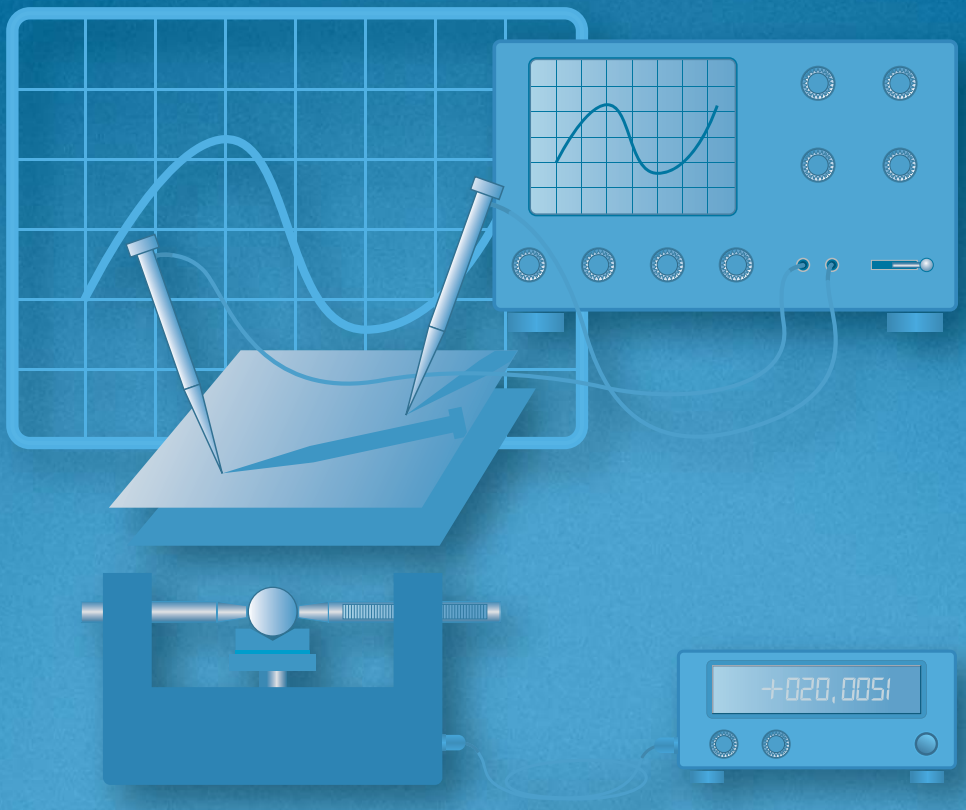
2020-04-06 - SOCOS



Quality Management in the Bosch Group | Technical Statistics

10. Capability of **Measurement and Test Processes**

BOSCH
Invented for life

Page intentionally left blank

Quality Management in the Bosch Group
Technical Statistics

Booklet 10

Capability of Measurement and Test Processes

Edition November 2019

2020-04-06 - SOCOS

Edition November 2019

The minimum requirements for capability criteria given in this booklet correspond to the requirements at the time of publishing the present edition. For the determination of current minimum requirements for capability criteria the CDQ 0301 is binding.

Contents

List of flow charts

2020-04-06 - SOCOS

1 Introduction

The procedures described in the present booklet are part of the Bosch quality management. They are incorporated by means of the management manual [MM] and the central directive [CDQ 0301].

The present issue of Booklet 10 complies with the requirements according to [AIAG MSA]. Procedure 2 directly conforms to [AIAG MSA].

In the present booklet, standardized terminology is preferably used which also ensures unambiguousness in a legal case due to its definition and international acceptance (see chapter *"Definition of terms"*).

NOTE: The definitions of various standardized terms only differ marginally in currently applicable standards, i.e. the terms are used synonymously in practice. In these cases, terms are used in this document which are commonly used in the respective context (example: measuring system, measuring equipment, measuring instrument).

Verification of capability and monitoring of stability of measurement processes are done to ensure that a measuring system can measure a quality characteristic at the place of operation with sufficiently low systematic measurement error and variation of the measured values (related to the tolerance of the characteristic). The available procedures for continuous (variable) characteristics are complemented by procedures for the assessment of test processes for discrete (attributive) characteristics. A comprehensive description of numerous special procedures is beyond the scope of this booklet.

NOTE: Procedures for multi-dimensional (multivariate) characteristics are added, when the applicable ISO standard is available (in preparation at the time of publication of the present issue of booklet 10).

This booklet is divided into the chapters 1 – 6 containing the essential minimum information for every user and the appendix. The appendix contains notes and amendments and – as far as a demand could be recognized due to re-occurring inquiries – explanation of theoretical background that demands more in-depth mathematical knowledge. Thus, the appendix is preferably targeted at readers with corresponding information needs.

For measuring and testing, repeatably measurable or testable measurement standards and serial parts or reference parts are required as measuring or test objects. If a measurement or test process is shown to be non-capable, the causes have to be investigated. Here, systematic and random errors of the measuring or test system as well as the influence of measuring and test objects and operators have to be determined. Measuring aids, fixtures as well as the measurement and testing strategy and environmental conditions also have an effect.

Statistical analyses are performed using a suitable statistics software (e. g. solara.MP®). Measurement results from procedure 2 and 3 should preferably be analyzed using analysis of variances (ANOVA). Results may deviate from the results obtained from software-supported analysis and from the evaluation examples shown, if the analysis (as an exception) is done manually and the intermediate results are rounded, which should generally be avoided, and/or if the outdated average range method (ARM) is used.

2 Scope

Thoroughly and professionally performed and documented test planning is a prerequisite [CDQ 0301].

The verification of capability has to be provided by means of measurements and tests at the place of operation of the measuring or test systems and statistical analyses of the results. It is only reasonable for measuring and test systems that conduct a sufficiently large number of similar recurring measurements and tests (e. g. in the production flow) and it is valid for the examined characteristic only. If measurements and tests of different characteristics are done with the same measuring or test system, an individual verification of capability is required for each characteristic.

Verification of capability for measurement processes for continuous (variable) characteristics:

- Generally, it is a pre-requisite that the capability criteria according to procedure 1 (type-1 study) are met in order to perform one or more of the procedures 2 – 5.
- If operator influence is possible measurement process capability must normally be verified with procedure 1 together with procedure 2 (type-2 study). If operator influence is not possible, the capability must be verified with procedure 1 together with procedure 3 (type-3 study).
- If the linearity has not been proven sufficiently by the manufacturer or during the regular calibration of the measurement equipment, and if linearity is of special importance for the specific application, a linearity study has to be performed according to procedure 4.
- Procedure 5 is additionally intended for measurement processes with presumably insufficiently stable long-term behavior, since capability results of the procedures 1 – 4 refer to the time when the study was carried out.

In case of frequently changing measurement tasks (e. g. in development and test departments), measurement uncertainties should be determined rather than capabilities. If conformity statements according to [ISO 14253] are required, measurement uncertainties have to be determined categorically instead of or in addition to capabilities.

In order to ensure accuracy and traceability to international measurement standards, measuring and test equipment is subject to an initial inspection (e. g. during incoming inspection). Subsequently, it is subject to the control of inspection, measuring and test equipment [CDQ 1001], so that it is reinspected for systematic measurement errors at specified intervals (e. g. according to [VDI 2618], [VDI 2622]). Correct adjustment according to the manufacturer's specifications is crucial.

Measurement results always include uncertainties. Thus, calibration and control of measuring equipment according to [ISO 10012] requires specifying measurement uncertainties. The measurement uncertainty is determined using other methods, e. g. according to [Booklet 8] or [GUM].

The application of the procedures 1 – 4 is limited or inappropriate in case of some measurands, such as hardness or torque, as well as inhomogeneous measuring objects and product characteristics with only one upper or lower specification limit.

The statistical analyses of the procedures 1 – 5 are based on normally distributed measurement results. Otherwise, the procedures cannot be used directly as described below.

Procedures 6 and 7 are intended for the verification of the capability of test systems for the assessment of discrete (attributive) characteristics.

If procedures contained in this issue of booklet 10 cannot be applied for justified reasons, other procedures according to [AIAG MSA] have to be examined for their applicability and used. If these procedures – either unchanged or modified – cannot be applied, procedures described in the literature have to be examined for their applicability and used. As an exception, special procedures can be developed. The intended procedure has to be documented and agreed upon with the QM department and the customer.

NOTE: For contractual agreements with suppliers as well as internal or external customers, it is recommended to proficiently examine and specify the applicability of these procedures beforehand. Imprecise all-inclusive agreements such as "Verification of capability according to booklet 10" are not recommended.

Notes on documentation

Each capability study requires a corresponding documentation that may exceed the information contained in the standardized forms for the respective analyses.

Minimum information required:

- Unambiguous identification of the test plan (e. g. ID number, title, release, date) which includes an exact description of the measuring or test system, the measurement or test method, the measurement or test position, etc.
- Date and time of beginning and completion of measurements and tests, corresponding ambient temperature as possible, and humidity, air pressure, light intensity for visual inspection, etc. as needed;
- Unambiguous identification (e. g. ID number) of the calibration certificate of the measurement standard and/or the reference value, the uncertainty of calibration, the date of last calibration, the name of calibration laboratory;
- Identification of operators/appraisers and responsible person(s) either as ID codes or names;

  NOTE: If applicable, plant-specific directives regarding person-related data have to be observed;
- All measurement and test results that were used for the analysis (e. g. in a table);
- Specification limits;
- Information about evaluation strategy (e. g. solara.MP®, Bosch 2005, ANOVA), calculation formulae
- Results of evaluation (e. g. capability indexes) and classification (e. g. "capable", "not capable");
- Special incidents during the capability study, if applicable.

It must be ensured that this information is allocated unambiguously to each capability study and accessible if required. If some of this information is not contained in the standardized forms (e. g. if corresponding fields are missing), it has to be documented unambiguously using fields such as "Notes" or "Comments". Alternatively, clear references to separate documents containing this information can be entered in these fields (e. g. ID number, title, release, date).

NOTE: The standardized forms presented in this booklet comply with the version applicable at the time of publication of this issue of booklet 10. The up-to-date versions may have been upgraded and thus deviate from these forms (e. g. in the software solara.MP®*).*

3 Flow Charts

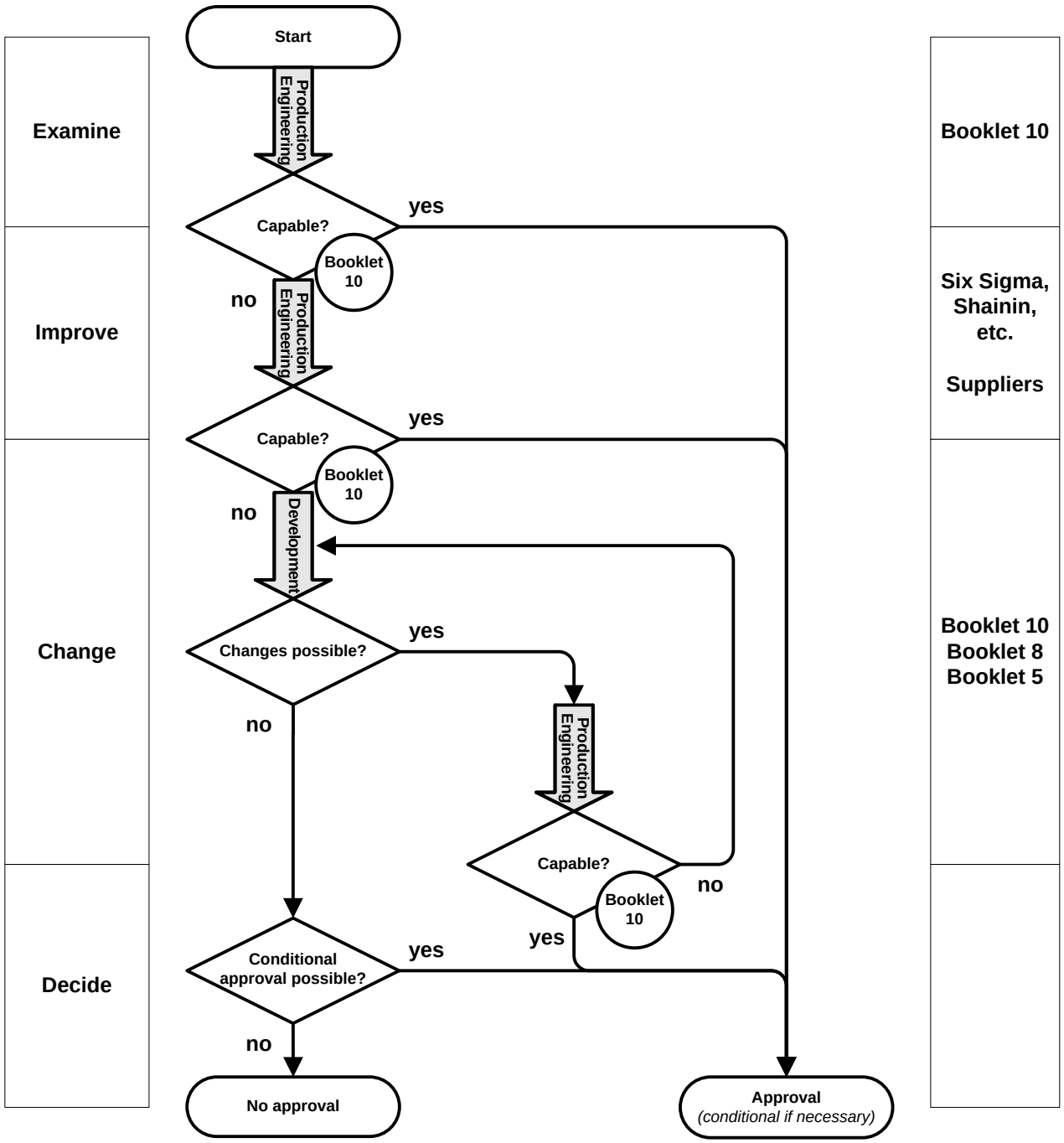


Fig. 1: Capability study – overview of stages and responsibilities

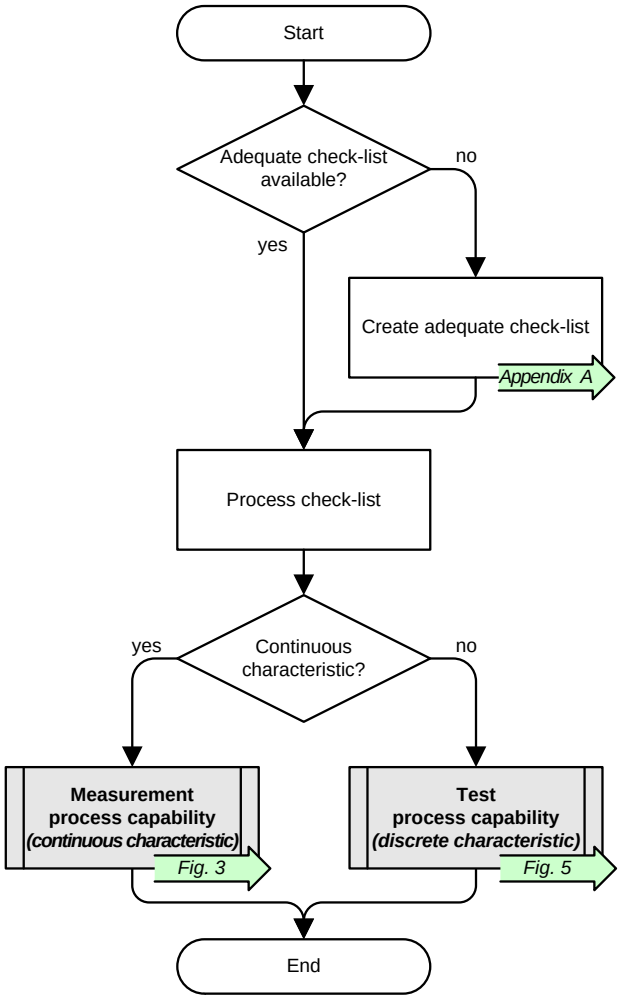


Fig. 2: Superordinate work flow of capability studies for measurement and test processes (level 1)

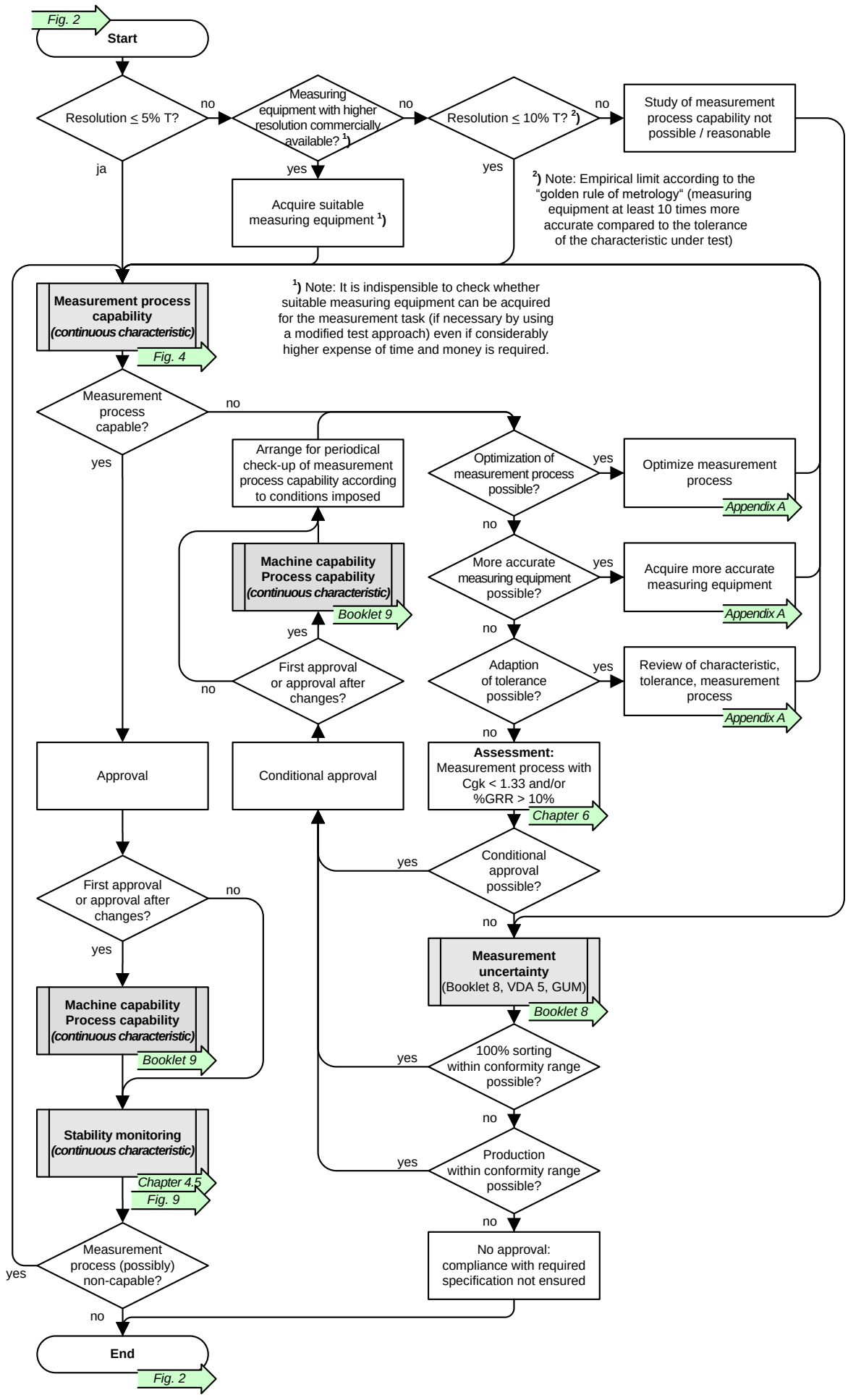


Fig. 3: Level-2 work flow of a capability study of measurement processes for continuous characteristics

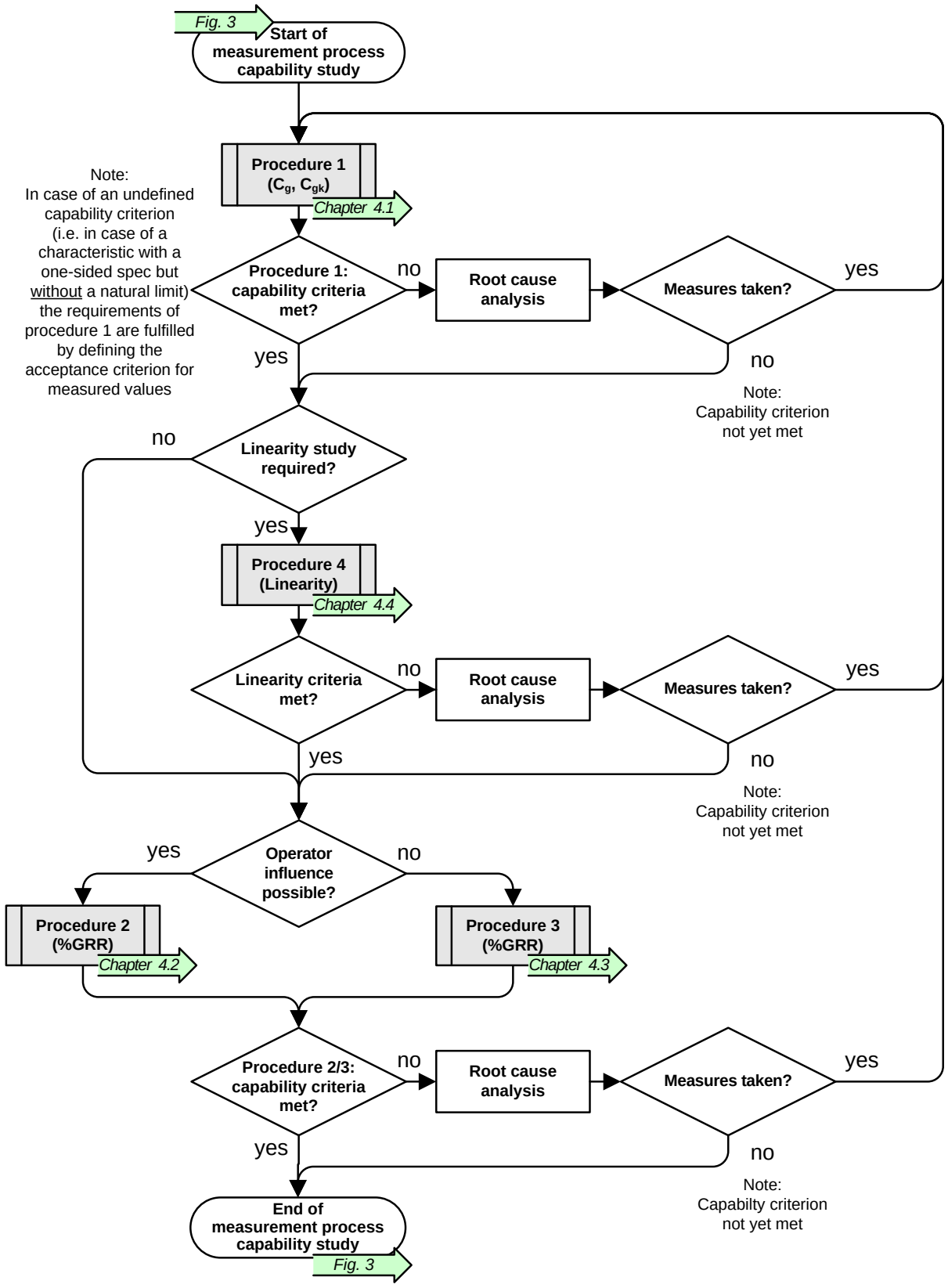


Fig. 4: Level-3 work flow of a capability study of measurement processes for continuous characteristics

