DOI: 10.2478/aucts-2018-0001

ACTA UNIVERSITATIS CIBINIENSIS – TECHNICAL SERIES
Vol. LXX 2018

ASSESSMENT OF HUMAN CAPABILITY, AN EFFECTIVE TOOL TO OBTAIN CONFIDENCE IN THE VISUAL INSPECTION PROCESS

SIMION Carmen
Faculty of Engineering/Department of Industrial Engineering and Management,
"Lucian Blaga" University, Sibiu, Romania, carmen.simion@ulbsibiu.ro

Abstract: *To consistently produce high quality products, a quality management system must be practically implemented in every organization. One of its core instrument is to ensure the capability of the measurement systems, which are the basis for decisions regarding the behavior of the product critical quality characteristics. Base on requirements of the quality management system, a Measurement System Analysis should be conducted for all measurement system which are mentioned in the organization quality plan. Most problematic measurement system issues come from measuring discrete data, which are usually the result of human judgment (subjective decision) when categorizing products such as good/bad (visual inspection). It was the aim of this paper to address such an issue presenting a case study made in a local company from the Sibiu region, in order to evaluate how capable are the appraisers to visually inspect steel chains. The results were analyzed using MINITAB statistical software with its module called Attribute Agreement Analysis. The conclusion was that the inspection process must be improved by operator training, developing visual aids/boundary samples, establishing standards and set-up procedures.*

Key words: visual inspection**,** measurement system analysis

1. Measurement System Analysis

To consistently produce high quality products, a quality management system must be implemented in every organization. One of its core instrument is to ensure the capability of the measurement systems, which are the basis for decisions regarding the behavior of the product critical quality characteristics. Base on requirements of the quality management system, a Measurement System Analysis (MSA) should be conducted for all measurement system which are mentioned in the organization quality plan.

MSA is an experimental and mathematical procedure to quantify the variation introduced to a process or product by the act of measuring. Measurement System Analysis is important to study the percentage of variation in a process that is caused by the measurement system, to compare measurements between operators, to compare measurements between two or more gages (devices used to obtain measurements), to provide criteria to accept new gage, to evaluate a suspect gage, to evaluate a gage before and after repair and to evaluate effectiveness of a training program [1].

A measurement system is a collection of operations, procedures, gages and other equipment, software and personnel used to assign a number or grade (qualificative, classification) to the characteristic being measured or categorized [2]. Industrial measurement system can be divided into two categories, one is variable measurement system, applied to continuous data and another is attribute measurement system, applies to discrete data.

Most problematic measurement system issues come from measuring discrete data, which are usually the result of human judgment (subjective decision) when categorizing products such as good/bad (visual inspection). This is because it is very difficult for all inspectors/appraisers to apply the same operational definition of what is “good” and what is “bad.” However, such measurement systems are seen throughout industries so it is important to quantify how well they are working [3].

Regardless of technological advancements in nondestructive inspection techniques, the primary and most important is visual inspection that will likely remain the first inspection method used in many field applications. When attempting to determine the soundness of any product for its intended application, visual inspection is normally the first step. Generally, almost any product can be visually examined to determine the accuracy of its fabrication.

Because requirements for visual inspection typically pertain to the inspector vision, it is of great

importance that the appraiser has sufficient training and experience before performing visual inspection, so the assessment of human capability is crucial to obtain the confidence in the inspection process.

The visual inspection of products involves a lot of problems. The most mattering of them is the one of the important variation of the results of inspection, leading decisions on the conformity of the product sometimes totally opposed. The principal difficulty is that human is the only measuring device which can be used for this type of inspection. Each appraiser must inspect the product and judge the conformity. Consequently, the subjectivity of measure is very present and the variability therefore high [4].

When the results of a measurement system are good or bad rather than a quantitative value, special procedures are necessary. The most commonly used methodologies used for attribute MSA: short method, hypothesis test analysis, signal detection theory and long method are defined in measurement systems analysis reference manual [2].

The attribute MSA is a set of trials conducted to assess the inspector ability to categorize products. An attribute MSA has multiple inspectors (two or more) that categorize, independently and in a random order, multiple products (usually between 20 and 30, more is preferable) multiple number of times (two or more trials); the products should represent the full range of process variation (good, bad and borderline).

