A CASE STUDY ON GAGE R&R IN AUTOMOTIVE INDUSTRY

Carmen SIMION[1*]

ABSTRACT In quality improvement projects it is a standard practice to assess the reliability of measurements before doing any analyses. In particular, in the measure phase of a Six Sigma program the measurement procedures need to be validated. A very important aspect is the measurement system variation, that should be small compared to the product variation or the specification interval (tolerance). The measurement variation can be characterized by accuracy and precision. Usually, to assess any measurement system used in the production environment, many companies use Repeatability & Reproducibility (R&R) studies. Such analyses allow to estimate the contribution of variation attributable to the measurement system itself. If these estimates indicate that the recorded measurements may be unreliable, this may impact all subsequent analyses, e.g. control charts, capability analyses and so one. It is the aim of this paper to address such issues presenting a case study from a major local automotive company. Minitab software was used to calculate the Repeatability & Reproducibility of the analyzed measurement system.
KEYWORDS measurement variation, precision, repeatability and reproducibility (R&R)

1. INTRODUCTION

Measurement data are used more often and in more ways than ever before; for instance, the decision to adjust a manufacturing process is now commonly based on measurement data [1]. So, when one bases decisions on measurement data, the quality of that data is critical because it will affect the ability to make the right decision.

The quality of measurement data is defined by the statistical properties of multiple measurements obtained from a measurement sysem operating under stable condition. A measurement system (MS) is the collection of instruments or gages, standards, operations, methods, fixtures, software, personnel (operators, appraisers, inspectors), environment and assumptions used to quantify a unit of measure or fix assessment to the feature characteristic being measured; a MS is the complete process used to obtain measurements [1].

Any time one measures the results of a process he will see variation. The observed (total) process variation comes from two sources, illustrated in figure 1 [1]: actual process variation (product variation or part to part variation) which refers to differences between parts/products made by the process and measurement variation (measurement system variation or production gage variation) i.e. imperfections in taking measurements. The relationship between the total, part and measurement variation is:

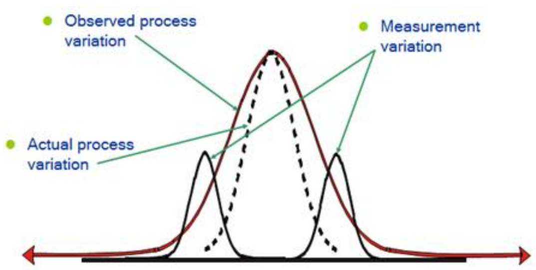


Fig. 1 Observed process variation [1].

$$\sigma^2_{Observed} = \sigma^2_{Actual} + \sigma^2_{MeasurementSystem} \quad (1)$$

where $\sigma^2_{Observed}$ is total variance of the product measurements, σ^2_{Actual} is the variance of the product and $\sigma^2_{MeasurementSystem}$ is the variance of the measurement system.

According to David C. Crosby, "If you don't know the capability of your measurement system, you don't know if your measurements or your products are good or bad" [2]. So, to be confident about what any process delivers, it is important to determine if the measurement system is capable to consistently measure the process output. To assess the adequacy of a measurement system by the variation introduced by the act of measuring (the measurement process), a mathematical procedure named Measurement Systems Analysis (MSA) may be used.

Measurement System Analysis is important [2], [3] to:

- Study the % of variation that is caused by MS
- Compare measurements between operators
- Compare measurements between two (or more) gages
- Provide criteria to accept new gage, before being placed in a production environment
- Evaluate a suspect gage
- Evaluate a gage before and after repair

[1] Lucian Blaga" University of Sibiu, B-dul Victoriei 10, 550024 Sibiu, Romania
Email :carmen.simion@ulbsibiu.ro

- Determine true process variation
- Evaluate effectiveness of a training program

It is recommended that MSA studies be peformed on part characteristics deemed safety critical or critical, part characteristics that contain a poor customer or in-house history record, low capability indeces (Cpk) and part characteristics that are difficult to measure or that have very tight tolerances.

The major components of the measurement process are illustrated in figure 2. The item to be measured can be a physical part, document or a scenario for customer service; operator refers to a person measuring the same products; reference is a standard that is used to calibrate the equipment; procedure is the method used to perform the test; equipment (gage) is the device used to measure the product (frequently, any device used on the shop floor) and environment is the surroundings where the measures are performed [3].