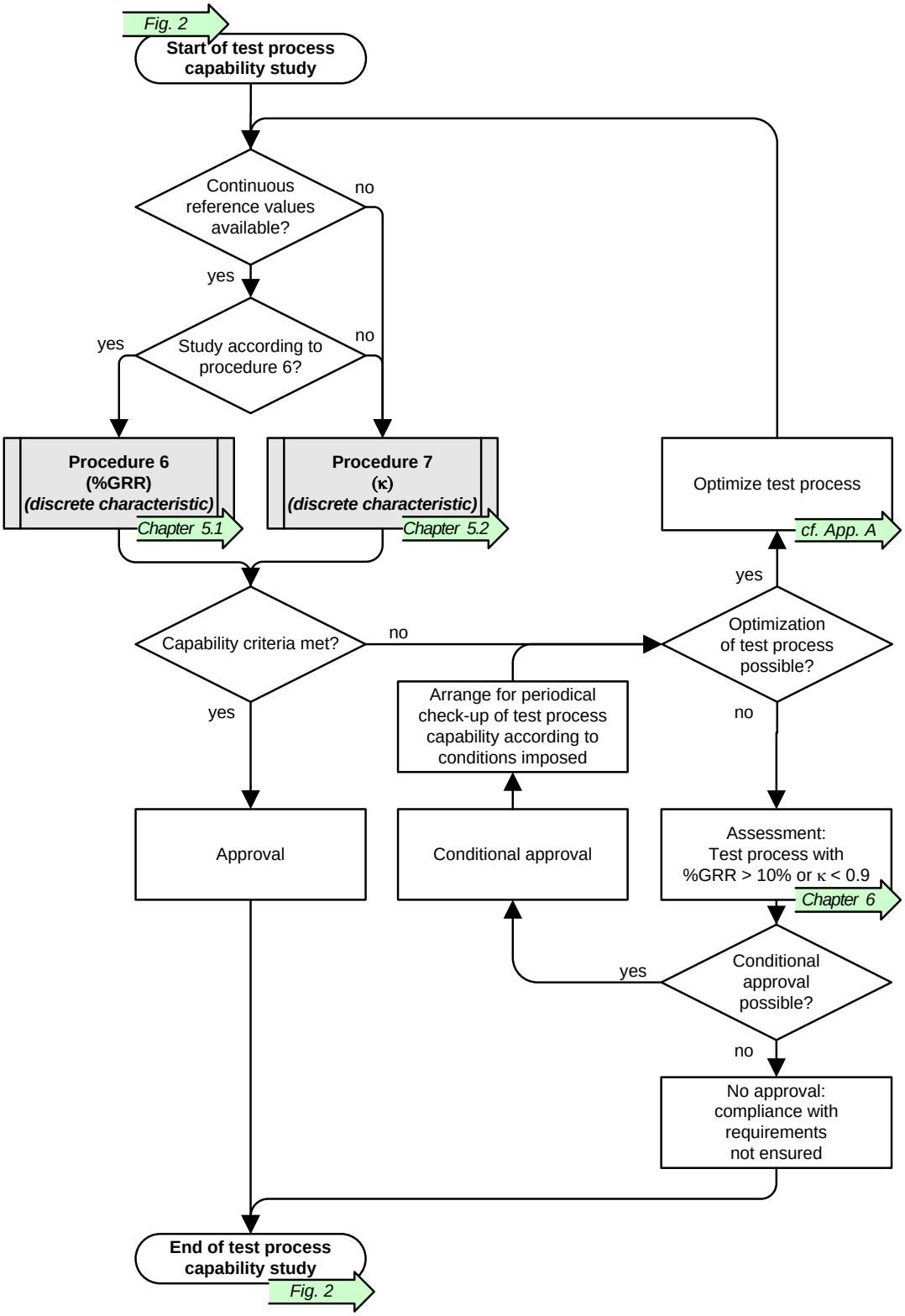


Fig. 5: Level-2 work flow of a capability study of test processes for discrete characteristics

NOTE: See chapter 5.3 for notes on stability monitoring in case of discrete characteristics.

4 Procedures for Verification of Measurement Process Capability by Means of Continuous Characteristics

4.1 Procedure 1 (type-1 study): Systematic measurement error and repeatability

Objective

Verification of the capability of a measurement process (as a test process for a particular characteristic) in terms of location and variation of measured values within the tolerance field of this characteristic.

NOTE 1: Procedure 1 is not part of [AIAG MSA] but a consistent upgrade of the analysis of systematic measurement errors described in [AIAG MSA]. It includes the minimum requirements according to [AIAG MSA]. Procedure 1 – in slightly varying versions – is part of the internal guidelines of numerous automotive manufacturers and is demanded by them.

NOTE 2: Procedure 1 has to be used before procedure 2 or 3, respectively. If there are several measuring systems that are identical in construction and if capability according to procedure 1 was already proven for one of these systems, it must be decided whether procedure 1 is required for the other measurement systems as well.

Requirements

Procedure 1 requires product characteristics with two-sided specification limits, i.e. with a lower and an upper limiting value (LSL and USL), so that the tolerance (T = USL – LSL) is defined. For characteristics with one-sided specification limits, i.e. with only one specified limiting value (LSL or USL) but a lower or upper *natural* limit (LSL* or USL*), the parameter T* = USL – LSL* or T* = USL* – LSL is used instead.

NOTE: A natural limit is defined as a limit that basically cannot be underrun or overrun for physical reasons. For example, the width of a joint or the roughness of a surface cannot become smaller than 0 so that 0 is a natural limiting value LSL = 0.*

However, if there is only one specification limit and no *natural* limit, then there is neither a tolerance T nor a parameter T*, i.e. the parameters C_g and C_{gk} cannot be calculated (see following paragraph *"Notes on procedure 1"* on how to proceed in this case).

Description of the procedure

Procedure 1 is carried out using a calibrated measurement standard that is measured 50 times but at least 25 times. If possible, the reference value x_m of the measurement standard should be in the middle of the tolerance range T of the characteristic that is to be measured with the measuring system. Measurements according to procedure 1 should be carried out at test points specified e. g. in the test plan.

From the measured values, the deviation from the reference value $\bar{x} - x_m$ (systematic measurement error, bias) and the standard deviation s of the measured values are calculated. From these results the capability indexes C_g and C_{gk} are calculated.

Requirements for the measurement standard

The measurement standard must be long-term stable and provide an unambiguous measurement result in case of measurements carried out under repeatability conditions. It must have the same characteristic as the production parts to be measured later with the measuring system. The measurement standard may be produced from a production part. It must be labeled clearly as a measurement standard, correctly calibrated and included in the control of inspection, measuring and test equipment. The calibration provides the traceable conventional true value for the respective characteristic of the calibrated part (see [CDQ 1001]).

The uncertainty of calibration U_{cal} of the measurement standard which is usually documented in the calibration certificate should be considerably smaller than the tolerance T of the product characteristic to be measured (rule of thumb for the ideal case: $U_{cal} < 0.01 \cdot T$; at least $U_{cal} < 0.1 \cdot T$ should be met).

If an adequate measurement standard is unavailable or if the calibration of such a standard is impossible, a capability study according to procedure 1 cannot be carried out (see chapter 2, last paragraph).

Conducting data collection

Data must be collected so that it best reflects the reality of later measurements at production parts. All influencing factors that take effect during production (except the influence of part variation) should – as far as possible – also take effect during the measurements according to procedure 1. The device settings (e. g. measuring sensor, sensor pressure, measuring range, analysis parameters) and measurement accessories should preferably be identical with serial measurements. It must be also ensured that all working steps between the individual measurements of the measurement series are done completely. That means that the measurement standard has to be removed from the clamping and re-inserted before

each measurement. Deviations from the described procedure are acceptable in justified cases. The reasons have to be documented in the records of the capability study. All parameters and settings have to be documented as well.

Analysis

Data to be analyzed:

Tolerance of characteristic to be measured:	$T = USL - LSL$	Reference value of measurement standard:	x_m
Number of measured values (sample size):	n	Measured values:	x_i (i = 1 ... n)

Required calculations:

Mean of measured values:	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	Standard deviation of measured values:	$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
Potential capability index:	$C_g = \frac{0.2 \cdot T}{6 \cdot s}$	Critical capability index:	$C_{gk} = \frac{0.1 \cdot T - \lvert \bar{x} - x_m \rvert}{3 \cdot s}$

The analysis is preferably carried out and documented by means of suitable statistics software (e. g. solara.MP®).

Capability criterion

Compliance with specified minimum values for C_g and C_{gk}. The current release of [CDQ 0301] is binding for these minimum values. At the time of publication of the present issue of booklet 10 the following limits apply: $C_g \geq 1.33$ and $C_{gk} \geq 1.33$.

Notes on procedure 1

- [AIAG MSA] recommends checking for significance of the systematic measurement error $\lvert \bar{x} - x_m \rvert$. For details see appendix C.1.
- For characteristics with an one-sided specification limit and **without** a natural limit (LSL* or USL*), only the systematic measurement error $\bar{x} - x_m$ and the standard deviation s are calculated. These results are used to define the acceptance range for **each individual measured value z** which is measured later **during the production process**:

  $z \leq USL + (\bar{x} - x_m) - 4 \cdot s = USL_0$ for characteristics with an one-sided upper limit,

  $z \geq LSL + (\bar{x} - x_m) + 4 \cdot s = LSL_0$ for characteristics with an one-sided lower limit.

  This means for practical application (e. g. during production) that the (usually) smaller critical limit USL_0 has to be used instead of USL, or the (usually) greater critical limit LSL_0 instead of LSL.

  NOTE 1: The exact position of the reference value x_m of the measurement standard is not relevant. However, it is recommended that a standard with x_m close to the respective limiting value should be used, deviation from USL or LSL approximately 10% ($x_m \approx 0.9 \cdot USL$ ... $1.1 \cdot USL$ or $x_m \approx 0.9 \cdot LSL$ ... $1.1 \cdot LSL$).

  NOTE 2: It is strongly recommended for this type of characteristic, to check additionally for significance of the measurement error (see Appendix C.1) and maybe also for linearity (see Appendix E). This applies in particular, if strongly dispersive measurement results have to be expected (e. g. in case of tear-off forces) as well as in case of customer requirements concerning exact procedures according to [AIAG MSA].

  NOTE 3: It must be made sure that the terms $\bar{x} - x_m$ are included in the calculation with their algebraic signs (but not their absolute values). Due to the one-sided limit, it is relevant whether the reference value x_m of the measurement standard (i.e. the conventional true value) is smaller or larger than the mean value of the measured values x_i.

  NOTE 4: In case of an insignificant systematic measurement error the term $\bar{x} - x_m$ can be omitted in the acceptance criteria.

  NOTE 5: The terms 4s in the acceptance criteria represent requirements analogous to $C_g \geq 1.33$ and $C_{gk} \geq 1.33$. For higher requirements such as $C_g \geq 1.67$ and $C_{gk} \geq 1.67$, the terms 4s have to be replaced by 5s, and for $C_g \geq 2.00$ and $C_{gk} \geq 2.00$, they have to be replaced by 6s.

  NOTE 6: A form for documentation is provided in appendix B, page 42 ff.

  See Appendix C.3 for details concerning the determination of acceptance criteria.

Flow chart

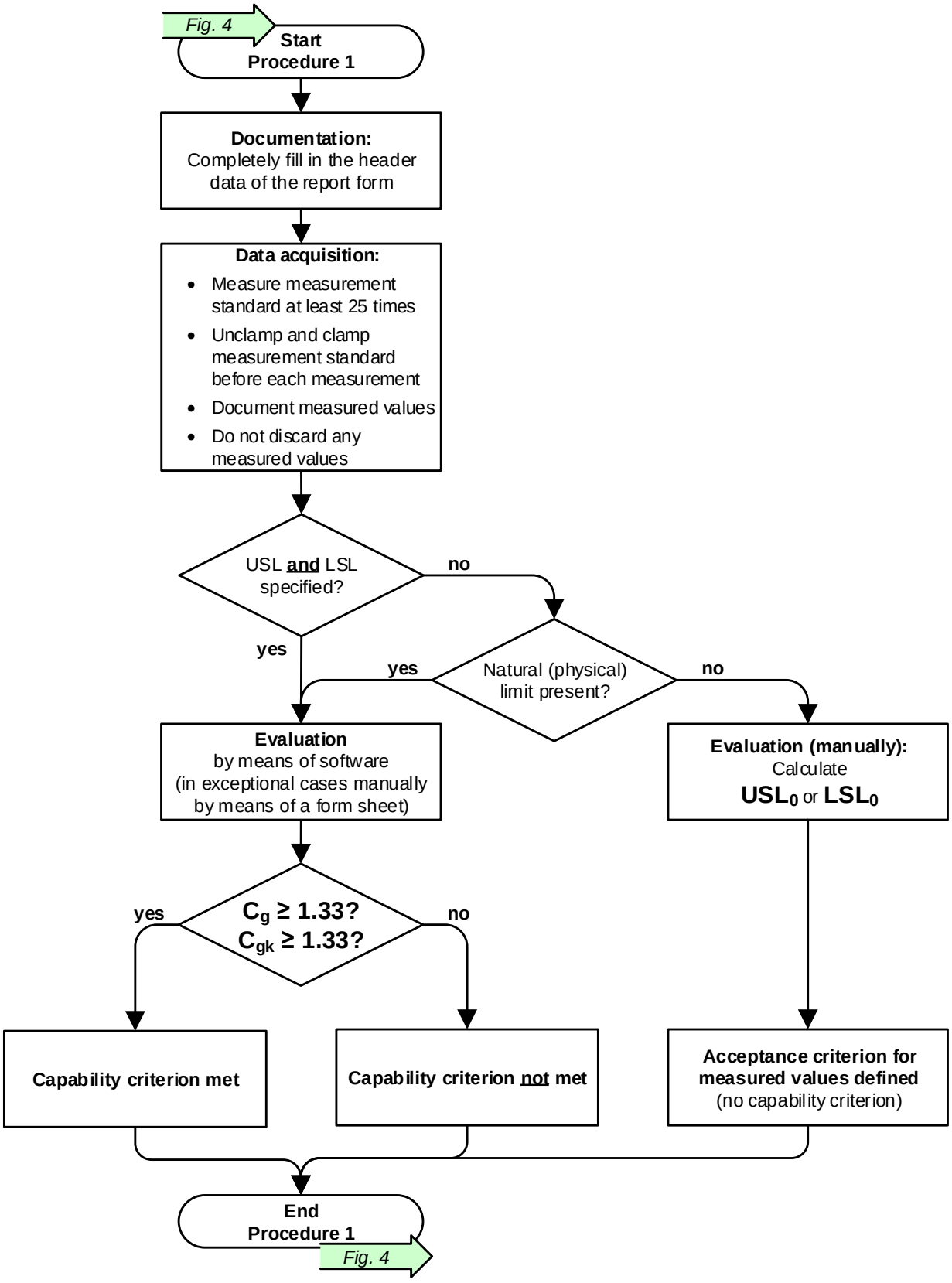


Fig. 6: Level-4 work flow of a capability study according to procedure 1 (type-1 study)

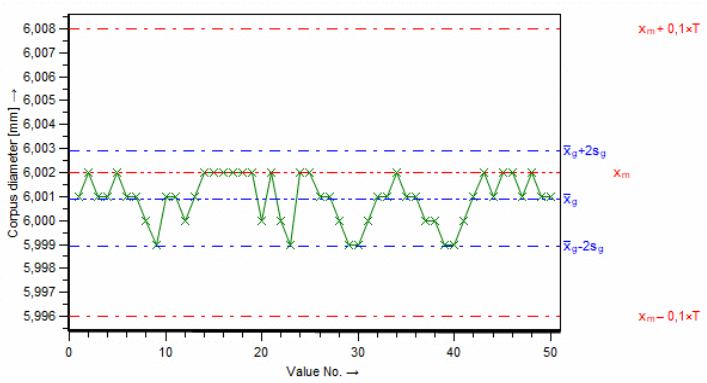
BOSCH	Measurement System Analysis Procedure 1	Sheet 1 / 1

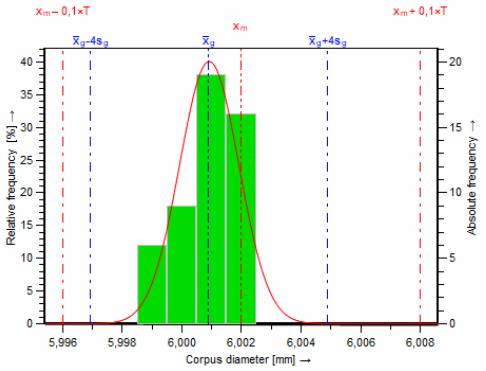
Area	:	MSE3	Operation	:	Mating corpus/needle	Characteristic	: Corpus diameter
Group/Dptm.	:	MOE7	Machine	:	PAKO 9	Char. No.	: 1
Workshop/sect.	:	W450	Machine No.	:	1003521	Nominal value	: 6.000
Product	:	Düse	Test station	:	JML0583W001	Lower allowance	: -0.030
Part	:	Lochdüse	Gage	:	JML0583W003	Upper allowance	: 0.030
Article number	:	0 433 171 914	Gage No.	:	6702779470004	Tolerance	: 0.060
Change status	:	20.01.2019	Gage Manuf.	:	BaP	Unit	: mm
			Resolution	:	0.001		

Comment : Manual operation; room temperature 20.2 °C

Standard: LY_0010W134#95 Standard No.: 6702780329 Standard/Ref. value: 6.002 Calibr. uncertainty: 0.0002

i	x_i	i	x_i	i	x_i	i	x_i	i	x_i
1	6.001	11	6.001	21	6.002	31	6.000	41	6.000
2	6.002	12	6.000	22	6.000	32	6.001	42	6.001
3	6.001	13	6.001	23	5.999	33	6.001	43	6.002
4	6.001	14	6.002	24	6.002	34	6.002	44	6.001
5	6.002	15	6.002	25	6.002	35	6.001	45	6.002
6	6.001	16	6.002	26	6.001	36	6.001	46	6.002
7	6.001	17	6.002	27	6.001	37	6.000	47	6.001
8	6.000	18	6.002	28	6.000	38	6.000	48	6.002
9	5.999	19	6.002	29	5.999	39	5.999	49	6.001
10	6.001	20	6.000	30	5.999	40	5.999	50	6.001





Drawing Values			Collected Values			Statistics		
$x_m + 0.1 * T$	=	6.008	$x_{max\ g}$	=	6.002	$\bar{x}_g + 3 * s_g$	=	6.00388
x_m	=	6.002	$\lvert B_i \rvert$	=	0,001100	$\bar{x}_g$	=	6.00090
$x_m - 0.1 * T$	=	5.996	$x_{min\ g}$	=	5.999	$\bar{x}_g + 3 * s_g$	=	5.99792
0.2 * T	=	0.012	R_g	=	0.003	6 s_g	=	0.00597
T	=	0.060	n_{ges}	=	50	s_g	=	0.00099488
Einheit	=	mm						

Test for Bias — Test result: significant ($\alpha \leq 0.1$ %)

				Minimum reference for capable meas. system		
	C_g =	1.61 ≤ 2.01 ≤ 2.41	0 1.33	$T_{min\ (Cg)}$	=	0.039701
	C_g =	1.30 ≤ 1.64 ≤ 1.98	0 1.33	$T_{min\ (Cgk)}$	=	0.050696
Resolution	%RE =	1.67 %	0 5	$T_{min\ (RE)}$	=	0.020000

Measurement system capable (RE, C_g, Cgk)

Bosch 2018 — Procedure 1

2020-04-06 - SOCOS

4.2 Procedure 2 (type-2 study): Repeatability and reproducibility (gage R&R) with operator influence

Objective

Verification of the capability of a measurement process (as a test process for a particular characteristic) in terms of its variation behavior using measurements of serial parts.

Requirements

Before using procedure 2, it should be checked whether operator influence on the measurement results has to be expected actually or whether procedure 3 should be applied (which is an alternative for procedure 2 in case of absent operator influence, see chapter 4.3). For example, operator influence has to be expected if

- the measurement is done manually (e. g. measurements with a caliper),
- the measurement procedure is not done automatically (e. g. as CNC program),
- the analysis of raw data is influenced by the operator (e. g. determination of validity ranges of a measured contour profile),
- a clamping device is not present which ensures placement of the measuring objects in an unambiguously reproducible position,
- a clamping device is present but the clamping forces are dependent on the force that the operator exerts to operate the clamping device.