The test is analyzed based on correct vs. incorrect answers to determine the goodness of the measuring system. Measures (statistics, scores) are calculated based on how often the appraisers correctly (agreement with the standard) characterize each product and how frequently they agree with themselves and each other.

The hypothesis test analysis methodology uses two primary methods of assessing the agreement of the attribute with the standard: the percentage or extent to which the appraisals agree with the standard and Kappa statistics, the percentage or extent to which adjustment is made between the agreement between the appraisals and the standard (after chance agreement has been removed):

- **agreement within appraiser** - the percentage or extent to which each appraiser agrees with himself or herself on all trials when each appraiser conducts more than one trial
- **agreement between appraisers** - the percentage or extent to which all appraisers agree with each other on all trials when more than one appraiser makes one or more appraisals
- **agreement of each appraiser vs. standard** - the percentage or extent to which each appraiser agrees with himself or herself as well as with the standard when a known standard is specified
- **agreement of all appraisers vs. standard** - the percentage or extent to which all appraisers agree with each other on all trials as well as with the standard when a known standard is specified
- **kappa** statistic indicating the degree of agreement of the assessments made by multiple appraisers when evaluating the same products; Fleiss's kappa - statistic used for assessing the reliability of agreement when appraiser(s) are selected at random from a group of available appraisers; Cohen's kappa - statistic used for assessing the reliability of agreement when the appraiser(s) are specifically chosen and are fixed.

The key in all measurement systems is having a clear assessment method and clear criteria for what to accept and what to reject. The hypothesis test analysis method, used in this analysis, consists mainly of qualificative/classification counting and division, and the results are evaluated using acceptability criteria from AIAG MSA reference manual presented in table 1 [2]. For any marginally acceptable or unacceptable measurement system, corrective action is required and when corrective action is completed, the attribute MSA must be redone.

Table 1: Decision criteria

Measures	Acceptable	Marginally acceptable	Unacceptable
Within Appraiser	> 90%	80% to 90%	< 80%
Each Appraiser vs. Standard	> 90%	80% to 90%	< 80%
Disagreement OK/NOK	< 2%	2% to 5%	> 5%
Disagreement NOK/OK	< 5%	5% to 10%	> 10%
Between Appraisers	> 90%	80% to 90%	< 80%
All Appraisers vs Standard	> 90%	80% to 90%	< 80%

Source: AIAG MSA reference manual [2]

The MSA manual reference says: "A general rule of thumb is that values of kappa greater than 0,75 indicate good to excellent agreement (with a maximum kappa = 1); values between 0,40 and 0,75 indicate marginal agreement and values les than 0,40 indicate poor agreement" [2].

2. Case study

The research was conducted in a local plant from Sibiu, part of a dynamic and modern german family group, that is active worldwide. With over 140 years of experience and taking round steel chains as a basis, the plant develop future-oriented solution concepts with chain systems and components in diverse forms and for a wide range of applications: sling and lashing systems, military technology, tyre protection chains, snow chains, conveyor & drives systems, industrial and hoist chains, forestry chains, object equipment and barriertech (fig. 1).