2. THE SUBJECT: REPEATABILITY & REPRODUCIBILITY

Measurement systems, like any other process, are subject to variation. The measurement system variation can be characterized by its location (accuracy) and width (precision) (Fig. 3) [4], [5].

Accuracy shows how accurate the measurement system is and describes the difference between the measurement and the part's actual value. Typically, it is broken into three components: stability, bias and linearity.

Precision shows how precise the measurement system is and describes the variation one sees when he/she measures the same part repeatedly with the same device. Typically, it is broken into two components: repeatability and reproducibility.

For any measurement system, there is always one or both of these problems. For instance, it may measures parts accurately but not precisely; or it is precise accurate but not accurate or it is neither accurate nor precise.

Repeatability is the inherent variation or capability of the equipment itself. Repeatability is the variation in measurements obtained with one measurement equipment used several times by one appraiser while measuring the identical characteristic on the same part. Repeatability is usually called as the "within appraiser" or "within system" variation.The concept of repeatability and its "bulls-eye" perspective are presented in figure 4 [4], [5].

The range chart is used to show the consistency of the measurement process through the repeatability concept.

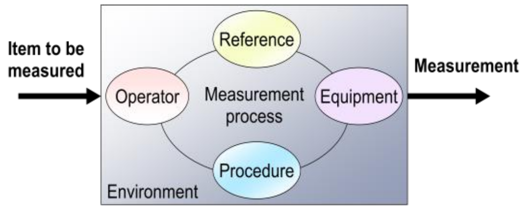


Fig. 2 Components of the measurement process [3].

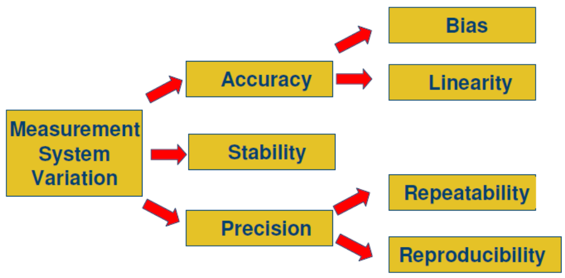


Fig. 3 Components of measurement variation [4].

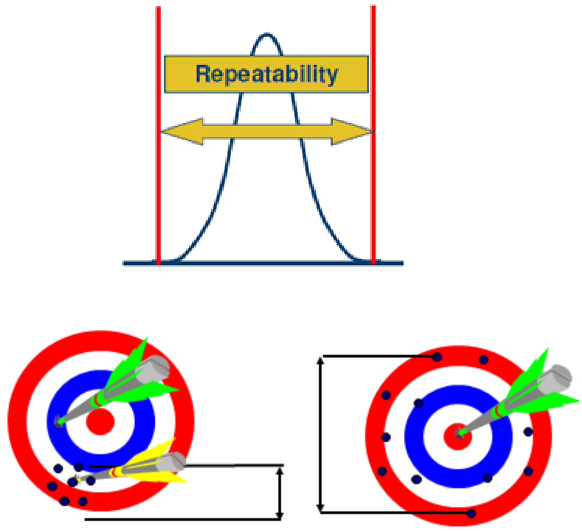


Fig. 4 Repeatability concept [4].

Reproducibility is usually referred to as appraiser variation (influence of appraisers on variability, for example their competence, perceptions, skills discipline and vigilance), so it is called as the "between appraisers" variation. This is often true when the manual equipments are influenced by the operators' skill. However, it is not true when the measurement processes (i.e., automated systems) is used where the operator is not a significant source of variation. Because of this reason, reproducibility is considered as the average variation between-conditions or between systems of measurement and it is called as the "between system" variation.

Reproducibility is the variation in the average of the measurements made by different appraisers using the same measurement equipment when measuring the identical characteristic on the same part.

The X-Bar chart is used to show the consistency of the measurement process through the reproducibility concept. The concept of reproducibility and its "bulls-eye" perspective are presented in figure 5 [4], [5].