A clear definition that is generally applicable to all practical situations is not possible. A decision must generally be made for the individual situation.

Description of the procedure

A type-2 study is done using at least 10 ($n \geq 10$) repeatably measurable and randomly selected serial parts as measuring objects. The characteristic values of these parts should preferably lie within the tolerance range. All factors should take effect that also will take effect during operation of the measuring system in series production. The selected serial parts are measured in random order by at least three ($k \geq 3$) operators in at least two ($r \geq 2$) measurement series under repeatability conditions (e. g. at the test points defined in the test plan, see also chapter *"Definition of terms"*). After completion of the first measurement series, each operator measures again the same serial parts in random order. If further measurement series are intended, the procedure is repeated in the same manner until all measurement series are completed. The next series must not be started before the preceding series has been completed. The measurement results have to be documented.

The measurement results are preferably analyzed by means of a statistics software (e. g. solara.MP®) using the ANOVA method (see Appendix D.2). Manual analyses with forms using the average range method (ARM, see Appendix D.3) as well as corresponding ARM analyses by means of software are no longer up-to-date and are generally not recommended.

If appropriate serial parts are unavailable for the measurements, the procedure cannot be applied. Suitable special procedures are required instead which have to be documented (see chapter 2, last paragraph).

Capability criterion

Compliance with the specified limiting value for the variation %GRR of the measurement process. The up-to-date release of [CDQ 0301] is binding for this limiting value. At the time of publication of the present issue of booklet 10, the following limits apply:

- $\%GRR \leq 10\%$ measurement process is capable (as a test process),
- $10\% < \%GRR \leq 30\%$ measurement process is conditionally capable (as a test process),
- $\%GRR > 30\%$ measurement process is not capable (as a test process).

NOTE: The reference value for %GRR is the tolerance T of the measured characteristic, i.e.

$$\%GRR = \frac{6 \cdot GRR}{T} \cdot 100\%;$$

also see the following notes and Appendix D.2 for calculations.

A type-2 study resulting in a non-capable measurement process is not necessarily due to the measuring system. For example, it may also be caused by the inhomogeneity of the characteristic of the production parts. An appropriate analysis is required.

Flow chart

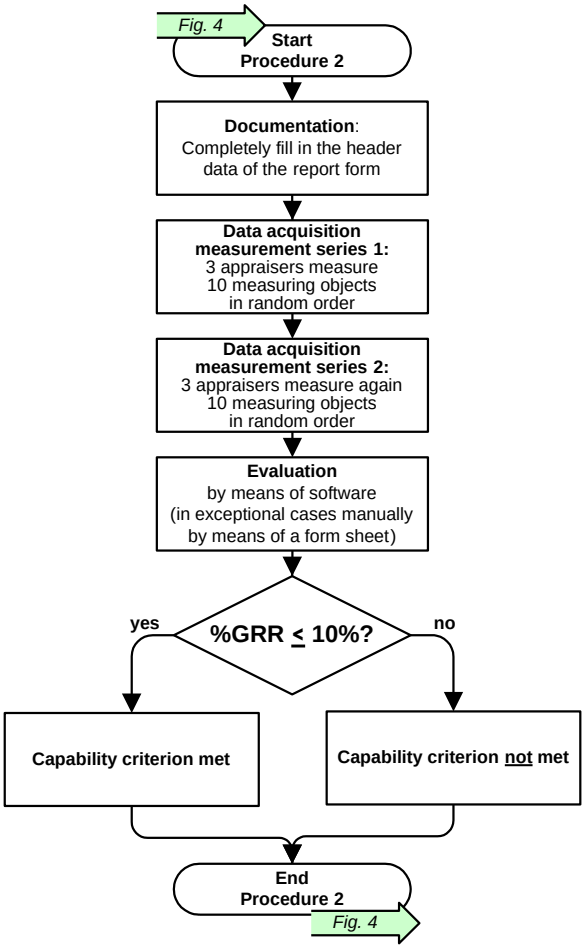


Fig. 7: Level-4 work flow of a capability study according to procedure 2 (type-2 study)

Notes

- [AIAG MSA] recommends three (r = 3) measurements per serial part.
- The tolerance T is not defined for characteristics with one-sided specification limits. If there is a natural (i.e. physical) lower or upper limit in addition to the specified limit, it must be examined if the parameter T* can be used instead of the tolerance T (see chapter 4.1, paragraph *"Requirements"*). If this is not the case or if there is not a natural limit, GRR is related to the **t**otal **v**ariation TV (for a definition, see Appendix D.2):

$$\%\mathrm{GRR} = \frac{\mathrm{GRR}}{\mathrm{TV}} \cdot 100\% = \frac{\mathrm{GRR}}{\sqrt{\mathrm{GRR}^2 + \mathrm{PV}^2}} \cdot 100\% \,.$$

- [AIAG MSA] *generally* recommends the total variation TV as reference value for GRR.
- [AIAG MSA] recommends the parameter ndc (**n**umber of **d**istinct **c**ategories) as an additional capability criterion which should not become smaller than 5 (for details, see Appendix D.1):

$$\mathrm{ndc} = \sqrt{2} \cdot \frac{\mathrm{PV}}{\mathrm{GRR}} = 1.41 \cdot \frac{\%\mathrm{PV}}{\%\mathrm{GRR}} \geq 5 \,.$$

- For the above mentioned recommendations according to [AIAG MSA], it must absolutely be considered if there are customer requirements concerning exact compliance with the recommendations according to [AIAG MSA]. If in doubt, these recommendations should be complied with.
- In exceptional cases %GRR can be determined using a lower number of measuring objects (e. g. if a measuring system is acquired before start of production and the number of available samples is not sufficient at the time of approval). The causes have to be documented. In this case, the number of measurement series must be adapted (for details see Appendix D.5).

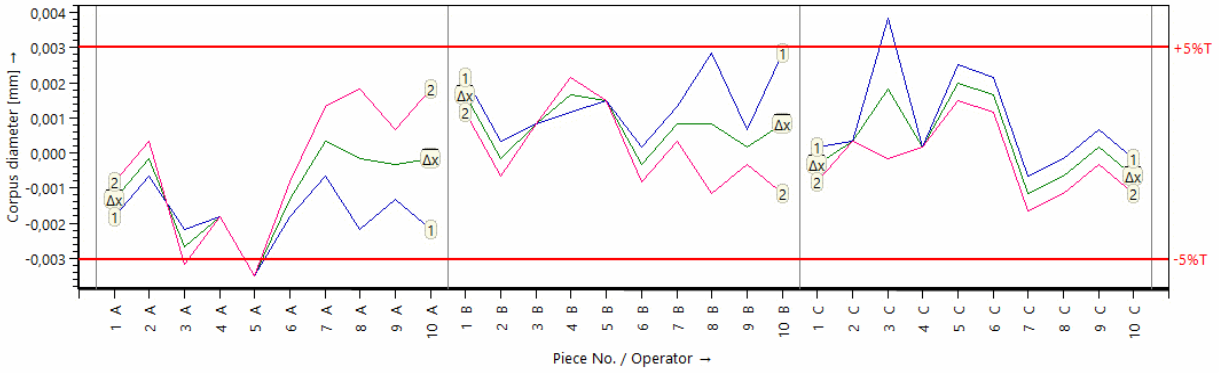
Measurement System Analysis Procedure 2

Sheet 1 / 1

Area	:	MSE3	Operation	:	Mating corpus/needle	Characteristic	:	Corpus diameter
Group/Dptm.	:	MOE7	Machine	:	PAKO 8	Char. No.	:	1
Workshop/sect.	:	W450	Machine No.	:	1004537	Nominal value	:	6.000
Product	:	Injector	Test station	:	JML0583W004	Lower allowance	:	-0.030
Part	:	Hole type nozzle	Gage	:	JML0563W001	Upper allowance	:	0.030
Article number	:	0 433 171 914	Gage No.	:	6702779430001	Tolerance	:	0.060
Change status	:	20.01.2019	Gage Manuf.	:	BaP	Unit	:	mm
			Resolution	:	0.001			

Comment : Manual operation; room temperature 20.2 °C

Standard: Standard No. Standard/Ref. value:



i	$x_{A;1}$	$x_{A;2}$	$x_{g,j}$	$s_{g,j}$	$x_{B;1}$	$x_{B;2}$	$x_{g,j}$	$s_{g,j}$	$x_{C;1}$	$x_{C;2}$	$x_{g,j}$	$s_{g,j}$
1	6.029	6.030	6.0295	0.0007	6.033	6.032	6.0325	0.0007	6.031	6.030	6.0305	0.0007
2	6.019	6.020	6.0195	0.0007	6.020	6.019	6.0195	0.0007	6.020	6.020	6.0200	0.0000
3	6.004	6.030	6.0035	0.0007	6.007	6.007	6.007	0.0000	6.010	6.006	6.0080	0.0028
4	5.982	5.982	5.9820	0.0000	5.985	5.986	5.9855	0.0007	5.984	5.984	5.9840	0.0000
5	6.009	6.009	6.0090	0.0000	6.014	6.014	6.014	0.0000	6.015	6.014	6.0145	0.0007
6	5.971	5.972	5.9715	0.0007	5.973	5.972	5.9725	0.0007	5.975	5.974	5.9745	0.0007
7	5.995	5.997	5.9960	0,0014	5.997	5.996	5.9965	0.0007	5.995	5.994	5.9945	0.0007
8	6.014	6.018	6.0160	0.0028	6.019	6.015	6.017	0.0028	6.016	6.015	6.0155	0.0007
9	5.985	5.987	5.9860	0,0014	5.987	5.986	5.9865	0.0007	5.987	5.986	5.9865	0.0007
10	6.024	6.028	6.0260	0.0028	6.029	6.025	6.027	0.0028	6.026	6.025	6.0255	0.0007

	Variance	Standard dev.	Confidence level 1 - α = 95%						
Repeatability	0.0000023556	0.0015348	EV:	0.0012799	≤	0.0015348	≤	0.0019174	%EV = 15.35%
Reproducibility	0,00000086806	0.00093169	AV:	0.00035980	≤	0.00093169	≤	0.006229	%AV = 9.32%
Interaction	pooling	pooling	IA:		≤		≤		%IA = ---
R&R	0.0000032236	0.0017954	GRR:	0.0015827	≤	0.0017954	≤	0.0064169	%GRR = 17.95%
Part variation	0.00038084	0.019515	PV:	0.0126070	≤	0.019515	≤	0.036405	%PV = 195.15%
Total variation	0.00038406	0.019598	TV:			0.020			

Test design			Reference quantity		
Number of trials	=	2	Process variation	=	0
Number of appraisers	=	3	Tolerance	=	0.060
Number of parts	=	10	Required Cp value	=	

Resolution	%RE	=	1.67%	0 5
Number of distinct categories (ndc)	ndc	=	15	0 5
Repeatability & Reproducibility	%GRR	=	17.95%	0 10 30
Minimum reference figure f. capable measurem. system	$T_{min\ (\%GRR)}$	=	0.108	
Min. ref. fig. f. conditionally capable measurem. system	$T'_{min\ (\%GRR)}$	=	0.0359	

Measurement system is conditionally capable (RE, %GRR)

Bosch 2018 — Procedure 2

2020-04-06 - SOCOS

4.3 Procedure 3 (type-3 study): Repeatability and reproducibility (gage R&R) without operator influence

Objective

Verification of the capability of a measurement process (as a test process for a particular characteristic) in terms of its variation behavior using measurements of serial parts without operator influences.

NOTE 1: Procedure 3 is only a special case of procedure 2.

NOTE 2: In contrast to procedure 1, procedure 3 includes possible interactions between the measurement procedure and the measuring object in the capability study. It concerns the possibly present influence of the production part variation on the measurement as well as the influence of the measurement on the behavior of the production parts. These interactions, which should be reduced to an unavoidable minimum, can be detected with a measurement standard used in procedure 1 only to a certain extent. If distinct enough, they can result in a proven capability according to procedure 1, but not according to procedure 3.

Requirements

Before using procedure 3, it must be checked thoroughly that any operator influence on the measurement results can definitively be excluded. Usually the operator cannot influence the process if

- the position of the measuring objects is clearly fixed by clamping devices and the clamping force cannot be influenced by the operator,
- the measurement procedure and the subsequent data analysis are done fully automatically without operator influence.

A clear definition that is generally applicable to all practical situations is not possible. A decision must generally be made for the individual situation. If in doubt, use procedure 2 (see chapter 4.2).

Description of the procedure

A type-3 study is done using at least 25 ($n \geq 25$) repeatably measurable and randomly selected serial parts as measuring objects. The characteristic values of these parts should preferably lie within the tolerance range. All factors should take effect that also will take effect during operation of the measuring system in series production. The selected serial parts are measured in random order in at least two ($r \geq 2$) measurement series under repeatability conditions (e. g. at the test points defined in the test plan, see also chapter *"Definition of terms"*). After completion of the first measurement series, the same serial parts are measured again in random order. If further measurement series are intended, the procedure is repeated in the same manner until all measurement series are completed. The next series must not be started before the preceding series has been completed. The measurement results have to be documented.

The measurement results are preferably analyzed by means of a statistics software (e. g. solara.MP®) using the ANOVA method (see Appendix D.2). Manual analyses with forms using the average range method (ARM, see Appendix D.3) as well as corresponding ARM analyses by means of software are no longer up-to-date and are generally not recommended.

If appropriate serial parts are unavailable for the measurements, the procedure cannot be applied. Suitable special procedures are required instead which have to be documented (see chapter 2, last paragraph).

Capability criterion

Compliance with the specified limiting value for the variation %GRR of the measurement process. The up-to-date release of [CDQ 0301] is binding for this limiting value. At the time of publication of the present issue of booklet 10, the following limits apply:

- $\%GRR \leq 10\%$ measurement process is capable (as a test process),
- $10\% < \%GRR \leq 30\%$ measurement process is conditionally capable (as a test process),
- $\%GRR > 30\%$ measurement process is not capable (as a test process).

NOTE: The reference value for %GRR is the tolerance T of the measured characteristic, i.e.

$$\%GRR = \frac{6 \cdot GRR}{T} \cdot 100\%;$$

also see the following notes and Appendix D.2 for calculations.

A type-3 study resulting in a non-capable measurement process is not necessarily due to the measuring system. For example, it may also be caused by the inhomogeneity of the characteristic of the production parts. An appropriate analysis is required.

Flow chart

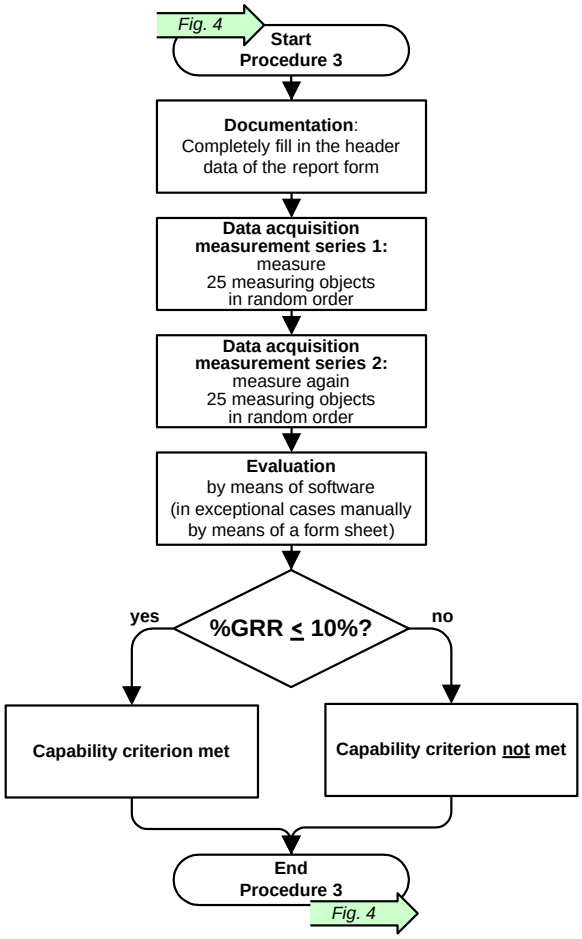


Fig. 8: Level-4 work flow of a capability study according to procedure 3 (type-3 study)

Notes

- [AIAG MSA] recommends three (r = 3) measurements per serial part.
- The tolerance T is not defined for characteristics with one-sided specification limits. If there is a natural (i.e. physical) lower or upper limit in addition to the specified limit, it must be examined if the parameter T* can be used instead of the tolerance T (see chapter 4.1, paragraph *"Requirements"*). If this is not the case or if there is not a natural limit, GRR is related to the **t**otal **v**ariation TV (for a definition, see Appendix D.2):

$$\%GRR = \frac{GRR}{TV} \cdot 100\% = \frac{GRR}{\sqrt{GRR^2 + PV^2}} \cdot 100\% .$$

- [AIAG MSA] *generally* recommends the total variation TV as reference value for GRR.
- [AIAG MSA] recommends the parameter ndc (**n**umber of **d**istinct **c**ategories) as an additional capability criterion which should not become smaller than 5 (for details, see Appendix D.1):

$$ndc = \sqrt{2} \cdot \frac{PV}{GRR} = 1.41 \cdot \frac{\%PV}{\%GRR} \geq 5 .$$

- For the above mentioned recommendations according to [AIAG MSA], it must absolutely be considered if there are customer requirements concerning the exact compliance with the recommendations according to [AIAG MSA]. If in doubt, these recommendations should be complied with.
- In exceptional cases %GRR can be determined using a lower number of measuring objects (e. g. if a measuring system is acquired before start of production and the number of available samples is not sufficient at the time of approval). The causes have to be documented. In this case, the number of measurement series must be adapted (for details see Appendix D.5).

Measurement System Analysis Procedure 3

Sheet 1 / 1

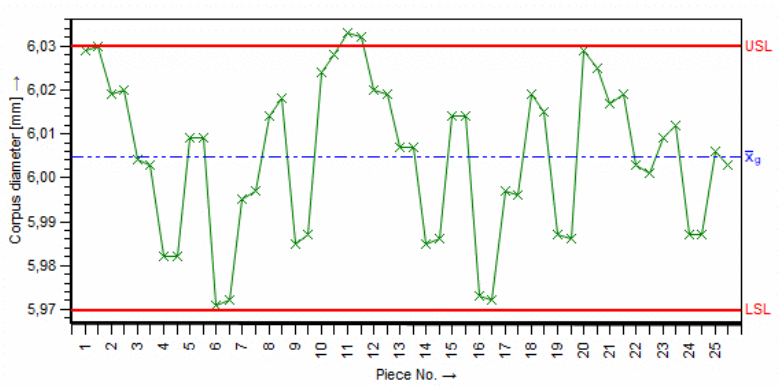
Area	:	MSE3	Operation	:	Mating corpus/needle	Characteristic	: Corpus diameter
Group/Dptm.	:	MOE7	Machine	:	PAKO 9	Char. No.	: 1
Workshop/sect.	:	W450	Machine No.	:	1003521	Nominal value	: 6.000
Product	:	Injector	Test station	:	JML0583W001	Lower allowance	: -0.030
Part	:	Hole type nozzle	Gage	:	JML0563W003	Upper allowance	: 0.030
Article number	:	0 433 171 914	Gage No.	:	6702779430004	Tolerance	: 0.060
Change status	:	20.01.2019	Gage Manuf.	:	BaP	Unit	: mm
			Resolution	:	0.001		

Comment :

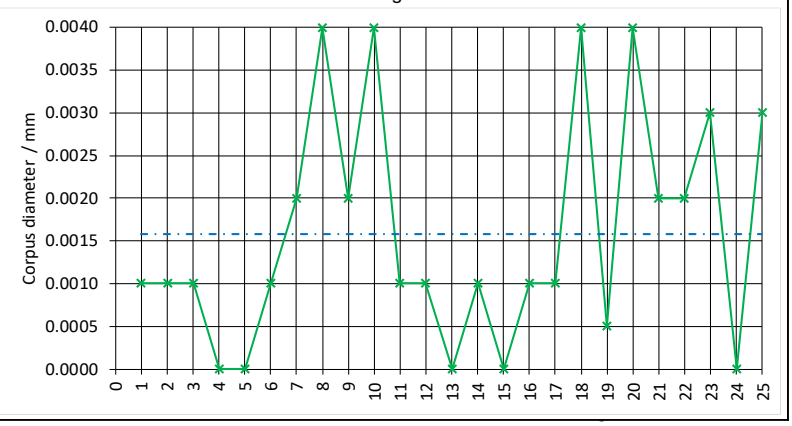
Standard: Standard No. Standard/Ref. value:

i	$x_{A;1}$	$x_{A;2}$	$x_{g,j}$	$s_{g,j}$
1	6.029	6.030	6.0295	0.0007
2	6.019	6.020	6.0195	0.0007
3	6.004	6.003	6.0035	0.0007
4	5.982	5.982	5.9820	0.0000
5	6.009	6.009	6.0090	0.0000
6	5.971	5.972	5.9715	0.0007
7	5.995	5.997	5.9960	0.0014
8	6.014	6.018	6.0160	0.0028
9	5.985	5.987	5.9860	0.0014
10	6.024	6.028	6.0260	0.0028
11	6.033	6.032	6.0325	0.0007
12	6.020	6.019	6.0195	0.0007
13	6.007	6.007	6.0070	0.0000
14	5.985	5.986	5.9855	0.0007
15	6.014	6.014	6.0140	0.0000
16	5.973	5.972	5.9725	0.0007
17	5.997	5.996	5.9965	0.0007
18	6.019	6.015	6.0170	0.0028
19	5.987	5.9865	5.9865	0.0007
20	6.029	6.025	6.0270	0.0028
21	6.017	6.019	6.0180	0.0014
22	6.003	6.001	6.0020	0.0014
23	6.009	6.012	6.0105	0.0021
24	5.987	5.987	5.9870	0.0000
25	6.006	6.003	6.0045	0.0021

Measurement Values



Ranges



	Variance	Standard dev.	Confidence level 1 - α = 95%	
Repeatability	0.0000021600	0.0014697	EV: 0.0011526 ≤ 0.0014697 ≤ 0.0020288	%EV = 14.70%
R&R	0.0000021600	0.0014697	GRR: 0.0011526 ≤ 0.0014697 ≤ 0.0020288	%GRR = 14.70%
Part variation	0.0003133200	0.017701	PV: 0.0126070 ≤ 0.0177010 ≤ 0.0266200	%PV = 177.01%
Total variation	0.0003840600	0.017762	TV: 0,018	

Versuchsplan			Reference quantity		
Number of trials	=	2	Process variation	=	0
Number of parts	=	25	Tolerance	=	0.060
			Required Cp value	=	

Resolution	%RE	=	0.0167	0 5
Number of distinct categories (ndc)	ndc	=	17	0 5
Repeatability & Reproducibility	%GRR	=	14.7%	0 10 30
Minimum reference figure f. capable measurem. system	$T_{min\ (\%GRR)}$	=	0.0882	
Min. ref. fig. f. conditionally capable measurem. system	$T_{min\ (\%GRR)}$	=	0.0294	

Measurement system is conditionally capable (%RE, %GRR)

Bosch 2018 — Procedure 3

2020-04-06 - SOCOS

4.4 Procedure 4: Linearity

Objective

Verification of a sufficiently linear relation between the values of a physical quantity to be measured and the corresponding measured values determined by the measuring system. This procedure determines whether the systematic measurement error of the measuring system is within the acceptable limits regarding the measuring range relevant for the measurement.

