
Calculation of ndc changed to 1.41(6.25/1.81) = 4.87 =4
% total variation values changed from 5.15 to 6
Page 191: Impact of GRR on the Capability Index – Formulas
Note at the bottom of this section value of 5.15 changed to 6 and last sentence deleted.
Page 205: Glossary
Definitions added 'Bayes Theorum'
Definition of Confidence Interval changed from original
Page 206: Glossary
Definition of Control Chart, additional information added
Page 207: Glossary
Definition added 'Inter-rater agreement'
Definition added 'Kappa (Cohens)'
Page 209: Glossary
Definition enhanced 'Stability'

Pages 211 – 213:
Reference list updated

The information contained in this document has been checked for accuracy to the best of our ability. However, we cannot guarantee that there are no other changes which we have not identified.

q-das@q-das.de • Tel. +49 (0)6201 3941-0