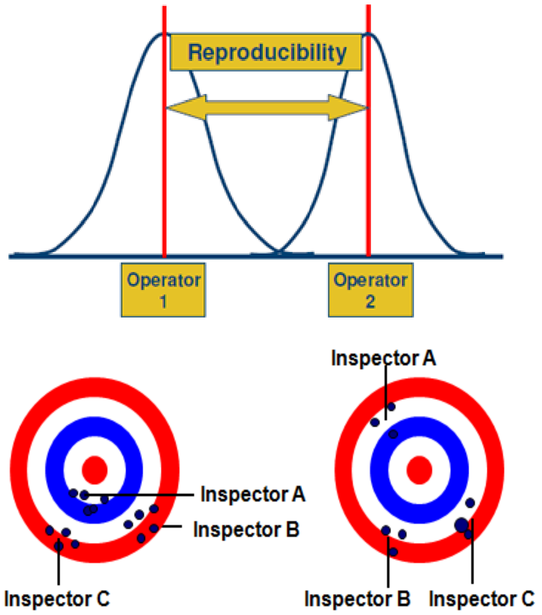


Fig. 5 Reproducibility concept [4].

Gage repeatability and reproducibility (Gage R&R) is an estimate of the combined variation of repeatability and reproducibility. Gage R&R studies determine how much of the observed process variation is due to measurement system variation.

Variable gage R&R is a well known procedure for evaluating measurement systems. A typically study [1] utilizes 2-3 operators/inspectors for one equipment that is measuring a single characteristic. Each appraiser measures 5-10 parts selected from a process 2-3 times. Depending on the cost and time involved one can add more appraisers and measurements and replications.

There are various methodes to conduct R&R studies. From all these, "Average and Range" and "Analysis of Variation, ANOVA" methods are the most prevalent and important. They compute the total measurement system variability, and allows the total measurement system variability to be separated into:

- part-to-part variation or process variation (σ^2_{Actual})
- gage variation ($\sigma^2_{Repeatability}$)
- appraiser variation ($\sigma^2_{Appraiser}$)
- appraiser-by-part variation ($\sigma^2_{Appraiser\ x\ Actual}$) or operator stability, only under the ANOVA method

In a good measurement system, repeatability and reproducibility (R&R) should be as small as possible and the part-to-part variation should be as large as possible, meaning that the measurement system can effective distinguish differences between parts.

Gage R&R is the variance equal to the sum of within appraiser and between appraisers variances and appraiser-by-part variation - only under the ANOVA method:

$$\sigma^2_{R\&R} = \sigma^2_{Repeatability} + \sigma^2_{Appraiser} + \sigma^2_{Appr.\ x\ Actual} \quad (2)$$

In order to decide whether a measurement system is useful for a certain goal, some rules of thumb are necessary. According to AIAG [1], four acceptability criteria must be assessed for variable gage R&R study (Fig. 6). Ideally, the values (in percentage) of all these categories should be in the *green* zone for acceptance; examining the visual aids below shows commonly used judgement criteria for each category:

- % Contribution represents the percent contribution to the overall variation made by each variance component (each variance component divided by the total variation, then multiplied by 100); these percentages add to 100

- % Study Variation represents the percent of the study variation for each component (the standard deviation for each component divided by the total standard deviation). These percentages do not add to 100.

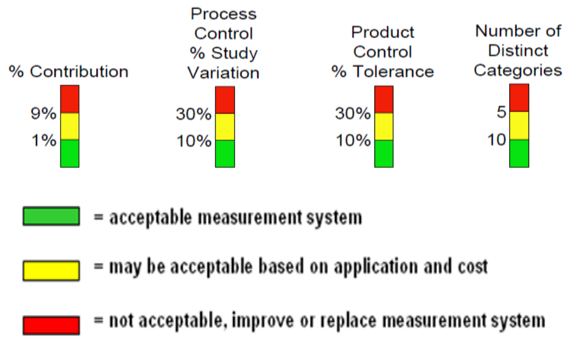


Fig. 6 General rules for gage R&R acceptability.

- % Tolerance is based on upper and lower specification limits (USL and LSL); it is used to compare the measurement system variation with the tolerance

If the measurement system is used for process improvement such as reducing part-to-part variation, then % Study Variation may be a better measure; if the measurement system is used to evaluate parts relative to specifications, then % Tolerance variation may be a better measure.

- The number of distinct categories is based on process variation and tells how many separate groups of parts the system is able to distinguish; to distinguish a higher number of distinct categories, it is necessary a more precise gage.

A statisical software program will produce these values once the data is properly entered.