NOTE: For an ideal measuring system, the output value (measured value) is always identical with the quantity value supplied at the input (e. g. by a measurement standard). This always applies regardless of the measuring system indicating the output value on a linear or a non-linear (e. g. logarithmic) scale. For example, 5 volts at the input must always be indicated as an output value of 5 volts, 10 volts at the input always as output value 10 volts, etc. The term "linearity" exclusively refers to this relationship between input and output values that can be plotted as a characteristic curve. This relationship is not exactly linear for real measuring systems.

Requirements

Measuring systems are subject to the control of inspection, measuring and test equipment [CDQ 1001]. The linearity of a measuring system (as described above) is usually tested by the manufacturer and subsequently as part of its regular calibration. Thus, an additional check as part of a capability study is usually not required.

However, special applications can require proving sufficiently linear behavior of the measuring system over the entire measuring range which is relevant. Examples are

- adjustable, settable gain (characteristic curve),
- logarithmic scale,
- error limit related to full scale.

Conducting a linearity study

[AIAG MSA] includes a procedure to verify the linearity of a measuring system (as described above). However, this procedure does not provide reliable results under all conditions (see Appendix E.1).

- Unless the procedure according to [AIAG MSA] is explicitly required, procedure 1 can be carried out once for each of several measurement standards with reference values x_i appropriately (e. g. equidistantly) distributed over the relevant measuring range. The type-1 capability of the measuring system has to be proven for each reference value x_i (see chapter 4.1).

  *NOTE 1: This approach is **not** a linearity study in a strict sense. It provides information regarding the capability of the measuring system at the investigated reference points x_i only. There is no information on the intermediate spans. It is recommended to use at least 5 measurement standards with different reference values. For economical reasons, however, it is often impossible to provide more than two measurement standards. In this case, the type-1 studies should preferably be carried out at the limits of the tolerance range.*

  NOTE 2: If measuring equipment is concerned which is part of a measuring system, the applicability of the results to the entire measuring system must be assessed.

- If the procedure according to [AIAG MSA] is explicitly demanded (e. g. due to a customer requirement), Appendix E.1 has to be observed.

  NOTE: If data obtained from several type-1 studies with different measurement standards are already available, these data can be used for the analysis according to [AIAG MSA]. Measuring again is not necessary.

4.5 Procedure 5: Stability

Objective

Validation of consistently correct measurement results by monitoring the long-term behavior of a measurement process and corresponding evaluation of the stability of the measuring system (similar to a $\bar{x} - s$ -SPC control chart whereas a measurement process is not controllable in terms of a SPC process).

NOTE: A series of measurements can be considered a measurement process that "produces" measured values. Thus, the known SPC procedures and rules can be used similarly for measurement processes in order to maintain a permanently mastered state of statistical control (stability over time).

Flow chart

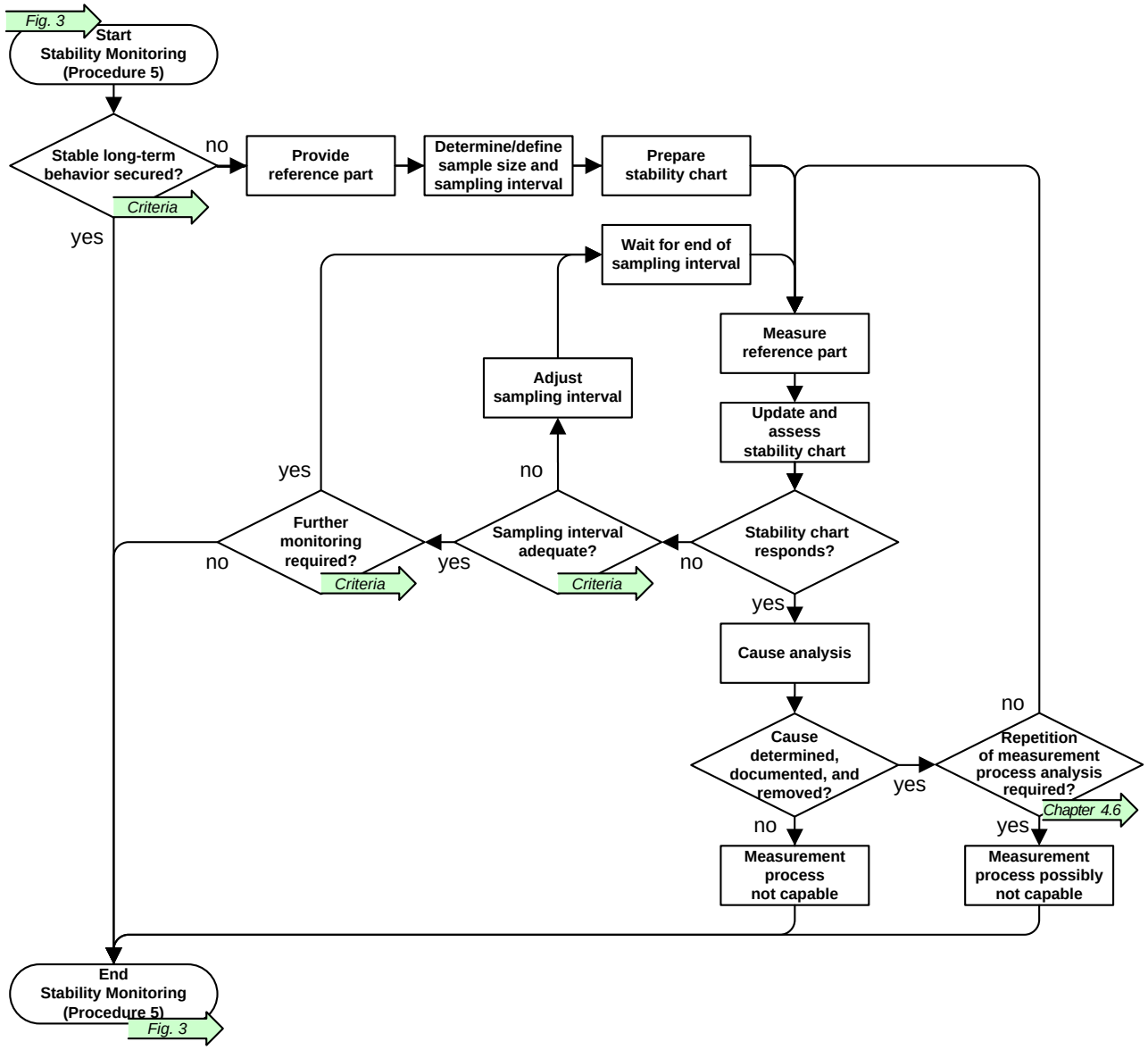


Fig. 9: Stability monitoring for continuous characteristics (procedure 5)

4.5.1 Preparing stability monitoring

Assessment of long-term stability

At first, it must be thoroughly examined if a stable long-term behavior can be expected and if it is sufficiently ensured. The following examples are typical criteria for long-term stability:

- capability indices far above or below the required minimum or maximum values (e. g. $C_{gk} > 2$, %GRR < 5%);
- no (considerable) changes of environmental conditions have to be expected (e. g. temperature, humidity, vibrations);

- no (frequent) change of operating personnel in case of possible operator influence on the measurement process;
- inspection intervals of the control of inspection, measuring and test equipment (i.e. the frequency of calibration and adjustments) are adapted to long-term behavior of the measuring system (e. g. drift);
- comprehensive positive experience with stable long-term behavior of measuring systems which are identical or similar in construction;
- no (confirmed or unconfirmed) defective parts or complaints from the production process regarding the characteristic to be measured;
- no (confirmed or unconfirmed) erroneous measurements; measurement results were synchronized with customer where appropriate;
- no universally used complex measuring system for different measurement tasks and requirements;
- no (considerable) stress of the measuring system concerning wear (e. g. fixtures, clamps, calipers);
- no possibly drifting measuring system (i.e. sensors, adjustable electrical parameters).

If in doubt, a stable long-term behavior has to be ensured by means of stability monitoring.

Reference part (stability part)

To conduct stability monitoring, a reference part (stability part) with known reference value x_m is required. This part can be a measurement standard or a serial part (properly modified, if necessary) that corresponds to the requirements of the standard used in procedure 1 which are relevant for the measurements. When using a serial part, the reference value can be calculated as mean value of at least 10 measurements using a calibrated measuring system. The reference part (stability part) must be clearly labeled.

Sample size

The reference part (stability part) is measured at least three times ($n \geq 3$) in process-specifically specified time intervals (sampling intervals).

For technical and/or economical reasons, it may be necessary to reduce the number of measurements per time interval to less than three ($n < 3$). In these cases, an individual value chart may be maintained alternatively. These exceptions have to be described.

NOTE: [AIAG MSA] does not provide the use of individual value charts.

Control limits for stability charts

	Lower **c**ontrol **l**imit (LCL)	**U**pper **c**ontrol **l**imit (UCL)
$\bar{x}$-chart (mean values):	$LCL = x_m - u_p \cdot \frac{s}{\sqrt{n}}$	$UCL = x_m + u_p \cdot \frac{s}{\sqrt{n}}$
s-chart (standard deviations):	$LCL_s = B'_{Eun} \cdot s$	$UCL_s = B'_{Eob} \cdot s$
Individual value chart:	$LCL = x_m - E'_E \cdot s$	$UCL = x_m + E'_E \cdot s$

For x_m the following values can be used:
- the reference value of the reference part (stability part) or
- the mean value of a previous test run (see [AIAG MSA], chapter 3, paragraph B).

For s the following values can be used:
- 2.5% of the characteristic tolerance T (=T/40) or
- the standard deviation from a previous test run (see [AIAG MSA], chapter 3, paragraph B) or
- the standard deviation from procedure 1 (not recommended because of short-term study).

The sample size is used for n, i.e. the number of measurements per sample.

u_p. B'_{Eun}, B'_{Eob} and E'_E corresponding to the sample size n are taken from the following table for confidence level 99%. For individual value charts, it must be decided how many measured values are combined in one group of size n (pseudo-sample). n = 3 is well-established.

n	u_p	B'_{Eun}	B'_{Eob}	E'_E
3	2.58	0.071	2.302	2.935
4	2.58	0.155	2.069	3.023
5	2.58	0.227	1.927	3.090

Values for further sample sizes and confidence levels can be calculated according to Appendix F.

4.5.2 Sampling interval

A sufficiently appropriate sampling interval is always determined by the respective measurement process and its behavior over time. Thus, generally applicable rules are not possible.

Generally, it must be examined carefully if an initial sampling interval as short as possible or shortening of an already fixed sampling interval is necessary (e. g. testing several samples per shift). The following examples are typical criteria for the requirement of short intervals:

- instable measurement process;
- capability indices at the limits (e. g. C_{gk} around 1.33 and/or %GRR around 10%);
- function-critical / process-critical characteristic;
- new measurement / test methods;
- no empirical data available;
- tests neither time-consuming nor costly;
- high statistical power required.

When the sampling interval is determined or changed, it must always be considered that sufficiently short reaction times must be ensured at any time in order to secure accurately timed part access in the case of an error (i.e. traceability must be ensured).

Adapting the sampling interval: Established procedure during process launch

- All mean values are within the control limits; the variations from value to value are easily recognizable and unsystematic (random): The sampling interval is appropriate; actions are not necessary.
- All mean values are within the control limits, but only small or no variations from value to value are recognizable (see also “Middle third”): The sampling interval could be too short; increase the interval (e. g. double it); repeat adaptation several times, if necessary.
- Some mean values are outside the control limits: The sampling interval could be too long; decrease the interval (e. g. halve it); repeat adaptation several times, if necessary.

If several adaptations of the sampling interval are unsuccessful, cause and risk analyses have to be performed and, if necessary, suitable measures have to be taken.

Adapting the sampling interval: Established procedure during production

- All mean values are within the control limits: A check-up measurement at the beginning of each shift is generally sufficient.
- Some mean values are outside the control limits: Cause and risk analyses have to be performed and, if necessary, suitable measures have to be taken (e. g. calibration, adjustment, overhaul, replacement and, if necessary, followed by a re-determination of the sampling interval as done during process launch).

In case of very small tolerances, it may become necessary to calibrate the measuring system before each measurement. In this case, measurements for stability monitoring are not required.

4.5.3 Conducting stability monitoring

The reference part (stability part) is measured at least three times ($n \geq 3$) in process-specifically specified time intervals (sampling intervals). The measured values are documented in a table on the stability chart; mean value and standard deviation of each sample are calculated and entered in chronological order in the $\bar{x}$-chart or s-chart, respectively.

The $\bar{x}$-chart can be maintained using absolute values or values relative to the reference value x_m, i.e. the differences of the measured values to the reference value (residues) are collected.

Analysis

The stability of a measurement process is evaluated by means of the stability chart.

Stable measurement process

All values (usually the mean values) are within the control limits and vary unsystematically (randomly). There are no indications of instability.

If the measurement process is shown to be stable over a longer period of time according to the stability chart, the sampling interval may be increased (see chapter 4.5.2).

If a measurement process is shown to be stable according to a greater number of subsequent stability charts, stability monitoring may be completed. The following examples are typical criteria that **do not** allow for completion:

- abnormalities during control of inspection, measuring and test equipment or calibration;
- customer requirement for stability monitoring;
- no further validation of the quality requirements for this characteristic;
- function-critical and/or process-critical characteristic (e. g. a special characteristic, risk part);
- changes of measurement setup.

If in doubt, the stability monitoring has to be continued.

Instable measurement process

The values show a large and unsystematical variation over time and some values are outside the control limits.

Indicators of possible problems in the measurement process:

- The values form an unusual (non-random) sequence of points. For identification the so-called 7-point rule can be used, i.e. 7 or more successive mean values
  - are exclusively above or below x_m *(Run)*
  - or form a steadily ascending or descending order *(Trend)*.
- Within the middle third between the control limits
  - are more than 90%
  - or less than 40%

  of all values *(Middle third)*.

If instability and/or another problem are detected, the cause has to be determined. At first, it must be clarified if there is an influencing factor which is either due to the measurement process or due to the measuring object. To do this, another reference part (stability part) should be available.

NOTE 1: Established methods for root cause analyses are e. g. cause-and-effect diagrams / Ishikawa diagrams (5M), 5 x Why (determination of root causes by means of systematic questioning), Kepner Tregoe (KT), Shainin, Six Sigma (DMAIC); see also [EQT].

NOTE 2: Environmental influences (e. g. changes of temperature, humidity, etc.) are among the most common causes of exceeding the limits. These parameters should be documented when preparing the stability chart in order to determine and remove causes quickly and easily.

The cause has to be eliminated. If necessary, the measurement process must be improved and the requirement for a new capability study must be assessed (see chapter 4.6). Cause and taken measures must be documented (e. g. on the back of the stability chart).

4.6 Repetition of capability studies

During productive use, the capability of the measurement process must be ensured at all times (preferably by means of Procedure 5). The following criteria are typical examples that may make a new analysis of the measurement process and a new verification of capability necessary:

- after interventions in the measurement process (e. g. after exceeding control limits), the stability chart shows a significant difference compared to the status before the intervention;
- after adjusting the measuring system or components of the measuring system (e. g. individual measuring instruments during control of inspection, measuring and test equipment);
- upon restart after maintenance work where substantial disassemblies, modifications or replacements of crucial parts were necessary (e. g. measuring sensor, displacement transducer);
- upon start-up of new, overhauled or repaired measuring systems;
- in case of (later) tolerance cutbacks of the characteristic to be measured;
- in case of technical changes of the measuring system (e. g. setup, software);
- in case of parameter changes that may change the capability of the measurement process;
- if basic conditions of the measurement process are changed that may influence the capability of the measurement process (e. g. workflow, measurement strategy);
- after changes of the operating personnel (e. g. new staff members in case of procedure 2);
- if it is suspected that the measuring system does not work properly;
- if necessary before, and definitely after relocation of the measuring system.

In doubt, the measurement process analysis has to be repeated and the capability must be verified again.

Measurement System Analysis Procedure 5 (Stability)

Sheet 1 / 1

Area	: MSE3	Operation	: Mating corpus/needle	Characteristic	: Corpus diameter
Group/Dptm.	: MOE7	Machine	: PAKO 9	Char. No.	: 1
Workshop/sect.	: W450	Machine No.	: 1003521	Nominal value	: 6.000
Product	: Injector	Test station	: JML0583W001	Lower allowance	: -0.030
Part	: Hole type nozzle	Gage	: JML0583W003	Upper allowance	: 0.030
Article number	: 0 433 171 914	Gage No.	: 6702779470004	Tolerance	: 0.060
Change status	: 20.01.2019	Gage Manuf.	: BaP	Unit	: mm
		Resolution	: 0.001		

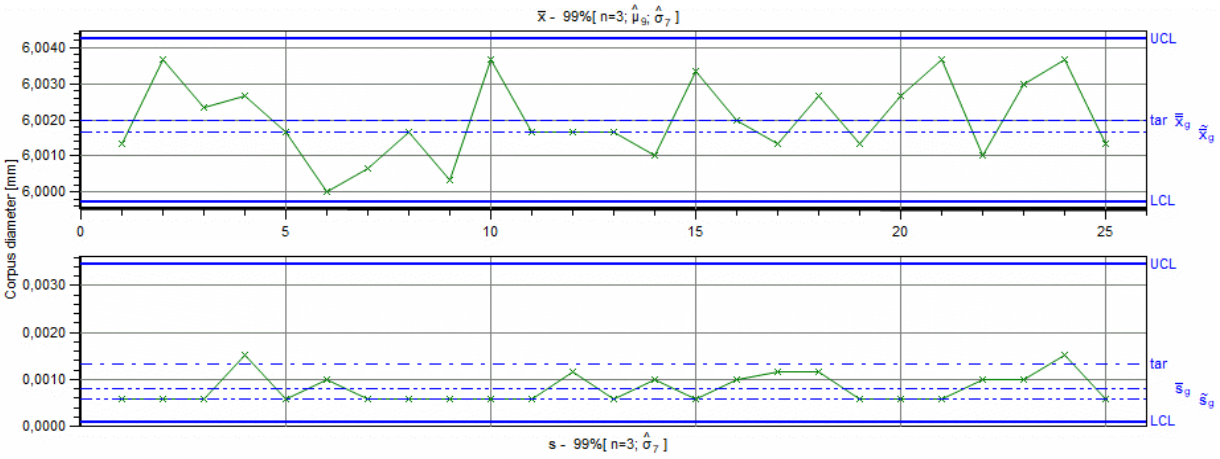
Kommentar :

Standard: LY_0010W134#95 Standard No. 6702780329 Standard/Ref. value: 6.002 Calibr. uncertainty: 0.0002

i	x_i	i	x_i	i	x_i	i	x_i	i	x_i
1	6.002	16	6.000	31	6.002	46	6.003	61	6.004
2	6.001	17	6.001	32	6.001	47	6.002	62	6.003
3	6.001	18	5.999	33	6.002	48	6.001	63	6.004
4	6.004	19	6.001	34	6.003	49	6.002	64	6.002
5	6.004	20	6.001	35	6.001	50	6.002	65	6.000
6	6.003	21	6.002	36	6.001	51	6.000	66	6.001
7	6.003	22	6.001	37	6.002	52	6.002	67	6.004
8	6.002	23	6.002	38	6.001	53	6.002	68	6.003
9	6.002	24	6.002	39	6.002	54	6.004	69	6.002
10	6.003	25	6.000	40	6.002	55	6.001	70	6.005
11	6.001	26	6.000	41	6.000	56	6.002	71	6.004
12	6.004	27	6.001	42	6.001	57	6.001	72	6.002
13	6.002	28	6.004	43	6.004	58	6.003	73	6.002
14	6.001	29	6.004	44	6.003	59	6.003	74	6.001
15	6.002	30	6.003	45	6.003	60	6.002	75	6.001

Drawing Values			Collected Values			Statistics			Quality Control Chart		
T_m	=	6.000				$\bar{\bar{x}}_g$	=	6.00200	xbar-chart, 99 %, n = 3		
LSL	=	5.970	$x_{min\ g}$	=	5.999	s_g	=	6.00090	UCL	=	6.00423
USL	=	6.030	$x_{max\ g}$	=	6.005	n_{eff}	=	75	Target	=	6.00200
T	=	0.060	R_g	=	0.006	k_{eff}	=	25	LCL	=	5.99977
									Violations of control limits: 0		
									s-chart, 99 %, n = 3		
									UCL	=	0.003453
									M	=	0.001329
									LCL	=	0.000106
									Violations of control limits: 0		

Bosch 2018 — Procedure 5, Stability



5 Procedures for Verification of Test Process Capability by Means of Discrete Characteristics

Note

The analysis using discrete or discretized characteristics is generally not recommended, since meeting up-to-date requirements for error rates requires sample sizes which are economically not justifiable. The verification of capability by means of continuous characteristics using procedures 1 - 5 should always be preferred.