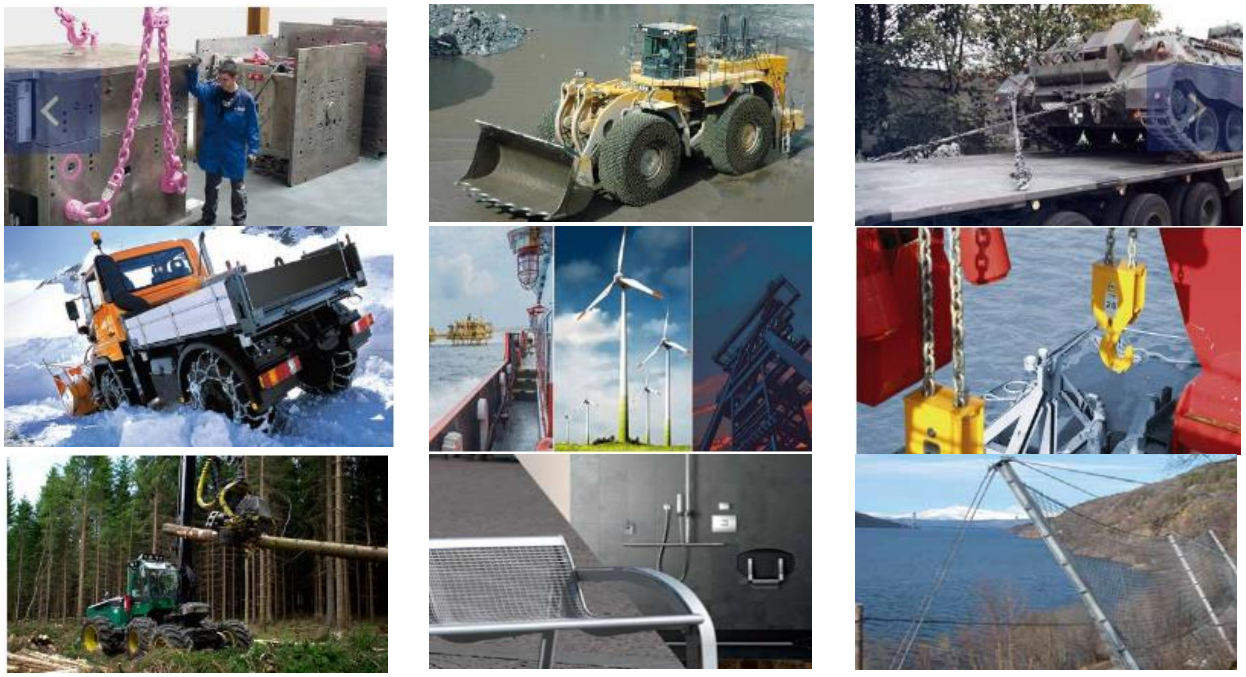


Figure 1: Product range
Source: http://web.rud.com/en-us/company.html

In order to evaluate the capability of the inspection process, the paper presents an attribute measurement study for the linking system between chains (chain linking), presented in figure 2. These linking systems are removable coupling elements designed to ensure their use in the entire product range and consist of two identical parts (subassemblies) obtained by forging, which are then machined, respectively hardened and milled to obtain the "teeth" required to couple the two subassemblies. A common problem that occurs in these forged parts are the cracks that appear, at the end of the milling area, in the outer curvature of the chain linking. Some cracks can be spotted with the naked eye and others by magnetic particle inspection. Because this is a current problem, it requires a rigorous training of inspectors that must ends with an MSA test designed to reveal valuable insights into the capabilities and weaknesses of inspectors.

Figure 2: The linking system between chains

The visual inspection results were processed using MINITAB software [5] with its module called Attribute Agreement Analysis and are presented from figure 3 to figure 7.