3. CASE STUDY

This gage R&R study was made in a major company from Sibiu county, a private company working mainly for the automotive industry. The company is a strong partner for some of the most important players in the automotive industry and is constantly upgrading its business relations in other industries also. Its products are assembled on cars like BMW, Mercedes, Peugeot, Renault, Audi, Ford, Fiat, Citroen and Kia and the main groups of products are: components for injection systems; blades, arms and arm heads for windshield wipers; turbo charger components (nozzle ring; central housing; rollers); steering gears and column components; pinions for steering gears; cold and hot coiled springs; stamped parts; drive shafts and components; mechanically welded assemblies; components for air conditioning installations; industrial equipment, tool shop parts, tooling, progressive stamping dies.

The company had intended to install a new gage to measure an internal diametre of the central housing for turbochargers (Fig. 7). The specification for the diameter is 4.2 ± 0.1 mm.

Before being placed the new gage – a digital caliper with a resolution of 0.01 mm (Fig. 8) in the production environment, the company had decided to use measurement system analysis by doing a basic gage R&R study to evaluate if the measurement system measures consistently and accurately, and adequately discriminates between parts.

Three aprraisers (A, B and C) were selected; they randomly sample 10 housings across all major sources of process variation (machine, time, shift, job change) to represent those typically produced. They code the housings to identify the measurements taken on each housing. The first aprraiser measures the 10 housings in random order. Then, the second and the third aprraiser measures the 10 housings in a different random order. Each appraiser measure each part three times.

Fig. 7 Turbo charger components: nozzle ring and central housing

Fig. 8 The Gage: Digital caliper

The measurements with digital caliper are given, partially, in figure 9.

Minitab 16 computer software [6] was used to calculate Repeatability & Reproducibility for the analyzed caliper, with both “Average-Range” and “ANOVA methods. The results for Caliper R&R study, session window output and graph window output, are presented from figure 10 to figure

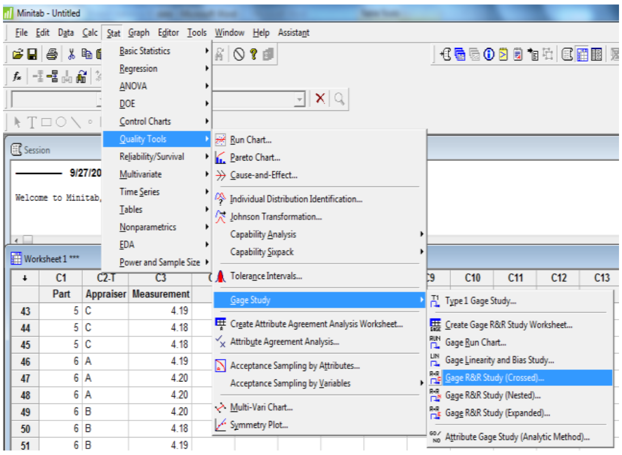


Fig. 9 Measurement results

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.0000132	18.30
Repeatability	0.0000126	17.37
Reproducibility	0.0000007	0.94
Part-To-Part	0.0000591	81.70
Total Variation	0.0000723	100.00

Process tolerance = 0.2

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.0036384	0.0218302	42.78	10.92
Repeatability	0.0035440	0.0212640	41.67	10.63
Reproducibility	0.0008233	0.0049397	9.68	2.47
Part-To-Part	0.0076869	0.0461216	90.39	23.06
Total Variation	0.0085045	0.0510271	100.00	25.51

Number of Distinct Categories = 2

Fig. 10 Session window output for Caliper / Xbar-R method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Part	9	0.0053122	0.0005902	10.5774	0.000
Appraiser	2	0.0000622	0.0000311	0.5575	0.582
Part * Appraiser	18	0.0010044	0.0000558	2.7901	0.002
Repeatability	60	0.0012000	0.0000200		
Total	89	0.0075789			

Alpha to remove interaction term = 0.25

Fig. 11 Session window output for Caliper / ANOVA method

If we look in the session window output for caliper (Fig. 10) at the percent contribution from Part-To-Part (81.70) we see that it is larger than that of Total Gage R&R (18.30). This is considered very good and tells us that much of the variation is due to differences between parts.

But if we look at the %Contribution, %Study Variation, %Tolerance (often informative) columns and at the number of distinct categories we see that for all of these, the acceptability criteria are not in the acceptance zone, so it is required a further analysis to find the sources of measurement error. The supposition was that the digital caliper was not adequate to measure this internal diametre (hole).