5.1 Procedure 6: Test decisions for discretized continuous characteristics

Objective

Verification of the capability of a test process regarding unambiguous test decisions when testing discretized characteristics.

Requirements

The procedure requires continuous reference values.

Description of the procedure

The study is done using a reference lot which comprises 50 reference parts from the production (serial parts). Their discrete characteristic values are determined and documented before starting the study.

At first, the continuous characteristic values of the reference parts (i.e. the reference values) have to be determined by measurement. The measurement uncertainty U, allocated to the measured values, must be known. Reference parts are required whose characteristic values cover a range beginning slightly below LSL – U and ending slightly above USL + U. The measurement result is documented for each reference part.

Next, each reference part is allocated unambiguously to one (of two possible) categories which corresponds to the measurement result (discretizing): *"within tolerance" = "+", "out of tolerance" = "–"*. The discretized results (i.e. the reference ratings) are documented.

Each reference part of the lot must be unambiguously identifiable so that the respective data can always be correctly allocated. This requirement must be implemented so that only authorized personnel but not the test personnel can identify the reference part. Possible implementations are e. g. 2D barcodes, complex number codes, labels only legible under UV light.

To do the study, the reference parts are used as test objects. They are tested under serial conditions in a random order that is unknown to the test personnel using the specified test equipment and test methods (e. g. according to the test plan) or an automatic test system. Each part is allocated to one (of two possible) categories. The test personnel must be adequately trained and instructed.

If the test results (i.e. the ratings) can be affected by the handling and/or subjective decision of the test personnel (e. g. when manual calipers are used for testing), the test objects must be tested by 3 appraisers in 3 test runs, respectively.

If handling and/or subjective decisions are irrelevant (e. g. in case of automatic test systems), the test objects must be tested in 4 test runs.

In either case, the order of the test objects must be randomly rearranged for each test run. The test results (i.e. the ratings *"+"* or *"–"*) are documented.

Analysis

If all ratings for one reference part match the reference rating, this rating (i.e. *"+"* or *"–"*) is entered into the column *"Code"* of the summary table. Otherwise *"x"* is entered into the column *"Code"* (see following example).

Next, the table is sorted by descending order of the continuous reference values (highest value on top). The sorted table shows two uncertainty ranges around the limiting values. Their width is an indicator for the variation of the test results and thus GRR.

Beginning at the top of the column *"Code"*, the last reference value tagged *"–"* and the first reference value tagged *"+"* is searched for and the width d2 is calculated as difference of both reference values. Similarly d1 is determined. Next, the mean value d = (d1 + d2) / 2 is calculated from both values. Finally, %GRR = d / T · 100% is calculated.

If the test equipment checks only against one of two limiting values, only one uncertainty range of width d can be determined which is used directly to calculate %GRR.

Capability criterion

The test process is considered capable when %GRR ≤ 10% and conditionally capable when 10% < %GRR ≤ 30% (corresponding to procedures 2 and 3).

Otherwise, the test process is incapable. The process has to be improved by taking suitable measures (e. g. instruction of test personnel, correct handling, changes of construction, alternative test equipment). If the result of a repeated test is negative again, then procedures 1 - 3 must be used.

Example

The example on the right shows test results once sorted according to the running number of the test objects and once again sorted according to the decreasing values of the reference values. Test objects whose ratings are not consistently in agreement are highlighted in grey. Sorting according to the reference values clearly identifies the uncertainty ranges.

Upper uncertainty range (d2):

3.642 mm is the smallest value that is still consistently rated "–"; 3.626 mm is the largest value that is still consistently rated "+":

d2 = 3.642 mm – 3.626 mm = 0.016 mm.

Lower uncertainty range (d1):

3.570 mm is the smallest value that is still consistently rated "+"; 3.546 mm is the largest value that is still consistently rated "–":

d1 = 3.570 mm – 3.546 mm = 0.024 mm.

Mean of uncertainty ranges (d):

d = (d2 + d1) / 2
= (0.016 mm + 0.024 mm) / 2
= 0.020 mm

Reproducibility and repeatability (%GRR):

Mean value d related to the tolerance T = 0.075 mm of the characteristic:

%GRR = d / T * 100%
= 0.020 mm / 0.075 mm* 100%
= 26.7%

Sorted by Test Object No.

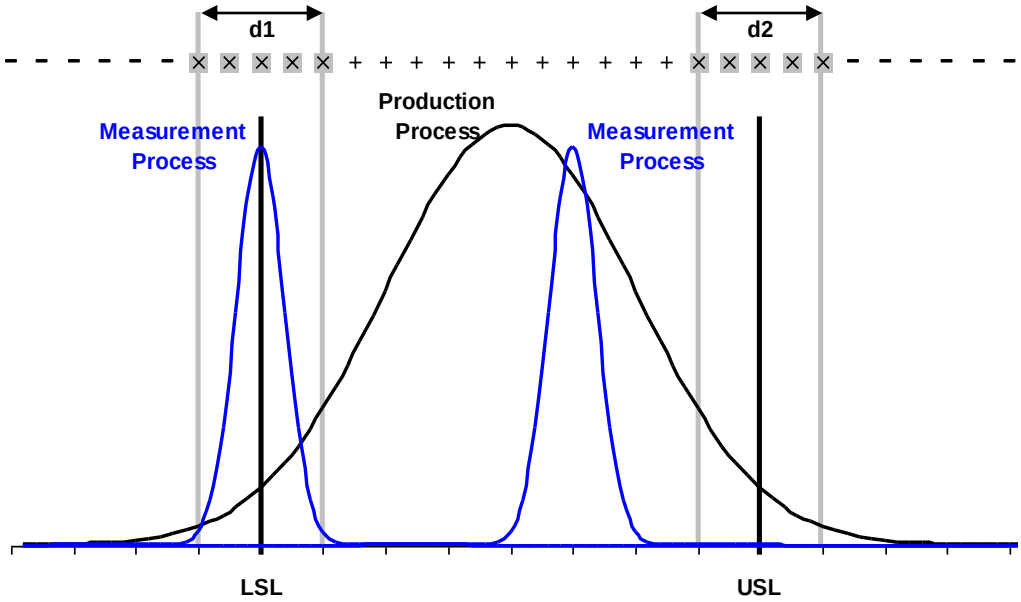
Test object no.	Continuous reference value	Discretized reference value	Appraiser A, Trial 1	Appraiser A, Trial 2	Appraiser A, Trial 3	Appraiser B, Trial 1	Appraiser B, Trial 2	Appraiser B, Trial 3	Code
1	3,632	+	-	-	+	+	+	-	x
2	3,649	-	-	-	-	-	-	-	-
3	3,587	+	+	+	+	+	+	+	+
4	3,552	-	+	-	-	-	-	-	x
5	3,621	+	+	+	+	+	+	+	+
6	3,645	-	-	-	-	-	-	-	-
7	3,652	-	-	-	-	-	-	-	-
8	3,599	+	+	+	+	+	+	+	+
9	3,634	+	-	-	+	-	+	-	x
10	3,625	+	+	+	+	+	+	+	+
11	3,572	+	+	+	+	+	+	+	+
12	3,552	-	+	-	+	-	-	+	x
13	3,595	+	+	+	+	+	+	+	+
14	3,561	-	+	+	+	-	+	+	x
15	3,617	+	+	+	+	+	+	+	+
16	3,585	+	+	+	+	+	+	+	+
17	3,531	-	-	-	-	-	-	-	-
18	3,582	+	+	+	+	+	+	+	+
19	3,544	-	-	-	-	-	-	-	-
20	3,574	+	+	+	+	+	+	+	+
21	3,595	+	+	+	+	+	+	+	+
22	3,642	-	-	-	-	-	-	-	-
23	3,621	+	+	+	+	+	+	+	+
24	3,565	+	+	+	-	-	-	+	x
25	3,593	+	+	+	+	+	+	+	+
26	3,622	+	+	+	+	+	+	+	+
27	3,632	+	-	-	-	+	+	+	x
28	3,664	-	-	-	-	-	-	-	-
29	3,546	-	-	-	-	-	-	-	-
30	3,652	-	-	-	-	-	-	-	-
31	3,586	+	+	+	+	+	+	+	+
32	3,641	-	-	-	-	+	-	+	x
33	3,614	+	+	+	+	+	+	+	+
34	3,600	+	+	+	+	+	+	+	+
35	3,591	+	+	+	+	+	+	+	+
36	3,632	+	-	-	-	+	+	-	x
37	3,570	+	+	+	+	+	+	+	+
38	3,603	+	+	+	+	+	+	+	+
39	3,578	+	+	+	+	+	+	+	+
40	3,597	+	+	+	+	+	+	+	+
41	3,587	+	+	+	+	+	+	+	+
42	3,614	+	+	+	+	+	+	+	+
43	3,613	+	+	+	+	+	+	+	+
44	3,592	+	+	+	+	+	+	+	+
45	3,560	-	+	+	+	-	-	+	x
46	3,626	+	+	+	+	+	+	+	+
47	3,632	+	-	-	-	+	+	-	x
48	3,573	+	+	+	+	+	+	+	+
49	3,559	-	+	+	-	-	-	-	x
50	3,609	+	+	+	+	+	+	+	+

Sorted by Reference Value

Test object no.	Referenzwert kontinuierlich	Referenzwert diskretisiert	Appraiser B, Trial 1	Appraiser B, Trial 1	Appraiser B, Trial 1	Appraiser B, Trial 1	Appraiser B, Trial 2	Appraiser B, Trial 3	Code	
28	3,664	-	-	-	-	-	-	-	-	
7	3,652	-	-	-	-	-	-	-	-	
30	3,652	-	-	-	-	-	-	-	-	
2	3,649	-	-	-	-	-	-	-	-	
6	3,645	-	-	-	-	-	-	-	-	
22	3,642	-	-	-	-	-	-	-	-	
32	3,641	-	-	-	-	+	-	+	x	
9	3,634	+	-	-	+	-	+	-	x	
1	3,632	+	-	-	+	+	+	-	x	d2
27	3,632	+	-	-	-	+	+	+	x	
36	3,632	+	-	-	-	+	+	-	x	
47	3,632	+	-	-	-	+	+	-	x	
46	3,626	+	+	+	+	+	+	+	+	
10	3,625	+	+	+	+	+	+	+	+	
26	3,622	+	+	+	+	+	+	+	+	
5	3,621	+	+	+	+	+	+	+	+	
23	3,621	+	+	+	+	+	+	+	+	
15	3,617	+	+	+	+	+	+	+	+	
33	3,614	+	+	+	+	+	+	+	+	
42	3,614	+	+	+	+	+	+	+	+	
43	3,613	+	+	+	+	+	+	+	+	
50	3,609	+	+	+	+	+	+	+	+	
38	3,603	+	+	+	+	+	+	+	+	
34	3,600	+	+	+	+	+	+	+	+	
8	3,599	+	+	+	+	+	+	+	+	
40	3,597	+	+	+	+	+	+	+	+	
13	3,595	+	+	+	+	+	+	+	+	
21	3,595	+	+	+	+	+	+	+	+	
25	3,593	+	+	+	+	+	+	+	+	
44	3,592	+	+	+	+	+	+	+	+	
35	3,591	+	+	+	+	+	+	+	+	
3	3,587	+	+	+	+	+	+	+	+	
41	3,587	+	+	+	+	+	+	+	+	
31	3,586	+	+	+	+	+	+	+	+	
16	3,585	+	+	+	+	+	+	+	+	
18	3,582	+	+	+	+	+	+	+	+	
39	3,578	+	+	+	+	+	+	+	+	
20	3,574	+	+	+	+	+	+	+	+	
48	3,573	+	+	+	+	+	+	+	+	
11	3,572	+	+	+	+	+	+	+	+	
37	3,570	+	+	+	+	+	+	+	+	
24	3,565	+	+	+	-	-	-	+	x	
14	3,561	-	+	+	+	-	+	+	x	
45	3,560	-	+	+	+	-	-	+	x	d1
49	3,559	-	+	+	-	-	-	-	x	
4	3,552	-	+	-	-	-	-	-	x	
12	3,552	-	+	-	+	-	-	+	x	
29	3,546	-	-	-	-	-	-	-	-	
19	3,544	-	-	-	-	-	-	-	-	
17	3,531	-	-	-	-	-	-	-	-	

Explanation

Variations of the production process, i.e. the variation of the characteristic values, and variations of the measurement process, i.e. the variation of the (continuous) measurement results for two differing characteristic values, are represented by the distributions in the following chart:



It is evident that all measurement results for a characteristic value which is sufficiently far from the limiting values lead to consistent (discrete) ratings (i.e. ”within tolerance” in the above example). In contrast, measurement results that do not lead to consistent ratings have to be expected for a characteristic value which is sufficiently close to a limiting value (i.e. individual measurement results are partly within and partly outside the tolerance range). Thus, the span of characteristic values without consistent ratings (d1, d2) is a suitable estimate for the spread of the (discrete) test process and therefore interpretable as reproducibility and repeatability GRR.

Measurement System Analysis Procedure 6 (Discrete Char.)

BOSCH

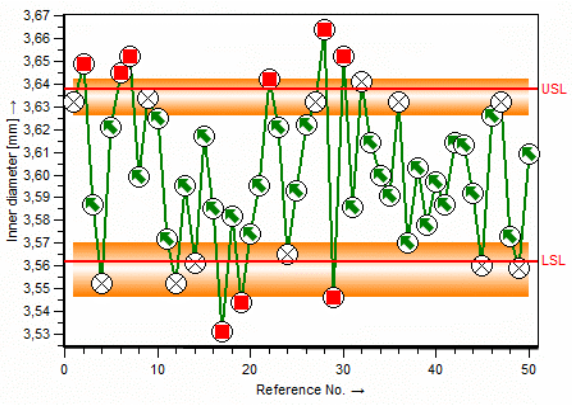
Sheet 1 / 1

Area	:	MSE3	Operation	:	Grinding inner diam.	Characteristic	:	Inner diameter
Group/Dptm.	:	MOE7	Machine	:	BOKO 3	Char. No.	:	1
Workshop/sect.	:	W450	Machine No.	:	1003954	Nominal value	:	3.600
Product	:	Injector	Test station	:	JML0782W001	Lower allowance	:	-0.0375
Part	:	Needle	Gage	:	LG_4H7N1	Upper allowance	:	0.0375
Article number	:	0 433 392 425	Gage No.	:	67027025840013	Tolerance	:	0.075
Change status	:	20.01.2019	Gage Manuf.	:	BaP	Unit	:	mm
			Resolution	:	0.002			

Comment :

Standard: Standard No. Standard/Ref. value:

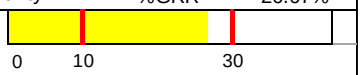
n	Ref. 1	$x_{A;1}$	$x_{A;2}$	$x_{A;3}$	$x_{B;1}$	$x_{B;2}$	$x_{B;3}$	
1	3,6320	–	–	+	+	+	–	☹
2	3,6490	–	–	–	–	–	–	☺
3	3,5870	+	+	+	+	+	+	☺
4	3,5520	+	–	–	–	–	–	☹
5	3,6210	+	+	+	+	+	+	☺
6	3,6450	–	–	–	–	–	–	☺
7	3,6520	–	–	–	–	–	–	☺
8	3,5990	+	+	+	+	+	+	☺
9	3,6340	–	–	+	–	+	–	☹
10	3,6250	+	+	+	+	+	+	☺
11	3,5720	+	+	+	+	+	+	☺
12	3,5520	+	–	+	–	–	+	☹
13	3,5950	+	+	+	+	+	+	☺
14	3,5610	+	+	+	–	+	+	☹
15	3,6170	+	+	+	+	+	+	☺
16	3,5850	+	+	+	+	+	+	☺
17	3,5310	–	–	–	–	–	–	☺
18	3,5820	+	+	+	+	+	+	☺
19	3,5440	–	–	–	–	–	–	☺
20	3,5740	+	+	+	+	+	+	☺
21	3,5950	+	+	+	+	+	+	☺
22	3,6420	–	–	–	–	–	–	☺
23	3,6210	+	+	+	+	+	+	☺
24	3,5650	+	+	–	–	–	+	☹
25	3,5930	+	+	+	+	+	+	☺
26	3,6220	+	+	+	+	+	+	☺
27	3,6320	–	–	–	+	+	+	☹
28	3,6640	–	–	–	–	–	–	☺
29	3,5460	–	–	–	–	–	–	☺
30	3,6520	–	–	–	–	–	–	☺
31	3,5860	+	+	+	+	+	+	☺
32	3,6410	–	–	–	+	–	+	☹
33	3,6140	+	+	+	+	+	+	☺
34	3,6000	+	+	+	+	+	+	☺
35	3,5910	+	+	+	+	+	+	☺
36	3,6320	–	–	–	+	+	–	☹
37	3,5700	+	+	+	+	+	+	☺
38	3,6030	+	+	+	+	+	+	☺
39	3,5780	+	+	+	+	+	+	☺
40	3,5970	+	+	+	+	+	+	☺
41	3,5870	+	+	+	+	+	+	☺
42	3,6140	+	+	+	+	+	+	☺
43	3,6130	+	+	+	+	+	+	☺
44	3,5920	+	+	+	+	+	+	☺
45	3,5600	+	+	+	–	–	+	☹
46	3,6260	+	+	+	+	+	+	☺
47	3,6320	–	–	–	+	+	–	☹
48	3,5730	+	+	+	+	+	+	☺
49	3,5590	+	+	–	–	–	–	☹
50	3,6090	+	+	+	+	+	+	☺



Drawing Values			
	USL	=	3,63750
	LSL	=	3,56250
	T	=	0,07500
Number of reference measurements		=	1
Number of reference parts		=	50
Number of appraisers		=	2
Number of trials per appraiser		=	3

Analysis of Non-Conformancies			
Number of non-conformancies	$n_{<>}$	=	12
	$n_{max<>}$	=	0
	$n'_{max<>}$	=	1

Signal Detection Approach			
Non-conformance range	d_{OSG}	=	0.01600
Non-conformance range	d_{USG}	=	0.02400
Average non-conformance range	d	=	0.02000
Repeatability & Reproducibility	%GRR	=	26.67%



Hypothesis Tests			
$Kappa_{FI}$	K'_{FI}	=	0.6060
$Kappa_{Co}$	K'_{Co}	=	0.6063
$Kappa_{min}$	K'_{min}	=	---

Bosch 2018 — Procedure 6, Discrete Characteristics

2020-04-06 - SOCOS

5.2 Procedure 7: Test decisions for discrete and discretized continuous characteristics

Objective

Verification of the capability of a test process regarding unambiguous test decisions when testing discrete or discretized continuous characteristics.

NOTE: This procedure can be used with and without continuous reference values.

Description of the procedure

The study is done using a reference lot which comprises reference parts whose discrete characteristic values are determined and documented before starting the study.

- **Reference parts with continuous characteristic values**

  If continuous characteristic values of the reference parts can be determined, they have to be determined by measurement. The measurement uncertainty U, allocated to the measured values, must be known. Reference parts are required whose characteristic values cover a range beginning slightly below LSL − U and ending slightly above USL + U. The measurement result is documented for each reference part.

  Next, each reference part is allocated unambiguously to a countable rating category which corresponds to the measurement result (discretizing): e. g. *"within tolerance" / "out of tolerance"* or *"good" / "bad"* or corresponding numeric codes such as *"1" / "0"*. The discretized results (i.e. the reference ratings) are documented.

- **Reference parts with discrete characteristic values**

  Reference standard (boundary samples catalogue): For a repeatable and unambiguous identification of certain characteristics (attributes) of test objects, a reference standard (boundary samples catalogue) is necessary, against which the test objects are compared. This is a documentation of all attributes of test objects that are to be identified by the test process. The catalogue can be implemented as a collection of physically existing parts with corresponding attributes or as a collection of corresponding photographic images in case of visual inspections, as a collection of sound samples in case of acoustic inspections, etc.

  Categorizing: If continuous characteristic values of the reference parts can not be determined (e. g. in case of visual inspections), each reference part is allocated to a countable rating category according to its attributes using the reference standard for comparison (boundary samples catalogue). The results are documented: e. g. *"good" / "bad"* or corresponding numeric codes such as *"1" / "0"*.

  Number of categories: More than two categories are possible, e. g. *"good" / "rework" / "bad"* or corresponding numeric codes such as *"2" / "1" / "0"*. However, experience shows that multi-stage tests (see Appendix G.4) generally lead to more reliable results than multiple categories.

- **Reference lot (master)**

  Lot size: The lot size should be as large as possible (100 to 200 reference parts are recommended; at least 50 reference parts are required according to [AIAG MSA]). The lot size should follow the optimum between partly conflicting general conditions such as requirements for the statistical power of the test, acceptable effort, available capacities and economy.

  Composition: All attributes relevant for the test must be contained in the reference lot, i.e. all attributes that are to be identified by the test process. The reference lot should be composed according to the current frequencies of the individual attributes in the production lot, e. g. according to a Pareto analysis over the last production interval (last 3 months recommended).

  Identifiability: Each reference part of the lot must be unambiguously identifiable so that the respective data can always be correctly allocated. This requirement must be implemented so that only authorized personnel but not the test personnel can identify the reference part. Possible implementations are e. g. 2D barcodes, complex number codes, labels only legible under UV light.

- **Conducting the study**

  To do the study, the reference parts are used as test objects. They are tested under serial conditions in a random order that is unknown to the test personnel using the specified test equipment and test methods (e. g. according to the test plan) or an automatic test system. Each part is allocated to a rating category. The test personnel must be adequately trained and instructed.

  If the test results (i.e. the ratings) can be affected by the handling and/or subjective decision of the test personnel (e. g. when manual calipers are used for testing or in case of visual inspections), the test objects must be tested by (at least) 3 appraisers in (at least) 3 test runs, respectively.