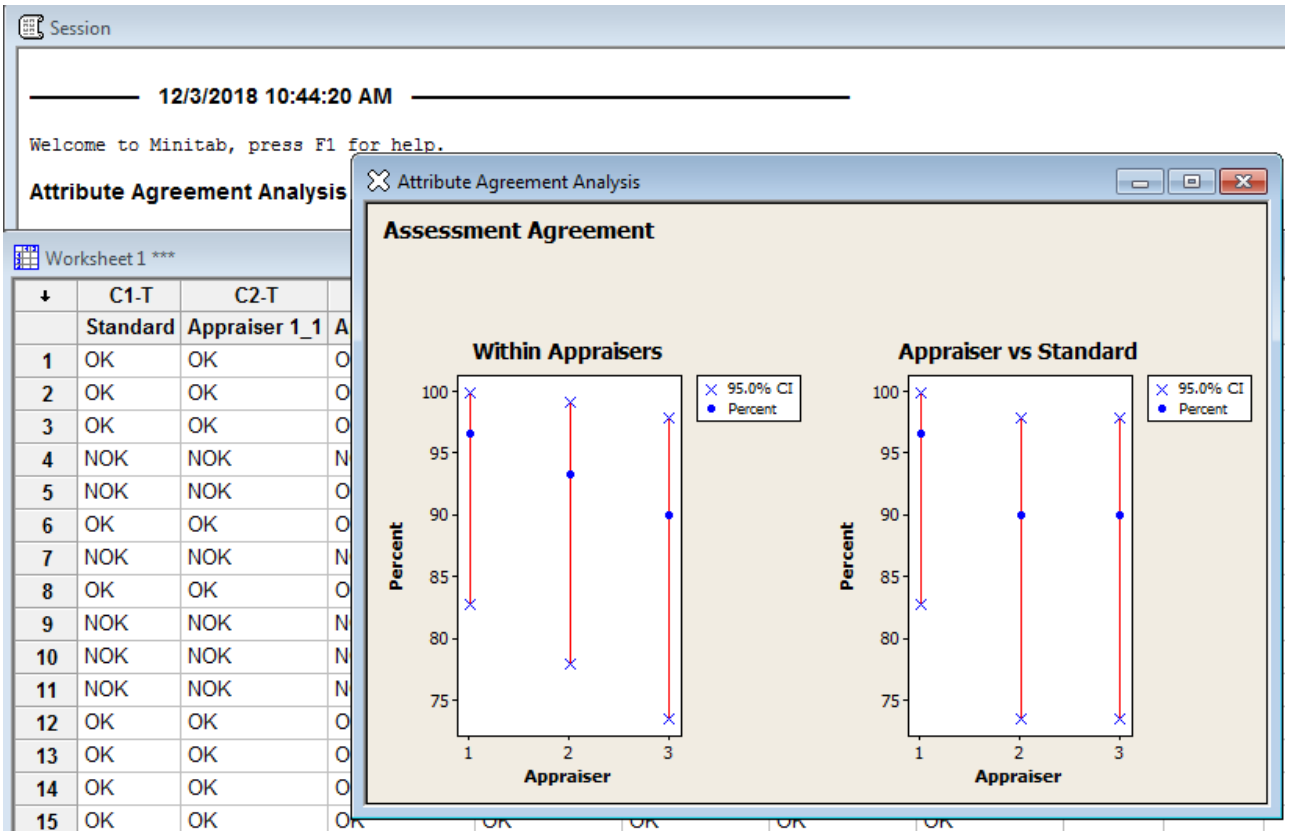
Analysing results of the research, the conclusions are as follows:

- Individual Repeatability of all appraisers (Within Appraisers) is above 90%, so acceptable (fig. 3, left side and figure 4): this means that appraisers are consistently with themselves.

Also, Fleiss's and Cohen's values of kappa statistics indicate good agreement for appraiser 3 and very good agreement for appraiser 1 and 2 (figure 4).

- Individual Effectiveness of all appraisers (Each Appraiser vs. Standard) is also above 90%, so acceptable (fig. 3 – right side and figure 5): this means that appraisers are in agreement with the true status (standard) of the linking system.

Also, Fleiss's and Cohen's values of kappa statistics indicate very good agreement for all appraisers (figure 5).



	C1-T	C2-T	
	Standard	Appraiser 1_1	A
1	OK	OK	O
2	OK	OK	O
3	OK	OK	O
4	NOK	NOK	N
5	NOK	NOK	O
6	OK	OK	O
7	NOK	NOK	N
8	OK	OK	O
9	NOK	NOK	N
10	NOK	NOK	N
11	NOK	NOK	N
12	OK	OK	O
13	OK	OK	O
14	OK	OK	O
15	OK	OK	O

Figure 3: Minitab worksheet and graphs

Within Appraisers

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95 % CI
1	30	29	96.67	(82.78, 99.92)
2	30	28	93.33	(77.93, 99.18)
3	30	27	90.00	(73.47, 97.89)

Matched: Appraiser agrees with him/herself across trials.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa
1	NOK	0.931429
	OK	0.931429
2	NOK	0.861111
	OK	0.861111
3	NOK	0.797980
	OK	0.797980

Cohen's Kappa Statistics

Appraiser	Response	Kappa
1	NOK	0.931507
	OK	0.931507
2	NOK	0.861111
	OK	0.861111
3	NOK	0.798206
	OK	0.798206

Figure 4: Within appraiser results

- Assessment Disagreement (figure 5) results show that (only) appraiser 2 classified a conforming linking system as a non conforming on 1 occasion but all appraisers were inconsistent in their judgments.