The two-way ANOVA table, shown in figure 11, includes terms for the part, appraiser, and appraiser-by-part interaction. Because the p-value for the appraiser-by-part interaction is less than 0.25 (p=0.002), Minitab does not remove the interaction term from the model and does not generates a second ANOVA table [6], [7].

Figure 12 shows the graph window output for Caliper R&R study, generated by Minitab [6], [7]:

- The *Components of Variation chart* graphically represents the Caliper R&R table in the Session window output. Minitab breaks down the Variation in the measurement system into specific sources; each cluster of bars represents a source of variation. For the housing data, the difference in parts accounts for most of the variation.

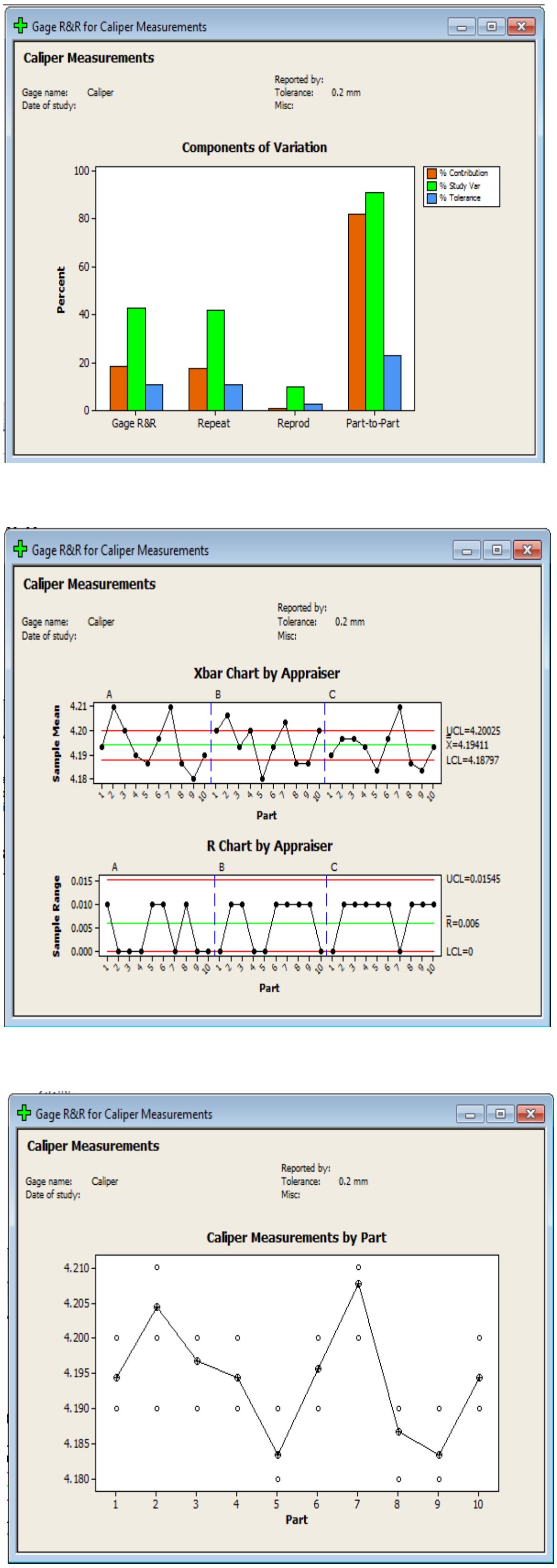
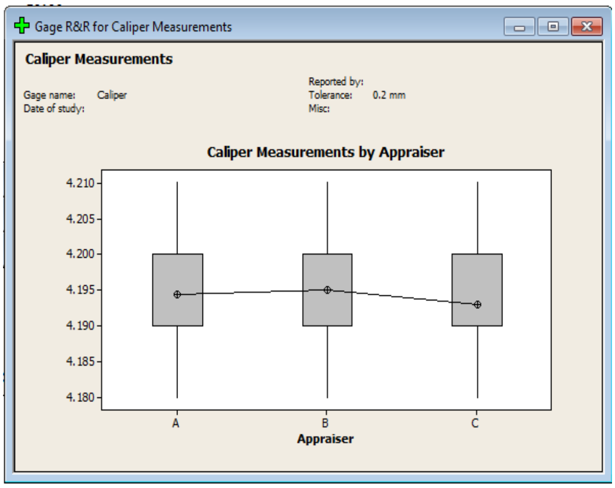