  If handling and/or subjective decisions are irrelevant (e. g. in case of automatic test systems), the test objects must be tested in multiple test runs (6 test runs are recommended).

  In either case, the order of the test objects must be randomly rearranged for each test run. The test results (i.e. the ratings) are documented.

Analysis

The unambiguousness of test decisions is analyzed by means of pair-wise agreements of individual ratings. The parameter κ ("kappa") is used as a quantitative measure:

$$\kappa = \frac{\text{Observed non-random agreements}}{\text{Possible non-random agreements}}$$

Details of the calculation are explained in Appendix G.

The analysis comprises the following comparisons and calculations of the corresponding parameters κ:

- Within appraisers: Compare all test runs of each individual appraiser without comparing to the reference (*repeatability*).
- Between appraisers: Compare all test runs of all appraisers without comparing to the reference (*reproducibility*).
- Compare all test runs of each individual appraiser to the reference.
- Compare all test runs of all appraisers to the reference.

Deviating from [AIAG MSA] the analysis is performed using Fleiss' kappa statistics [Fleiss] which is more generally applicable. If the analysis according to [AIAG MSA] using Cohen's kappa statistics is explicitly demanded (e. g. due to customer requirements), then proceed according to [AIAG MSA].

NOTE 1: Complementing documentation on the topics "cross-table method" and "analysis according to [AIAG MSA]" is available at C/QMM and on the C/QMM intranet pages.

NOTE 2: The analysis according to [AIAG MSA] does not intend "within-appraiser" comparisons and comparing "all test runs of all appraisers to the reference" is not possible.

Capability criterion

The capability is classified by means of the parameter κ ("kappa"):

- $\kappa \geq 0.9$ test process is capable,
- $0.9 > \kappa \geq 0.7$ test process is conditionally capable,
- $\kappa < 0.7$ test process is not capable.

The minimum of all determined κ-values is relevant for the final classification of the test process.

If the test process is conditionally capable or not capable, it must be improved by taking suitable measures (e. g. instruction of test personnel, correct handling, changes of construction, alternative test equipment).

BOSCH
Quality Management

Test Process Analysis Procedure 7

Record No.: 9911015

Sheet 1 of 2

Product / Test Object

Product: Housing

Part: Cover

Part / Drawing No.: A 111 999 222

Revision: 05 / 02/29/2009

Characteristic

Designation: Surface quality

Characteristic No.: 15

☐ **Continuous Characteristic**

Nominal Value: n/a

Upper Limit: n/a

Lower Limit: n/a

Tolerance: n/a

Unit: n/a

☑ **Discrete Characteristic**

Measuring & Test Equipment

Location: XxP/W000999

Test / Measuring Station: Visual inspection

Test / Measuring Equipment No.: 123 456 789

Designation: Boundary samples catalogue

Calibration Certificate No.: n/a

Measurement Uncertainty: n/a

Test Method: Visual inspection, manually, room temperature 20.2ºC, light intensity 250 cd (Candela)

Test Scenario

Number of reference parts	N_O	=	50
Number of appraisers	N_A	=	3
Number of trials per appraiser	N_T	=	3
Number of rating categories	N_C	=	2

Rating Categories

0 - Not OK

1 - OK

Test Data: See sheet 2 ff

Analysis

Appraiser name	Symbol	Within appraiser without reference				Each appraiser against reference			
		κ (Kappa)	κ ≥ 0.90 capable	0.70 ≤ κ < 0.90 conditionally capable	κ < 0.70 not capable	κ (Kappa)	κ ≥ 0.90 capable	0.70 ≤ κ < 0.90 conditionally capable	κ < 0.70 not capable
Miller	A	0.7600		X		0.8802		X	
Smith	B	0.8451		X		0.9226	X		
King	C	0.7029		X		0.7747		X	
		Between appraisers without reference				All appraisers against reference			
all		0.7936		X		0.8592		X	

Overall Result

Minimum of all results: Kappa = 0.7029

Kappa ≥ 0.90: ☐ capable

0.70 ≤ Kappa < 0.90: ☑ conditionally capable

Kappa < 0.70: ☐ not capable

Comment: none

Date: 02/29/2009 Department: W025 Name: J. Q. Public Signature: J. Q. Public

Procedure 7: Test Results (Ratings) **Record No. 9911015, Sheet 2 of 2**

Test Object No.	Reference Value (continuous)	Reference Value (discrete or discretized)	Appraiser (Symbol) – Trial No.											
			Miller			*Smith*			*King*					
			A-1	A-2	A-3	B-1	B-2	B-3	C-1	C-2	C-3			
1	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
2	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
3	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
4	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
5	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
6	*n/a*	*1*	*1*	*1*	*0*	*1*	*1*	*0*	*1*	*0*	*0*			
7	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*0*	*1*			
8	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
9	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
10	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
11	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
12	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*1*	*0*			
13	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
14	*n/a*	*1*	*1*	*1*	*0*	*1*	*1*	*1*	*1*	*0*	*0*			
15	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
16	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
17	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
18	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
19	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
20	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
21	*n/a*	*1*	*1*	*1*	*0*	*1*	*0*	*1*	*0*	*1*	*0*			
22	*n/a*	*0*	*0*	*0*	*1*	*0*	*1*	*0*	*1*	*1*	*0*			
23	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
24	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
25	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
26	*n/a*	*0*	*0*	*1*	*0*	*0*	*0*	*0*	*0*	*0*	*1*			
27	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
28	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
29	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
30	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*1*	*0*	*0*	*0*			
31	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
32	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
33	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
34	*n/a*	*0*	*0*	*0*	*1*	*0*	*0*	*1*	*0*	*1*	*1*			
35	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
36	*n/a*	*1*	*1*	*1*	*0*	*1*	*1*	*1*	*1*	*0*	*1*			
37	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
38	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
39	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
40	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
41	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
42	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
43	*n/a*	*1*	*1*	*0*	*1*	*1*	*1*	*1*	*1*	*1*	*0*			
44	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
45	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
46	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
47	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
48	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			
49	*n/a*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*			
50	*n/a*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*			

Rating categories: 0 - not OK; 1 - OK n/a - not applicable

5.3 Notes on stability monitoring and repetition of capability studies

[AIAG MSA] neither contains a recommendation nor a method for monitoring the long-term stability of test processes for discrete characteristics. It may seem logical to proceed similar to procedure 5 by using stability charts for these test processes that are similar to np-charts or p-charts used for SPC processes (see [AIAG SPC]). However, sample sizes of n > 50 are usually required for these charts. Thus, with regard to the required test effort, there is no advantage over repeating the complete capability study.

[AIAG MSA] also does not include any notes or recommendations on time intervals that should be applied for repeating the capability study. A common practice is to monitor the error rate of the production process and to verify the capability of the test process again in case of significant changes. However, it must be clarified here that changes of the error rate can be caused by changes in the test process as well as changes in the production process. Thus, they are not a clear indication.
If there is a possibility to monitor discrete characteristics as part of successive process steps (indirectly and preferably using a continuous characteristic), this possibility should be used in addition or as an alternative.

Moreover, the following criteria are typical examples that may require a new capability study:

- after adjusting the test system or components of the test system (e. g. camera for visual inspection during control of inspection, measuring and test equipment);
- upon restart after maintenance work where substantial disassemblies, modifications or replacements of crucial parts where necessary (e. g. camera for visual inspection);
- upon start-up of new, overhauled or repaired test systems;
- upon (subsequent) tolerance cutbacks in case of discretized continuous characteristics;
- in case of technical changes of the test system (e. g. setup, software);
- in case of completions or significant changes of the reference standard (boundary samples catalogue);
- if basic conditions of the test process are changed that may influence the capability of the test process (e. g. workflow, testing strategy);
- after changes of the operating personnel (e. g. new staff members);
- if it is suspected that the test system does not work properly;
- if necessary, before and definitely after relocation of the test system.

If in doubt, the test process analysis has to be repeated and the capability must be verified again.

6 Assessment of Non-Capable Measurement and Test Processes

The following approach is reasonable for measurement and test processes whose unconditional capability cannot be verified:

- Root cause analysis (e. g. cause-and-effect diagram, 5 x Why);
- Review of limiting factors and coordination between production engineering and development departments (e. g. with regard to tolerances, production concept, measurement strategy);
- Use of FMEA results that are available for the respective characteristic;
- Documentation of measures (e. g. in the control plan).

It is necessary to document comprehensibly that the compliance with the demanded specifications is assured and, if necessary, agreed with the customer.

The form in Appendix B (page 40ff) can be supportive for the decision whether a conditional approval can be accounted for.

Appendix

A Examples of Checklists for Measurement Process Analyses

Measurement Process Analysis

Appendix 1: Checklist for planning and optimization of a measurement process

	Checked	OK	Measures
Measuring equipment, setting gauges			
Measuring, clamping and retaining forces	☐	☐	
Definition of measuring and test points	☐	☐	
Holding fixtures, alignment of measuring object and measuring sensor	☐	☐	
Sampling elements	☐	☐	
Guidance, friction, wear	☐	☐	
Positioning, tilt of measuring object	☐	☐	
Test sequence, warming-up phase	☐	☐	
Quality of setting gauge(s) and standard(s)	☐	☐	
Measuring method, strategy			
Sampling or non-contact	☐	☐	
Reference element, basis of decision-making	☐	☐	
Measuring speed, settling time	☐	☐	
Multi-point measurement or scanning instead of individual measurement values	☐	☐	
Mean value of repeated measurements	☐	☐	
Measuring ranges	☐	☐	
Measurement software, statistics software	☐	☐	
Calibration of measuring chain	☐	☐	
Setup procedure (e.g. before each measurement)	☐	☐	
Ambient conditions			
Vibrations, oscillations	☐	☐	
Dust, oil mist, draught, humidity	☐	☐	
Temperature fluctuations, solar radiation	☐	☐	
Electrical interference, voltage peaks	☐	☐	
Energy fluctuations (air flow, electrical power)	☐	☐	
Measuring object			
Cleanliness, washing residues	☐	☐	
Surface quality, burrs	☐	☐	
Imperfect shape, reference basis	☐	☐	
Material properties (e.g. temperature coefficient)	☐	☐	
Operator, working instructions			
Briefing, training, accurateness, handling	☐	☐	
Cleanliness (e.g. greasy hands), heat transfer	☐	☐	
...			

Measurement Process Analysis

Appendix 2: Checklist for selection of measuring equipment

	Checked	OK	Measures
Resolution < 5%	☐	☐	
Linearly measuring test equipment used?	☐	☐	
Absolutely measuring test equipment used?	☐	☐	
More robust measuring equipment can be used (e.g. support, guidance, operating levers, transmitters, fixations)?	☐	☐	
Operator-independent measuring equipment can be used?	☐	☐	
New (non-contact) measurement methods can be used?	☐	☐	
Do the measuring systems have interfaces for automatic data transfer (AQDEF format)?	☐	☐	
More suitable measuring equipment is available or can be acquired?	☐	☐	

...

Measurement Process Analysis

Appendix 3: Checklist for review of characteristic and tolerance

	Checked	OK	Measures
Influence of characteristic on function of the production part (e.g. DRBFM and/or FMEA considered? Design of characteristic ensures function?)	☐	☐	
Alternative characteristic as a "substitute" (e.g. tightness instead of roundness)	☐	☐	
Effects of alternative characteristic on process capability and process control (function, reliability)	☐	☐	
Tolerance adaptation (e.g. using statistical tolerancing)	☐	☐	
Alternative materials and/or substances can be used?	☐	☐	
Alternative production method or parameters can be used? (e.g. DoE and/or suitable test runs were conducted?)	☐	☐	
Consultation with			
Production planning / manufacturing engineering	☐	☐	
Production	☐	☐	
Quality management	☐	☐	
Development	☐	☐	
Sales and/or customers	☐	☐	
Purchasing	☐	☐	

...

B Forms for Manual Analysis

Note: The forms for procedures 1 – 3 and 5 – 6 correspond to the forms contained in the previous issue of booklet 10. Their use is not recommended. Software-supported analysis should be preferred.

BOSCH
Quality Management

Measurement Process Analysis Procedure 1

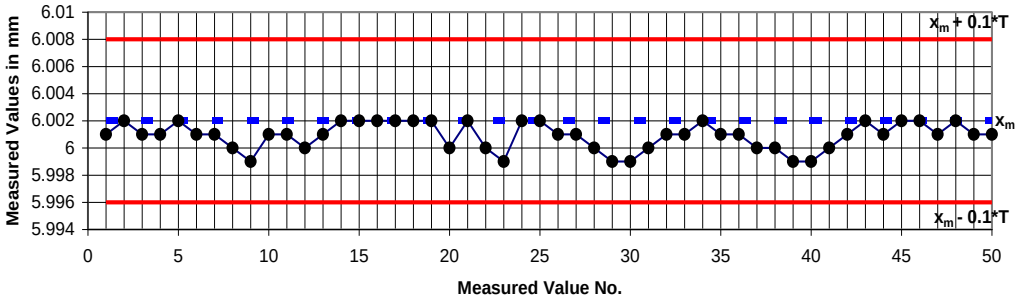
Record No.: 9911015
Sheet 1 of 1

Measuring Equipment		Characteristic		Measurement Standard	
Location:	W025	Measuring Object:	Shaft	Designation:	Positioning cylinder
Designation	Length meter	Drawing No.:	1460320000	Equipment No.:	LY8N 6.000 No. 1
Equipment No.:	JML9Q002	Designation of Characteristic:	Outer diameter	Reference Value x_m:	6.0020 mm
Resolution:	0.001 mm	Nominal Value:	6.000 mm	Uncertainty U_{cal}:	0.0005 mm
		Tolerance:	0.060 mm		

Measurement: *Manual operation; measurement point: middle of cylinder; room temperature: 20.2°C*

Table values in: *mm* **Deviations from:** *-*

1 - 5	6 - 10	11 - 15	16 - 20	21 - 25	26 - 30	31 - 35	36 - 40	41 - 45	46 - 50
6.001	6.001	6.001	6.002	6.002	6.001	6.000	6.001	6.000	6.002
6.002	6.001	6.000	6.002	6.000	6.001	6.001	6.000	6.001	6.001
6.001	6.000	6.001	6.002	5.999	6.000	6.001	6.000	6.002	6.002
6.001	5.999	6.002	6.002	6.002	5.999	6.002	5.999	6.001	6.001
6.002	6.001	6.002	6.000	6.002	5.999	6.001	5.999	6.002	6.001



Reference value x_m= 6.0020 mm **Mean value** $\bar{x}$ = 6.0009 mm **Standard deviation s =** 0.0010 mm

Resolution ≤ 5% T? [X] yes [ ] no

Systematic measurement error:
For n = 25 the systematic measurement error is significant if $|\bar{x} - x_m| > 0.413 \cdot s$
For n = 50 the systematic measurement error is significant if $|\bar{x} - x_m| > 0.284 \cdot s$
$|\bar{x} - x_m|$ = 0.0011 mm 0.413*s = 0.00041 mm 0.284*s = 0.00028 mm
[ ] insignificant [X] significant [ ] not assessed

Capability indexes:

$$C_g = \frac{0.2 \cdot T}{6 \cdot s} = \frac{0.2 \cdot 0.06\ mm}{6 \cdot 0.001\ mm} = 2.01$$

$$C_{gk} = \frac{0.1 \cdot T - |\bar{x} - x_m|}{3 \cdot s} = \frac{(0.1 \cdot 0.06 - |6.0009 - 6.002|)\ mm}{3 \cdot 0.001\ mm} = 1.64$$

Cg ≥ 1.33 and Cgk ≥ 1.33? [X] yes [ ] no

Comment: *The systematic measurement error cannot be corrected*

Date: *02/29/2009* Department: *W025* Name: *John Q. Public* Signature: *John Q. Public*

2020-04-06 - SOCOS

BOSCH
Quality Management

Measurement Process Analysis Procedure 2

Record No.: 9911015
Sheet 2 of 2

Measuring Equipment

Location:	W025
Designation	Length meter
Equipment No.:	JML9Q002
Resolution:	0.001 mm

Characteristic

Measuring Object:	Shaft
Drawing No.:	1460320000
Designation of Characteristic:	Outer diameter
Nominal Value:	6.000 mm
Tolerance:	0.060 mm

Result procedure 1: C_g = 2.01 C_{gk} = 1.64 Record / Sheet / Date: 9911015 / 1 / 02/29/2009

Measurement: Manual operation; measurement point: middle of cylinder; room temperature: 20.2°C

Values: Table values in: mm Deviation from: —

	Appraiser A: Mr. Miller			Appraiser B: Ms. Smith			Appraiser C: Mr. King			
Meas. Objekt No.	Meas. Series 1	Meas. Series 2	Range $R_{A,i}$	Meas. Series 1	Meas. Series 2	Range $R_{B,i}$	Meas. Series 1	Meas. Series 2	Range $R_{C,i}$	Mean Value $\bar{x}_i$
1	6.029	6.030	0.001	6.033	6.032	0.001	6.031	6.030	0.001	6.031
2	6.019	6.020	0.001	6.020	6.019	0.001	6.020	6.020	0.000	6.020
3	6.004	6.003	0.001	6.007	6.007	0.000	6.010	6.006	0.004	6.006
4	5.982	5.982	0.000	5.985	5.986	0.001	5.984	5.984	0.000	5.984
5	6.009	6.009	0.000	6.014	6.014	0.000	6.015	6.014	0.001	6.013
6	5.971	5.972	0.001	5.973	5.972	0.001	5.975	5.974	0.001	5.973
7	5.995	5.997	0.002	5.997	5.996	0.001	5.995	5.994	0.001	5.996
8	6.014	6.018	0.004	6.019	6.015	0.004	6.016	6.015	0.001	6.016
9	5.985	5.987	0.002	5.987	5.986	0.001	5.987	5.986	0.001	5.986
10	6.024	6.028	0.004	6.029	6.025	0.004	6.026	6.025	0.001	6.026
	$\bar{x}_A$ = 6.0039		$\bar{R}_A$ = 0.0016	$\bar{x}_B$ = 6.0058		$\bar{R}_B$ = 0.0014	$\bar{x}_C$ = 6.0054		$\bar{R}_C$ = 0.0011	R_p = 0.0580

Analysis

Note: K-Factors only valid for n = 10 measuring objects, k = 3 appraisers and r = 2 measurement series per appraiser

Mean range		$\bar{\bar{R}} = \frac{1}{3} \cdot (\bar{R}_A + \bar{R}_B + \bar{R}_C) =$	0.00137 mm
Equipment Variation EV (Repeatability)	$K_1 = 0.8862$	$EV = K_1 \cdot \bar{\bar{R}} =$	0.00121 mm
Range of appraiser mean values		$R_{\bar{x}} =$	0.00190 mm
Appraiser Variation AV (Reproducibility)	$K_2 = 0.5231$	$AV = \sqrt{(K_2 \cdot R_{\bar{x}})^2 - \frac{EV^2}{n \cdot r}} =$	0.00096 mm
Gage Repeatability and Reproducibility GRR		$GRR = \sqrt{EV^2 + AV^2} =$	0.00154 mm
		$\%GRR = 6 \cdot \frac{GRR}{T} \cdot 100\% =$	***15.4%***
Part Variation PV	$K_3 = 0.3146$	$PV = K_3 \cdot R_p =$	0.01825 mm
Number of distinct catagories ndc	ndc ≥ 5 ?	$ndc = \sqrt{2} \cdot \frac{PV}{GRR} = 1.41 \cdot \frac{PV}{GRR} =$	**16**

%GRR ≤ 10% ☐ ***capable*** 10% < %GRR ≤ 30% ☒ ***conditionally capable*** %GRR > 30% ☐ ***not capable***

Comment: Improvement of %GRR currently unfeasible; risk analysis required

Date: 02/29/2009 Department: W025 Name: J. Q. Public Signature: J. Q. Public

Note: Calculations on this form are done using the average range method (ARM, see Appendix D.3) which is not recommended and should only be used in exceptions.

BOSCH
Quality Management

Measurement Process Analysis Procedure 3

Record No.: 9911015
Sheet 2 of 2

Measuring Equipment

Location:	W025
Designation	Length meter
Equipment No.:	JML9Q002
Resolution:	0.001 mm

Characteristic

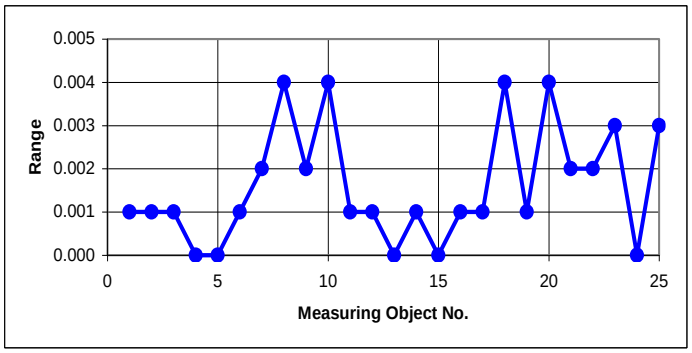
Measuring Object:	Shaft
Drawing No.:	1460320000
Designation of Characteristic:	Outer diameter
Nominal Value:	6.000 mm
Tolerance:	0.060 mm

Result procedure 1: C_g = 2.01 C_{gk} = 1.64 Record / Sheet / Date: 9911015 / 1 / 02/29/2009

Measurement: Automatic operation; measurement point: middle of cylinder; room temperature: 20.2°C

Values: Table values in: mm Deviation from: —

Meas. Object No.	Meas. Series 1	Meas. Series 2	Range R_i	Mean Value $\bar{x}_i$
1	6.029	6.030	0.001	6.030
2	6.019	6.020	0.001	6.020
3	6.004	6.003	0.001	6.004
4	5.982	5.982	0.000	5.982
5	6.009	6.009	0.000	6.009
6	5.971	5.972	0.001	5.972
7	5.995	5.997	0.002	5.996
8	6.014	6.018	0.004	6.016
9	5.985	5.987	0.002	5.986
10	6.024	6.028	0.004	6.026
11	6.033	6.032	0.001	6.033
12	6.020	6.019	0.001	6.020
13	6.007	6.007	0.000	6.007
14	5.985	5.986	0.001	5.986
15	6.014	6.014	0.000	6.014
16	5.973	5.972	0.001	5.973
17	5.997	5.996	0.001	5.997
18	6.019	6.015	0.004	6.017
19	5.987	5.986	0.001	5.987
20	6.029	6.025	0.004	6.027
21	6.017	6.019	0.002	6.018
22	6.003	6.001	0.002	6.002
23	6.009	6.012	0.003	6.011
24	5.987	5.987	0.000	5.987
25	6.006	6.003	0.003	6.005
			$\bar{R}$ = 0.0016	R_p = 0.0610



Analysis

Note: K-Factors only valid for n = 25 measuring objects and r = 2 measurement series

Mean Range $\bar{R}$ = 0.0016 mm

Equipment Variation EV (Repeatability) $K_1 = 0.8862$ $EV = K_1 \cdot \bar{R}$ = 0.0014 mm

Gage Repeatability & Reproducibility (0) GRR $GRR = EV$ = 0.0014 mm

$$\%GRR = 6 \cdot \frac{GRR}{T} \cdot 100\% = \textbf{14.2\%}$$

Range of Mean Values R_p = 0.0610 mm

Part Variation PV $K_3 = 0.25$ $PV = K_3 \cdot R_p$ = 0.0153 mm

Number of distinct catagories ndc $ndc = \sqrt{2} \cdot \frac{PV}{GRR}$ = **15**

ndc ≥ 5 ?