- Reproducibility of Measurement System (Between Appraisers), shows that all three appraisers agreed with each other on both assessments to 83,33% (on 25 out of 30 inspected linking system), so marginally acceptable (figure 6) and Fleiss's values of kappa statistic indicate very good agreement, too.

- The last metric, overall Effectiveness of the Measurement System (All Appraisers vs Standard) tells that for 25 out of 30 linking system inspected all three appraisers agreed with the true status of product, which represents 83,33%, so marginally acceptable; Fleiss's and Cohen's values of kappa statistics indicate very good agreement.

Each Appraiser vs Standard

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95 % CI
1	30	29	96.67	(82.78, 99.92)
2	30	27	90.00	(73.47, 97.89)
3	30	27	90.00	(73.47, 97.89)

Matched: Appraiser's assessment across trials agrees with the known standard.

Assessment Disagreement

Appraiser	# OK / NOK	Percent	# NOK / OK	Percent	# Mixed	Percent
1	0	0.00	0	0.00	1	3.33
2	1	7.69	0	0.00	2	6.67
3	0	0.00	0	0.00	3	10.00

OK / NOK: Assessments across trials = OK / standard = NOK.
NOK / OK: Assessments across trials = NOK / standard = OK.
Mixed: Assessments across trials are not identical.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa
1	NOK	0.965714
	OK	0.965714
2	NOK	0.862857
	OK	0.862857
3	NOK	0.898457
	OK	0.898457

Cohen's Kappa Statistics

Appraiser	Response	Kappa
1	NOK	0.965753
	OK	0.965753
2	NOK	0.863014
	OK	0.863014
3	NOK	0.898494
	OK	0.898494

Figure 5: Each appraiser versus standard results

Between Appraisers

Assessment Agreement

# Inspected	# Matched	Percent	95 % CI
30	25	83.33	(65.28, 94.36)

Matched: All appraisers' assessments agree with each other.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
NOK	0.854251	0.0471405	18.1214	0.0000
OK	0.854251	0.0471405	18.1214	0.0000

Cohen's Kappa Statistics

You must have two appraisers and single trial per appraiser to compute Kappa.

Figure 6: Between appraisers' results

The 83,33% score indicates not a very high efficiency of this inspection process (ideally must be over 90%), but it is acceptable. Therefore, an action plan should be drawn up with a view of the improvement of the appraiser training process, the environment (light), the limit samples, etc.

All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95 % CI
30	25	83.33	(65.28, 94.36)

Matched: All appraisers' assessments agree with the known standard.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
NOK	0.909009	0.0745356	12.1956	0.0000
OK	0.909009	0.0745356	12.1956	0.0000

Cohen's Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
NOK	0.909087	0.0744200	12.2156	0.0000
OK	0.909087	0.0744200	12.2156	0.0000

Figure 7: All appraisers versus standard results

3. Conclusions

Human measurement systems are often used in a lot of processes to perform visual inspection, so their assessment is important to see where are the problems, in order to eliminate them and to guide the process improvement. Visual inspection is a very effective inspection method, and it should be the primary method included in any effective quality control program.

Because most processes require at least some form of subjective judgment, visual inspection must be carried out by a trained person, in which any product defect is detected by the aid of a naked eye. The attribute study is just one of many MSA tools an organization may use to better understand their processes and verify the validity and utility of the data collected through their measurement systems. Such understanding contributes to the organization's efforts to improve quality and respond to change through informed decision making

4. References

1. Open Source Six Sigma, *Lean Six Sigma Black Belt manual*, 3th edition, Scottsdale, Arizona (2007).
2. Automotive Industry Action Group AIAG, *Measurement Systems Analysis Reference Manual*, 4th edition, Detroit, Michigan (2010).
3. Chieh, C. J., Making Sense of Attribute Gage R&R Calculations, Information on *http://www.isixsigma.com/tools-templates/measurement-systems-analysis-msa-gage-rr/making-sense-**a**ttribute-gage-rr-calculations/*
4. Guerra, A. S., Pillet, M., Maire, J. L., Control of variability for man measurement, *12th IMEKO TC1-TC7 joint Symposium on Man, Science and Measurement*, pp. 45-50, Annecy, France (2008).
5. Minitab 16 Statistical Software, State College, PA: Minitab, Inc., 2010.