Fig. 12 Graph Window Output for Caliper



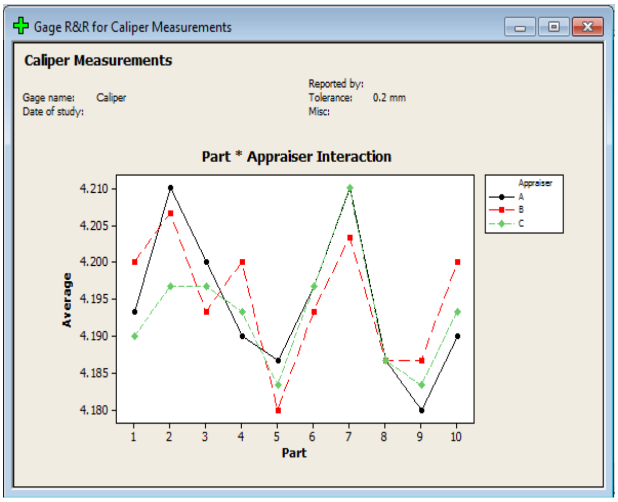


Fig. 12 Graph Window Output for Caliper - continued

- *The x-Bar Chart by Appraiser* compares the part-to-part variation to the repeatability component. Control limits represent measurement error and each point on the X-bar chart is the average of the measures made on the same part by the same operator. Ideally, this graph should show lack of control (points above or below the control limits) because the parts chosen for a gage R&R study should represent the typical part-to-part variability. It is desirable to observe more variation between part averages than what is expected from repeatability variation alone.

For these data, a lot of the points in the Xbar are outside the control limits, indicating variation is mainly due to differences between parts.

- The *R Chart by Appraiser* is a control chart of ranges that graphically displays operator consistency. Each point on the R chart is the range (difference between the highest and lowest) of all measures on the same part by the same operator. All of the points should be within the control limits; if any points on the R-chart fall above the upper control limit (UCL), the operator is having difficulty consistently measuring the parts. The UCL takes into account the numberof times each operator measures a part. If operators measures consistently, the ranges are small relative to the data and the points fall within the control limits.

In this case, the R-chart shows that the appraisers recorded the values for each part with a similar amount of variability.

- *Measurements by Part* chart shows all measurements in the study arranged by part. The measurements are represented by empty circles and the averages by solid circles. A line connects the average measurements for each part. Each point represents at least one measurement of that part by at least one operator (one point may represent more than one measurement). Differences in the scatter of the points may indicate part bias. Ideally, multiple measurements for each part show little variation (the empty circles for each part are close together) and averages vary enough so that differences between parts are clear.

In this case, in the by Part graph there are large differences between parts, as shown by the nonlevel line.

- *Measurements by Appraiser* chart help to mdetermine whether the measurements and variability are consistent across operators. The by operator graph shows all study measurements arranged by operator. Dots represent the measurements; black circles represent the means. Each point represents at least one measurement of a part by one operator (one point may represent more than one measurement). A line connects the average measurements for each operator; if the line is parallel to the x-axis, the operators are measuring the parts similarly and if it is not parallel, the operators are measuring the parts differently. Differences among the operators may indicate operator bias.

In the by Appraiser graph there are small differences between operators (they measure the parts with approximately the same variation), as shown by the level line. Operator C appears to measure slightly lower than the others.

- Each point on *Part by Appraiser Interaction* charts represents the average measurements taken by each operator on each part in the study. A line connects the averages for a single operator. Ideally, the lines will follow the same pattern and the part averages will vary enough that differences between parts are clear: lines are virtually identical means

- operators are measuring the parts the same; one line is consistently higher or lower than the others means that operator is measuring parts consistently higher or lower than the others; lines are not parallel or they cross means the operators ability to measure a part depends on which part is being measured (an interaction between operator and part)

In this case, the divergence in the points indicates an operator by part interaction.

The conclusion of this first study was that this new measuring system is not acceptable and the decision was to replace the digital caliper by a three points internal micrometer with a resolution of 0.001 mm (Fig. 13) and another Gage R&R study was made (Fig. 14, 15 and 16).