%GRR ≤ 10% ☐ *capable* 10% < %GRR ≤ 30% [X] *conditionally capable* %GRR > 30% ☐ *not capable*

Comment: Improvement of %GRR currently unfeasible; risk analysis required

Date: 02/29/2009 Department: W025 Name: J. Q. Public Signature: J. Q. Public

Note: Calculations on this form are done using the average range method (ARM, see Appendix D.3) which is not recommended and should only be used in exceptions.

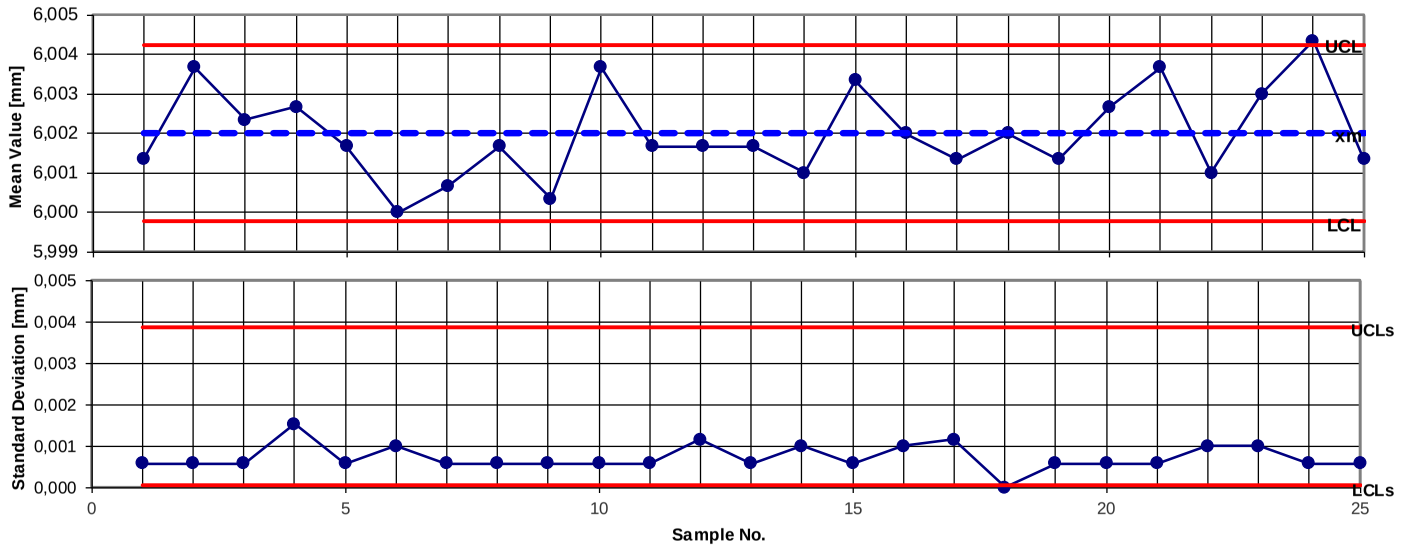
BOSCH
Quality Management

Stability
(Measurement Process)
Procedure 5

Production part	Shaft 1460 320 000
Work step	Stability monitoring
Characteristic	Outer diameter 6 mm

Plant / workshop / department	W025
Stability control chart prepared by	Uk
Date	10.10.2019
Sheet no.	4
Designation of measuring equipment	Length meter
Measuring equipment no.	JML9Q002
ID of setup measurement standard	LY8N 6.000 No. 1
Reference value x_m	6,002 mm
Measurement uncertainty U_{cal}	0,0005 mm
Table values: Deviation from	6,000 mm
Table values: Unit	0,001 mm

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
x_1	2	4	3	3	2	0	1	1	0	4	2	3	2	2	4	3	2	2	1	3	4	2	4	5	2
x_2	1	4	2	1	1	1	1	2	0	4	1	1	1	0	3	2	2	2	2	3	3	0	3	4	1
x_3	1	3	2	4	2	-1	0	2	1	3	2	1	2	1	3	1	0	2	1	2	4	1	2	4	1
x_4																									
x_5																									
$\bar{x}$	1,3	3,7	2,3	2,7	1,7	0,0	0,7	1,7	0,3	3,7	1,7	1,7	1,7	1,0	3,3	2,0	1,3	2,0	1,3	2,7	3,7	1,0	3,0	4,3	1,3
s	0,6	0,6	0,6	1,5	0,6	1,0	0,6	0,6	0,6	0,6	0,6	1,2	0,6	1,0	0,6	1,0	1,2	0,0	0,6	0,6	0,6	1,0	1,0	0,6	0,6



Sample size n	3
Sampling interval	2 times per shift
C_{gk} from procedure 1	1,64

Characteristic

Designation	Outer diameter
Nominal value	6,000 mm
USL	6,030 mm
LSL	5,970 mm
Tolerance T	0,060 mm

Control limits (confidence level 99 %)

$UCL = x_m + (2.58/\sqrt{3}) \cdot (T/\sqrt{40})$	6,0042 mm
$LCL = x_m + (2.58/\sqrt{3}) \cdot (T/\sqrt{40})$	5,9998 mm
$UCL_s = 2.571 \cdot T/40$	0,003856 mm
$LCL_s = 0.037 \cdot T/40$	0,000055 mm

Date	11.10	10.11	10.11	10.11	10.12	10.12	10.12	10.12	10.13	10.13	10.13	10.13	10.14	10.14	10.14	10.14	10.15	10.15	10.15	10.15	10.18	10.18	10.18	10.18	19.10
Time	6:50 AM	11:40 AM	2:40 PM	9:35 PM	6:25 AM	10:50 AM	2:20 PM	9:05 PM	6:20 AM	10:15 AM	3:00 PM	9:35 PM	7:10 AM	11:50 AM	4:05 PM	9:30 PM	6:50 AM	12:10 PM	4:15 PM	9:20 PM	6:45 AM	12:05 PM	4:20 PM	9:15 PM	6:50 AM
Signature	Pf	Mu	Mu	Pf	Pf	Kk	Kk	Pf	Pf	Mu	Mu	Kk	Kk	Kk	Kk	Kk	Pf	Pf	Pf	Kk	Kk	Mu	Mu	Mu	Pf

Evaluated	10/20/2019	Ks
Approved	10/21/2019	Nm

BOSCH
Quality Management

Test Process Analysis Procedure 6

Record No.: 9911015
Sheet 6 of 7

Measuring & Test Equipment

Location: W025
Designation: Limit plug gauge
Equipment No.: LG3.6H11 No. 1

Characteristic

Measuring Object: Housing
Drawing No.: 1265120000
Designation of Characteristic: Inner diameter
Nominal Value: 3.600 mm
Tolerance: 0.075 mm

Measuring System

Designation: Bore hole measuring device
Equipment No.: JMK3N1/3.6 No.1
Traceability: Setting ring gage 3.600
Measurement Uncertainty: 0.002 mm

Test: Manual operation; 2 operators; room temperature 20.2°C

Evaluation: Within tolerance: + Ouside tolerance: - No agreement: x

Test Object No.	Reference		Oper. A		Oper. B		Code
	con-tinuous	dis-cretized	Smith		Miller		
			1	2	1	2	
28	3.664	-	-	-	-	-	-
7	3.652	-	-	-	-	-	-
30	3.652	-	-	-	-	-	-
2	3.649	-	-	-	-	-	-
6	3.645	-	-	-	-	-	-
22	3.642	-	-	-	-	-	-
32	3.641	-	-	-	+	-	x
9	3.634	+	-	-	-	+	x
27	3.632	+	-	-	+	+	x
1	3.632	+	-	-	+	+	x
36	3.632	+	-	-	+	+	x
47	3.632	+	-	-	+	+	x
46	3.626	+	+	+	+	+	+
10	3.625	+	+	+	+	+	+
26	3.622	+	+	+	+	+	+
5	3.621	+	+	+	+	+	+
23	3.621	+	+	+	+	+	+
15	3.617	+	+	+	+	+	+
33	3.614	+	+	+	+	+	+
42	3.614	+	+	+	+	+	+
43	3.613	+	+	+	+	+	+
50	3.609	+	+	+	+	+	+
38	3.603	+	+	+	+	+	+
34	3.600	+	+	+	+	+	+
8	3.599	+	+	+	+	+	+

Test Object No.	Reference		Oper. A		Oper. B		Code
	con-tinuous	dis-cretized	Smith		Miller		
			1	2	1	2	
40	3.597	+	+	+	+	+	+
21	3.595	+	+	+	+	+	+
13	3.595	+	+	+	+	+	+
25	3.593	+	+	+	+	+	+
44	3.592	+	+	+	+	+	+
35	3.591	+	+	+	+	+	+
3	3.587	+	+	+	+	+	+
41	3.587	+	+	+	+	+	+
31	3.586	+	+	+	+	+	+
16	3.585	+	+	+	+	+	+
18	3.582	+	+	+	+	+	+
39	3.578	+	+	+	+	+	+
20	3.574	+	+	+	+	+	+
48	3.573	+	+	+	+	+	+
11	3.572	+	+	+	+	+	+
37	3.570	+	+	+	+	+	+
24	3.565	+	+	+	-	-	x
14	3.561	-	+	+	-	+	x
45	3.560	-	+	+	-	-	x
49	3.559	-	+	+	-	-	x
12	3.552	-	+	-	-	-	x
4	3.552	-	+	-	-	-	x
29	3.546	-	-	-	-	-	-
19	3.544	-	-	-	-	-	-
17	3.531	-	-	-	-	-	-

d1 = 3.57 - 3.546 = 0.024 d2 = 3.642 - 3.626 = 0.016 d = (0.024 + 0.016) / 2 = 0.02 %GRR = **26.7%**

%GRR ≤ 10% [ ] capable
10% < %GRR ≤ 30% [X] conditionally capable
30% < %GRR [ ] not capable

Comment: Test process to be analyzed

Date: 02/29/2009 Department: W025 Name: J. Q. Public Signature: J. Q. Public

Note: On the form, the test results are shown already sorted according to their (continuous) reference values.

BOSCH
Quality Management

Assessment of Non-capable Measurement & Test Processes

Record No.: *9911015*
Sheet *1* of *2*

Measuring Equipment		Characteristic		Measurement Standard	
Location:	*W025*	Measuring / Test Object:	*Shaft*	Designation:	*Positioning cylinder*
Designation:	*Measurement sensor*	Drawing No.:	*1460320000*	Equipment No.:	*LY8N 6.000 No. 1*
Equipment No.:	*LX 0815 P1*	Designation of Characteristic:	*Outer diameter*	Reference Value x_m:	*6.0020 mm*
Resolution:	*0.001 mm*	Nominal Value:	*6.000 mm*	Uncertainty U_{kal}:	*0.0005 mm*
		Tolerance:	*0.060 mm*		

1. Capability indices of procedures 1 – 4 and 6 – 7

Check the applicable result for each procedure (n/a – procedure not applicable / not used)

Pro-cedure	n/a	capable	conditionally capable				not capable	
1			$1.20 \leq C_{gk} < 1.33$	X	$0.80 \leq C_{gk} < 1.20$		$C_{gk} < 0.80$	
2			$10\% < \%GRR \leq 20\%$		$20\% < \%GRR \leq 30\%$	X	$\%GRR > 30\%$	
3	X		$10\% < \%GRR \leq 20\%$		$20\% < \%GRR \leq 30\%$		$\%GRR > 30\%$	
4		X	$1.20 \leq MIN(C_{gk}) < 1.33$		$0.80 \leq MIN(C_{gk}) < 1.20$		$MIN(C_{gk}) < 0.80$	
4 (MSA)	X						*Zero line (partly) outside the confidence limits*	
5	X							
6	X		$10\% < \%GRR \leq 20\%$		$20\% < \%GRR \leq 30\%$		$\%GRR > 30\%$	
7	X		$0.8 \leq \kappa < 0.9$		$0.7 \leq \kappa < 0.8$		$\kappa < 0.7$	
Key figure 1			***1***		***2***	X	***7***	

Highest key figure achieved is relevant

2. External relevance of failures (effect for customer)

Assessment according to design and process FMEA (see booklet 14, CDQ0305, divisional & in-plant regulations)

Effect for customer	*none*	*insignificant to marginal*	*moderately serious to serious*	*very serious*
B-assessment according to FMEA	*1*	*2 - 4*	*5 - 8*	*9 - 10*
Key figure 2	***1***	***2***	***3*** X	***7***

Description of failure effect: *Outer diameter is too large, component cannot be used in the vehicle; 0 km complaint*

3. Internal relevance of failures (effect for Bosch)

Assessment according to design and process FMEA (see booklet 14, CDQ0305, divisional & in-plant regulations)

Effect for Bosch	*none*	*insignificant to marginal*	*moderately serious to serious*	*very serious*
B-assessment according to FMEA	*1*	*2 - 4*	*5 - 8*	*9 - 10*
Key figure 3	***1***	***2***	***3*** X	***7***

Description of failure effect: *Outer diameter is too large, component must be scrapped, increased failure costs*

BOSCH
Quality Management

Assessment of Non-capable Measurement & Test Processes

Record No.: 9911015
Sheet 2 of 2

4. Result of assessment and measures

Product of key figures: 2 (Key figure 1) x 3 (Key figure 2) x 3 (Key figure 3) = 18 (Key figure of assessment)

Result		Decision	Measures	Responsible	Due Date
1 - 2		**Conditional release** Periodical check-up for possible process improvement.	*Optimize measuring system (prevent from manual influence)* *Repeat MSA according to procedure 1*	*Mr. ABC* *Mr. ABC*	*10/30/2009* *10/30/2009*
3 - 6		**Conditional release** Periodical check-up for possible process improvement. Proof of effectiveness of the measures in order to avoid failure effects (customer complaints).	*Optimize measuring system (prevent from manual influence)* *Repeat MSA according to procedure 1* *Safeguard characteristic by additional test XXX ...*	*Mr. ABC* *Mr. ABC* *Mr. DEF*	*10/30/2009* *10/30/2009* *11/30/2009*
≥ 7	*X*	**No release** The functionality of the characteristic must be ensured by means of a capable indirect test. If external failure implications cannot be ruled out, arrangements are established together with the development department which is in charge for the product. The customer has to be informed.	*Define a new measuring method ...* *Acquire more precise measurement equipment ...* *Repeat MSA according to procedure ...* *Safeguard characteristic by additional test XXX ...*	*Mr. ABC* *Ms. XYZ* *Mr. DEF* *Mr. DEF*	*10/30/2009* *11/30/2009* *12/10/2009* *10/1/2009*

5. Approval *(according to in-plant directives)*

	Date:	Department:	Name:	Signature:
Person in charge	*9/15/2009*	*XyP/W025*	*Amann*	*Amann*
Head of department: Production	*9/16/2009*	*XyP/MOE*	*Bemann*	*Bemann*
Head of department: Quality management	*9/17/2009*	*XyP/QMM*	*Cefrau*	*Cefrau*
Head of department: Development	*9/18/2009*	*GB/EXY*	*Demann*	*Demann*

BOSCH
Quality Management

Measurement Process Analysis

Procedure 1 in case of a one-sided limit
Acceptance criterion by means of a measurement standard

Record No.: 9911015
Sheet 1 of 1

Measuring Equipment

Location:	*XyP / MOE1*
Designation	*Pull-off force measuring device*
Equipment No.:	*JME1 E281*
Resolution:	*0.1 N*

Characteristic

Measuring Object:	*Sensor cable*
Drawing No.:	*F00X1Y2345*
Designation of Characteristic:	*Tensile shearing force of welded connection*
Nominal Value:	*90.000 N*
Limiting Value:	*70.000 N*

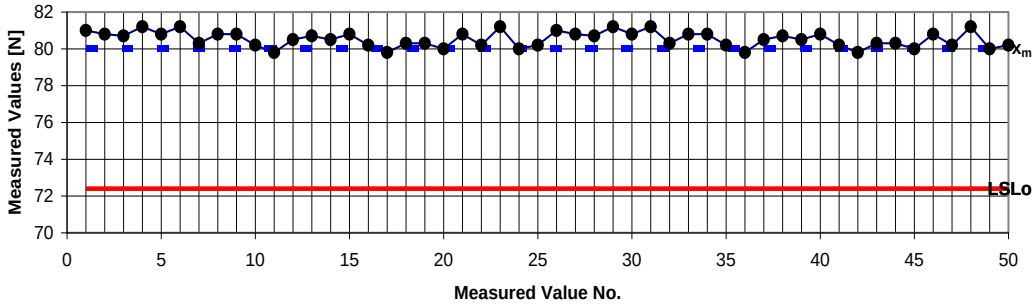
Measurement Standard

Designation:	*Load cell*
Equipment No.:	*JMP9 E2*
Reference Value x_m:	*80 N*
Uncertainty U_{cal}:	*0.2 N*

Measurement: *Manual operation using a special holding fixture for measurement standard (different to holding fixture for serial parts); room temperature 20°C*

Table values in: *N* **Deviations from:** *-*

1 - 5	6 - 10	11 - 15	16 - 20	21 - 25	26 - 30	31 - 35	36 - 40	41 - 45	46 - 50
81.000	81.200	79.800	80.200	80.800	81.000	81.200	79.800	80.200	80.800
80.800	80.300	80.500	79.800	80.200	80.800	80.300	80.500	79.800	80.200
80.700	80.800	80.700	80.300	81.200	80.700	80.800	80.700	80.300	81.200
81.200	80.800	80.500	80.300	80.000	81.200	80.800	80.500	80.300	80.000
80.800	80.200	80.800	80.000	80.200	80.800	80.200	80.800	80.000	80.200



Reference value x_m= 80 N **Mean value** $\bar{x}$ = 80.524 N **Standard deviation s =** 0.4182 N

Systematic measurement error:

For n = 25 the systematic measurement error is significant if $|\bar{x} - x_m| > 0.413 \cdot s$

For n = 50 the systematic measurement error is significant if $|\bar{x} - x_m| > 0.284 \cdot s$

$|\bar{x} - x_m|$ = 0.524 N 0.413*s = 0.1727 N 0.284*s = 0.1188 N

[ ] insignificant [X] significant [ ] not assessed

Acceptance criterion for measured values z that are determined during the production process:

[ ] Upper limiting value [N] $z \le USL_0 = USL + (\bar{x} - x_m) - 4 \cdot s - U_{cal}$

[X] Lower limiting value [N] $z \ge LSL_0 = LSL + (\bar{x} - x_m) + 4 \cdot s + U_{cal}$

*= 70 + (80.524 - 80) + 4 * 0.4182 + 0.2 =* **72.3968**

Comment: *The systematic measurement error cannot be corrected; reason: shear forces due to holding fixture; measure: minimization of shear forces by optimizing the holding fixture; the study will be repeated with optimized holding fixture*

Date: *02/29/2009* Department: *XyP / MOE1* Name: *John Q. Public* Signature: *John Q. Public*

2020-04-06 - SOCOS

C Amendments and Notes on Procedure 1

C.1 Significance of systematic measurement errors

For sample sizes n ≥ 20, [AIAG MSA] recommends to check the significance of the systematic measurement error $\left|\bar{x} - x_m\right|$ (see [AIAG MSA], footnote on page 88/89). The significance criterion of this so-called one-sample t-test is dependent on the confidence level 1 − α and on the sample size n. For the confidence level 95%, the systematic measurement error is considered **<u>insignificant</u>** if the criterion

$\frac{\left|\bar{x} - x_m\right|}{s} \le 0.413$ is met for sample size n = 25

or

$\frac{\left|\bar{x} - x_m\right|}{s} \le 0.284$ is met for sample size n = 50.

For a deviating confidence level and/or sample size the criterion must be adapted accordingly (see Appendix C.2).

This criterion implies a condition for the maximum difference of the parameters C_g and C_{gk}:

$$C_g - C_{gk} \le \frac{1}{3} \cdot \frac{\left|\bar{x} - x_m\right|}{s}.$$

NOTE: This condition results if the equation for C_{gk} is solved for $\left|\bar{x} - x_m\right| / s$ and C_g is substituted for that part of the resulting formula which corresponds to the equation for C_g (for equations see chapter 4.1).

The criterion for the respective sample size n inserted yields for

n = 25 the equivalent criterion $C_g - C_{gk} \le 0.138$

or for

n = 50 the equivalent criterion $C_g - C_{gk} \le 0.095$.