Fig. 13 The Gage: Digital three points internal micrometer

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.0000006	0.83
Repeatability	0.0000006	0.77
Reproducibility	0.0000000	0.06
Part-To-Part	0.0000763	99.17
Total Variation	0.0000769	100.00

Process tolerance = 0.2

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.0007991	0.0047944	9.11	2.40
Repeatability	0.0007679	0.0046072	8.75	2.30
Reproducibility	0.0002211	0.0013265	2.52	0.66
Part-To-Part	0.0087352	0.0524109	99.58	26.21
Total Variation	0.0087716	0.0526297	100.00	26.31

Number of Distinct Categories = 15

Fig. 14 Session window output for Micrometer / Xbar-R method

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Part	9	0.0070867	0.0007874	63.6595	0.000
Appraiser	2	0.0000047	0.0000023	0.1895	0.829
Part * Appraiser	18	0.0002226	0.0000124	14.0914	0.000
Repeatability	60	0.0000527	0.0000009		
Total	89	0.0073667			

Alpha to remove interaction term = 0.25

Fig. 15 Session window output for Micrometer / ANOVA method

If we look in the session window output for micrometer (Fig. 14) at the percent contribution from Part-To-Part (99.58) we see that it is larger than that of Total Gage R&R (9.11). This is considered very good and tells us that much of the variation is due to differences between parts and very little to the measurement system.

If we analyse the values for %Contribution, %Study Variation, %Tolerance and the number of distinct categories we see that for all of these, the acceptability criteria are in the acceptance (green) zone.

So according to AIAG guidelines, this system is acceptable because the variation due to the measuring system, either as a percent of study variation or as a percent of tolerance, is less than 10%, contribution is less than 1% and and the number of distinct categories is more than 10%.

Figure 16 shows the graph window output for Micrometer R&R study, generated by Minitab [6], [7]:

- The *Components of Variation chart* - the difference in parts accounts for most of the variation.
- *The x-Bar Chart by Appraiser* - most of the points in the Xbar are outside the control limits, indicating variation is mainly due to differences between parts.
- The *R Chart by Appraiser* - the appraisers recorded values for each part with a similar amount of variability.
- *Measurements by Part chart* - there are large differences between parts, as shown by the nonlevel line.

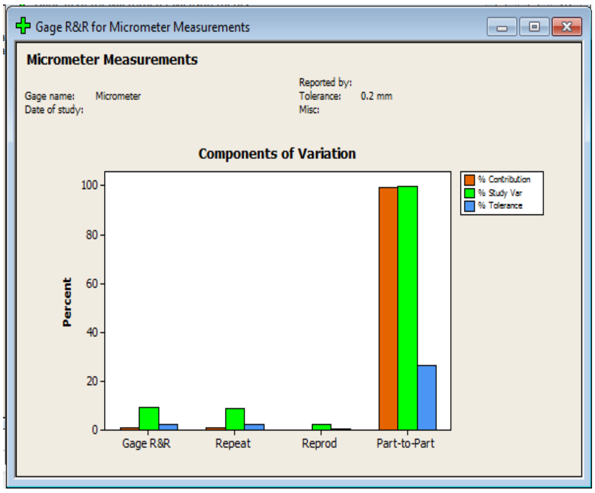


Fig. 16 Graph Window Output for Micrometer (1)

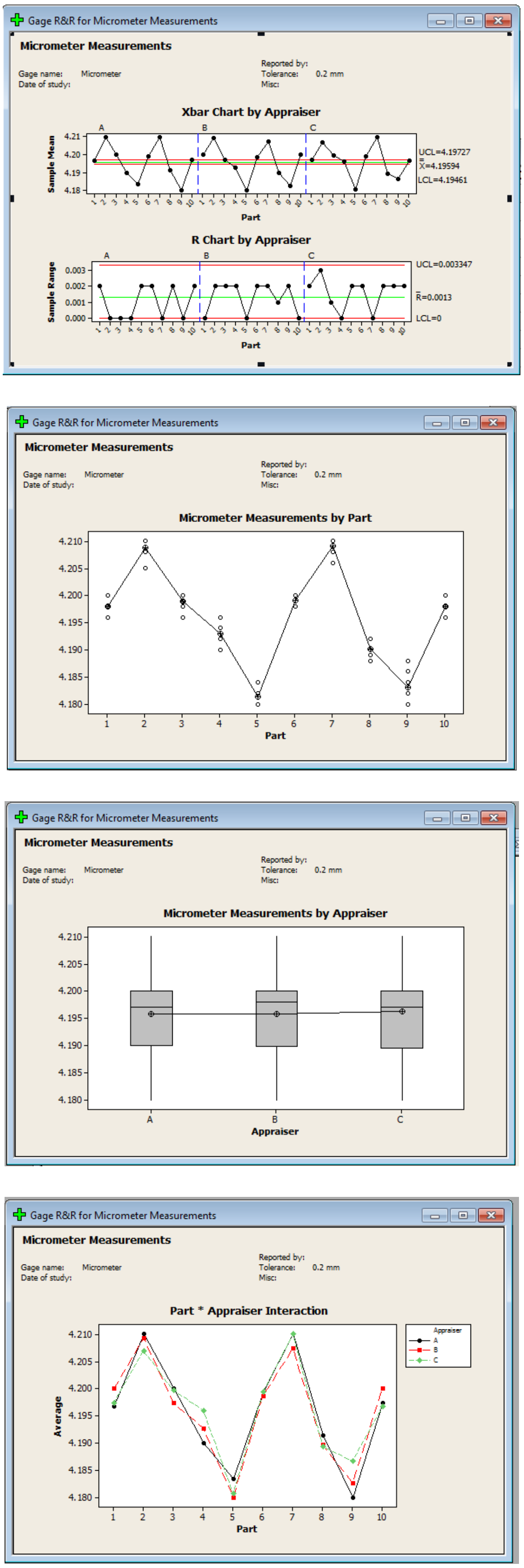


Fig. 16 Graph Window Output for Micrometer (2,3,4,5)

- *Measurements by Appraiser chart – appraisers* measure the parts with approximately the same variation, as shown by the level line
- *Part by Appraiser Interaction char*t - the lines follow one another closely and the differences between parts are clear. The operators seem to measure parts similarly.

The new housing measuring system contributes very little to the overall variation, as confirmed by both the session window output and graphs window output.

4. CONCLUSION

Regarding to the Raffaldi and Kappele [8] "If measurement variation can be reduced and gauge repeatability and reproducibility ratios improved, it is easier to differentiate between parts that are in or out of specification, allowing parts to be accepted or rejected with greater confidence". Hence, the gage R&R (GR&R) could use as an auditing tool which gives feedbacks on the improvement of measurement methods.

A successful quality improvement or statistical process control program needs good measurement systems. Measurement equipments which are functioning properly are critical for quality assurance, process control and process optimization activities.

The use of Gage R&R is a useful component in a measurement system analysis program.

Usually, the gauge repeatability and reproducibility (GR&R) study needs to be conducted prior to the process capability analysis for verifying the precision of measuring equipments and helping organizations improve their product quality. Therefore, how to ensure the quality of measurement becomes an important task for quality practitioners.

The efficiency of a measurement system is with exact relation to the accuracy of measurement instruments. Usual measurement tools such as calipers and micrometers are the ones that could cause the most anxiety with their wrong usage. The reason is that equipment and process measurements have to be controlled appropriately to obtain data of high quality for correct a analysis.

Identifying and reducing measurement variation is the whole reason for doing gage R&R study.

REFERENCES

[1] *Measurement Systems Analysis* Reference Manual, Fourth edition, Chrysler Group LLC, Ford Motor Company, General Motors Corporation (Automotive Industry Action Group, AIAG), Detroit-Michigan, USA, June 2010.
[2] Crosby, D. C., 1998, *A Managers Guide to Gauge R&R*, Rubber World 218, 16-17
[3] Wheeler D. J., Lyday R. W., 1989, *Evaluating the Measurement Process*, second ed., SPC Press Inc., Knoxville, Tennessee.
[4] Information on http://www.scribd.com/doc/191482362/08-Measurement-Systems-Analysis
[5] Information on http://www.free-six-sigma.com/six-sigma-tools.html
[6] Minitab 16 Statistical Software, 2010, State College, PA: Minitab, Inc.
[7] Information on http://www.minitab.com/uploadedFiles/Documents/sample-materials/FuelInjectorNozzles-EN.pdf
[8] Smith R.R., McCrary S.W., Callahan R.N., 2007, *Gauge repeatability and reproducibility studies and measurement system analysis: A Multi method exploration of the state of practice*, Journal of Quality Technology, 23/1, 1-11.