According to practical experience, this criterion leads to problems with high-quality measurement standards and high-quality measuring equipment (s small, criterion is not satisfied despite technically excellent small measurement errors) or it leads to problems with low-quality standards and equipment that cannot be seen directly (s large, criterion is satisfied despite technically unacceptably large measurement errors). This is due to the evaluation of the systematic measurement error relative to the variation s of the measurement process. It is not evaluated relative to the technically relevant tolerance of the characteristic to be measured. Thus,

$$C_g - C_{gk} \le \frac{1}{3} = 0.33$$

is sometimes used as a rule of thumb for just acceptable deviations in practice, i.e. systematic measurement errors up to s. The applicability must be assessed for every individual measurement process.

According to [AIAG MSA] a significant systematic measurement error should generally be corrected by modification to the measuring equipment (e. g. adjustment). If this is not possible, the systematic measurement error can be taken into account by correcting each measurement result (see [AIAG MSA], chapter III, paragraph B, page 95).

C.2 Determination of the significance criterion for systematic measurement errors

A one-sample t-test is used to test whether the mean value μ of the population agrees or disagrees with the reference value x_m. The mean value $\bar{x}$ of a sample of size n from this population is used as an estimator for μ.

Null hypothesis: $\mu = x_m$

Alternative hypothesis: $\mu \ne x_m$

The null hypothesis is accepted if $\left|\bar{x} - x_m\right| \le t_{f;\,1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$.

$t_{f;\,1-\alpha/2}$ is the (two-sided) quantile of the t-distribution for f = n – 1 degrees of freedom and confidence level 1 – α.

Thus, the significance criterion for α = 0.05 (confidence level 95%) and sample size n = 25 is calculated according to

$$\frac{\left|\bar{x} - x_m\right|}{s} \le \frac{t_{24;\,0.975}}{\sqrt{25}} = \frac{2.064}{5} = 0.413$$

and for n = 50 according to

$$\frac{\left|\bar{x} - x_m\right|}{s} \le \frac{t_{49;\,0.975}}{\sqrt{50}} = \frac{2.009}{7.071} = 0.284\,.$$

If the significance criterion is satisfied, the mean value μ and the reference value x_m are **not** significantly different. Significance criteria for further sample sizes n and α = 0.05:

n	$\frac{t_{n-1;\,1-\alpha/2}}{\sqrt{n}}$		n	$\frac{t_{n-1;\,1-\alpha/2}}{\sqrt{n}}$		n	$\frac{t_{n-1;\,1-\alpha/2}}{\sqrt{n}}$
5	1.241664		30	0.373406		55	0.270338
10	0.715357		35	0.343512		60	0.258327
15	0.553782		40	0.319816		65	0.247788
20	0.468014		45	0.300433		70	0.238442
25	0.412780		50	0.284197		75	0.230079

Significance criteria for deviating values of α and n are calculated in the same way. $t_{f;\,1-\alpha/2}$ can be found in tables or determined e. g. with the EXCEL worksheet function *TINV(α;f)*.

C.3 Characteristics with a one-sided limit and without a natural limit

The systematic measurement error $\bar{x} - x_m$ and the standard deviation s are calculated from the measured data of a type-1 study according to chapter 4.1. Both parameters are assumed to be temporally insignificantly changing characteristics of the measuring equipment (but not of the measuring object).

Let z be a single measured value, recorded during the production process. It is assumed that z belongs to a normal distribution with the standard deviation s (known from the type-1 study described in chapter 4.1), but with unknown mean value $\bar{z}$.

NOTE: It is assumed that a normal distribution with a standard deviation s and a certain mean value $\bar{z}$ would result if the measurements were repeated sufficiently frequently.

The additional requirement that z must belong to a certain distribution with 99.994% probability limits the possible distributions to the distributions between the following extreme positions:

- z coincides with the quantile +4s of the distribution with the mean value $\bar{z} = z - 4 \cdot s$ (upper edge position),
- z coincides with the quantile –4s of the distribution with the mean value $\bar{z} = z + 4 \cdot s$ (lower edge position).

Thus, the distributions which z can belong to, are limited to distributions with a mean value $\bar{z}$ in the range

$$z - 4 \cdot s \le \bar{z} \le z + 4 \cdot s\,.$$

Based on the supposed insignificant temporal change of the systematic measurement error $\bar{x} - x_m$ (known from the type-1 study described in chapter 4.1), it is further assumed that the conventional true value z_0 differs from $\bar{z}$ by exactly this measurement error:

$$\bar{z} - z_0 = \bar{x} - x_m\,.$$

This relationship solved for $\bar{z}$ and inserted, results in

$$z - 4 \cdot s \le z_0 + \bar{x} - x_m \le z + 4 \cdot s\,.$$

This inequality solved for z_0 yields the range where the conventional true value z_0 corresponding to the measured value z can be expected with a probability of 99.994%:

$$z - (\bar{x} - x_m) - 4 \cdot s \le z_0 \le z - (\bar{x} - x_m) + 4 \cdot s\,.$$

In case of an upper limit USL, it is requested that the conventional true value z_0 corresponding to the measured value z must not be greater than USL, i.e.

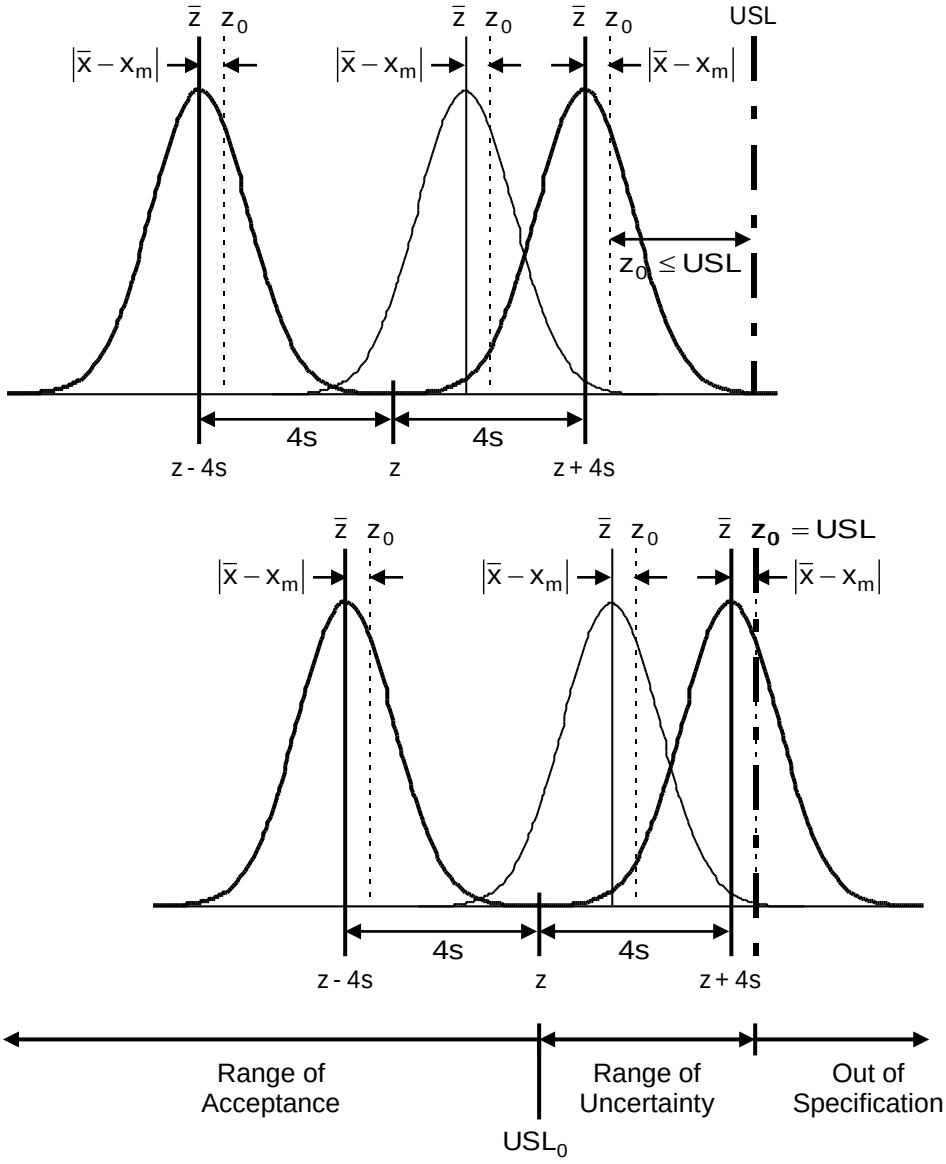
$$z_0 \le z - (\bar{x} - x_m) + 4 \cdot s \le USL\,.$$

Solved for z, the acceptance criterion for each individual measured value z results:

$$\boxed{z \le USL + (\bar{x} - x_m) - 4 \cdot s = USL_0}\,.$$

The following diagrams illustrate the explanations above. A measured value z is shown in an uncritical position (upper diagram) and in critical position (lower diagram) relative to a one-sided upper limit USL. Furthermore, the two distributions in the extreme positions as well as an example of a distribution in an intermediate position are shown, each with its mean value $\bar{z}$ and the corresponding conventional true value z_0. z can belong to any of these distributions.

z must only come as close to the limit USL as the greatest of all possible conventional true values z_0 does not violate USL ($z_0 \leq$ USL). In the lower diagram, z_0 coincides with USL (z_0 = USL) and z is located exactly at the upper edge of the acceptance range (z = USL_0), i.e. the acceptance range $z \leq USL_0$ is represented by the range "*left of z*" in the lower diagram.



In case of a lower limiting value LSL

$$LSL \leq z - (\bar{x} - x_m) - 4 \cdot s \leq z_0$$

is requested analogically, i.e. the acceptance criterion results from solving the unequality to z:

$$\boxed{z \geq LSL + (\bar{x} - x_m) + 4 \cdot s = LSL_0}$$

In cases where the expanded measurement uncertainty U_{cal} of the calibration of the measurement standard has to be considered (rule of thumb: $U_{cal} \geq 0{,}01 \cdot (|\bar{x} - x_m| + 4 \cdot s)$), the criteria above apply in the following modified form:

$$z \leq USL + (\bar{x} - x_m) - 4 \cdot s - U_{cal} = USL_0$$

$$z \geq LSL + (\bar{x} - x_m) + 4 \cdot s + U_{cal} = LSL_0 \,.$$

D Amendments and Notes on Procedures 2 and 3

D.1 Parameter ndc

ndc (**n**umber of **d**istinct **c**ategories) is the number of categories that can be distinguished by the measurement process. ndc describes the part variation PV related to the measurement process variation GRR. In order to ensure that the measuring equipment can measure differing part qualities with sufficiently distinct results, the part variation should be larger than the variation of the measuring equipment in the case of type-2 and type-3 studies. According to [AIAG MSA] ndc should not be less than 5:

$$\text{ndc} = \sqrt{2} \cdot \frac{\text{PV}}{\text{GRR}} \geq 5\,.$$

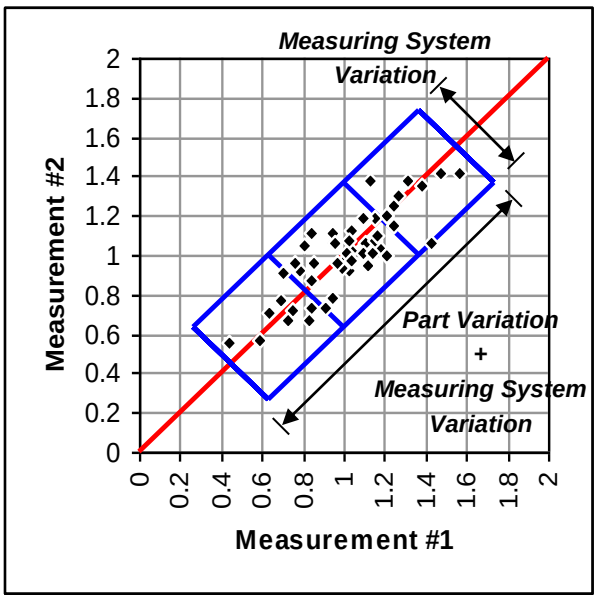
NOTE: Non-integer numbers for ndc are always rounded to the next integer.

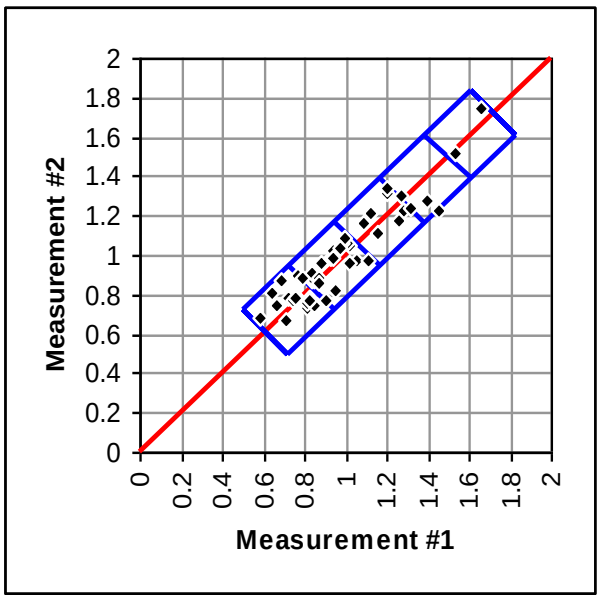
Illustrative interpretation of ndc

When using procedure 3, (at least) 25 parts are measured repeatedly, i.e. there are 2 measured values for each part. When plotting the measurement results against each other so that each part is represented by a data point whose x-coordinate represents the measured value of the first measurement and whose y-coordinate represents the measured value of the second measurement, a diagram is obtained as shown below (so-called iso-plot).

If the measurement results were different from part to part whereas the results of the first and second measurement were identical for each individual part, the corresponding data points in the diagram would be situated exactly on the diagonal. Thus, the scattering of the data points around the diagonal (i.e. their deviation from the diagonal) is an approximate measure for the variation GRR of the measuring equipment, while the scattering along the diagonal is an approximate measure for the part variation PV (plus a GRR-portion).

ndc can be interpreted as the number of squares that is necessary to cover the entire scattering area. The edge length of the squares is determined by the measurement process variation: the smaller the measurement process variation, the shorter the edge length of the squares and the more squares are necessary to cover the scattering area. Thus, ndc ≥ 5 corresponds to 5 or more squares. According to [AIAG MSA] the variation of the measurement process is sufficiently small in relation to the variation of the production process in this case.





The quantitative deduction of ndc is based on an approach similar to the signal-to-noise ratio [Wheeler].

Relationship to process-related %GRR

The capability criterion

$$\%GRR = \frac{GRR}{\sqrt{GRR^2 + PV^2}} \cdot 100\% \leq GRR_{max} \cdot 100\%$$

solved for PV/GRR and multiplied by $\sqrt{2} = 1.41$ yields

$$\sqrt{2 \cdot \left(\frac{1}{GRR_{max}^2} - 1 \right)} \leq \sqrt{2} \cdot \frac{PV}{GRR} = ndc\,.$$

Thus, ndc can be interpreted as an alternative representation of the process-related parameter %GRR. With $\%GRR_{max} = 0.3$, $ndc \geq 4.50 \approx 5$ is obtained; with $\%GRR_{max} = 0.1$, $ndc \geq 14.07 \approx 14$ is obtained.

Problems with process-related parameters

Process-related parameters do not contain any technically relevant criterion such as the tolerance of the characteristic to be measured. This can lead to the following misinterpretations:

- For very small PV, i.e. excellent results of the production process concerning the part variation, ndc approximates 0 and %GRR approximates 100%, i.e. the measurement process would have to be classified as not capable.
- For very large PV, i.e. poor results of the production process concerning the part variation, ndc is very large and %GRR approximates 0%, i.e. the measurement process would have to be classified as unconditionally capable.

Results of this kind must be analyzed thoroughly since the classification of the measurement process is mainly determined by the production process in these cases and may technically not be justified.

D.2 Procedure 2: Analysis using “analysis of variances (ANOVA)”

The basic idea of ANOVA (**an**alysis **o**f **va**riances) is the decomposition of the total variation into components that are allocated to certain influence quantities. Model for an individual measured value x_{ijm} with normally distributed random variables (without systematic measurement errors, bias):

		Variable	Distribution
	Mean value of all parts	μ	–
+	Influence of part no. i	α_i	$N(0, \sigma^2)$
+	Influence of appraiser no. j	β_j	$N(0, \omega^2)$
+	Influence of interactions between part i and appraiser j	λ_{ij}	$N(0, \gamma^2)$
+	Influence of measuring equipment (random measurement error)	ε_{ijm}	$N(0, \tau^2)$
=	Measured value x_{ijm}	x_{ijm}	

Thus, the *single* measured value x_{ijm} is calculated according to

$$x_{ijm} = \mu + \alpha_i + \beta_j + \lambda_{ij} + \varepsilon_{ijm}$$

with

i = 1, ..., n; n - Number of parts
j = 1, ..., k; k - Number of appraisers
m = 1, ..., r; r - Number of measurements per part and appraiser

Accordingly, the total variance of *all* measured values consists of individual components according to

$$VAR(x_{ijm}) = \sigma^2 + \omega^2 + \gamma^2 + \tau^2$$

which have to be determined. An estimator for the total variance is calculated from the measured values according to

$$\hat{VAR}(x_{ijm}) = s^2 = \frac{1}{nkr-1} \cdot TSS .$$

Here

$$TSS = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r}\left(x_{ijm} - \frac{x_{\bullet\bullet\bullet}}{nkr}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r} x_{ijm}^2 - \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

represents the sum of *all* squared measurement errors (**t**otal **s**um of **s**quares) and

$$\frac{x_{\bullet\bullet\bullet}}{nkr} = \frac{1}{nkr} \cdot \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r} x_{ijm}$$

represents the mean value of *all* measured values.

NOTE: A dot instead of an index denotes that the summation over that index has been carried out.

TSS is decomposed into the components SS (**s**um of **s**quares) according to

$$TSS = SS_P + SS_A + SS_{AP} + SS_E .$$

These SS components are allocated to the influences listed above:

Influence quantity	SS Component	Degree of freedom DF		
Parts	SS_P	DF_P	=	n - 1
Appraisers	+ SS_A	DF_A	=	k - 1
Interactions between appraisers and parts	+ SS_{AP}	DF_{AP}	=	(n - 1)·(k - 1)
Measuring equipment (random measurement error)	+ SS_E	DF_E	=	nk·(r - 1)
Total variation	= TSS	DF_{TSS}	=	nkr - 1

Balance of **d**egrees of **f**reedom (**DF**):

$$(n-1) + (k-1) + (n-1)\cdot(k-1) + nk\cdot(r-1) = nkr - 1.$$

The individual SS components are calculated according to

$$SS_P = \sum_{i=1}^{n} \frac{x_{i\bullet\bullet}^2}{kr} - \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

$$SS_A = \sum_{j=1}^{k} \frac{x_{\bullet j\bullet}^2}{nr} - \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

$$SS_{AP} = \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r} - \frac{x_{\bullet\bullet\bullet}^2}{nkr} - SS_P - SS_A = \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r} - \sum_{i=1}^{n} \frac{x_{i\bullet\bullet}^2}{kr} - \sum_{j=1}^{k} \frac{x_{\bullet j\bullet}^2}{nr} + \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

$$SS_E = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r} x_{ijm}^2 - \frac{x_{\bullet\bullet\bullet}^2}{nkr} - SS_P - SS_A - SS_{AP} = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{m=1}^{r} x_{ijm}^2 - \sum_{i=1}^{n}\sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r}$$

Each of the sums SS_x divided by the corresponding number of degrees of freedom DF_x yields MS_x (**m**ean of **s**quares):

$$MS_X = \frac{SS_X}{DF_X}.$$

The index X represents the indexes P, A, AP and E, respectively.

An F-test is used to analyze the significance of an influence quantity. The parameters are calculated as follows:

Influence quantity	Test statistic	Quantile of F-distribution
Parts	$F_P = \frac{MS_P}{MS_{AP}}$	$F_{P\,crit} = F_{DF_P;\,DF_{AP};\,1-\alpha}$
Appraisers	$F_A = \frac{MS_A}{MS_{AP}}$	$F_{A\,crit} = F_{DF_A;\,DF_{AP};\,1-\alpha}$
Interaction	$F_{AP} = \frac{MS_{AP}}{MS_E}$	$F_{AP\,crit} = F_{DF_{AP};\,DF_E;\,1-\alpha}$

An influence quantity is significant if the corresponding criterion $F_X > F_{X\,crit}$ is satisfied. $F_{X\,crit}$ can be found in tables or determined using the EXCEL worksheet function ***=FINV(α;DF$_{numerator}$;DF$_{denominator}$)***.

If all influence quantities are shown to be significant, the estimators for the individual variance components σ^2, ω^2, γ^2 and τ^2 are calculated:

$$\hat{\sigma} = PV = \sqrt{\frac{MS_P - MS_{AP}}{k\,r}}$$ **P**art **v**ariation

$$\hat{\omega} = AV = \sqrt{\frac{MS_A - MS_{AP}}{n\,r}}$$ **A**ppraiser **v**ariation (reproducibility)

$$\hat{\gamma} = INT = \sqrt{\frac{MS_{AP} - MS_E}{r}}$$ **Int**eraction appraiser – part

$$\hat{\tau} = EV = \sqrt{MS_E}$$ **E**quipment **v**ariation (repeatability)

$$GRR = \sqrt{EV^2 + AV^2 + INT^2}$$ **G**auge **r**epeatability & **r**eproducibility

$$TV = \sqrt{GRR^2 + PV^2}$$ **T**otal **v**ariation (including part variation)

Example

Data according to [AIAG MSA], page 118:

Measured values x_{ijm}		Appraiser j = 1 (appraiser A)			j = 2 (appraiser B)			j = 3 (appraiser C)		
		Measurement			Measurement			Measurement		
		m = 1	m = 2	m = 3	m = 1	m = 2	m = 3	m = 1	m = 2	m = 3
Part	i = 1	0.29	0.41	0.64	0.08	0.25	0.07	0.04	-0.11	-0.15
	i = 2	-0.56	-0.68	-0.58	-0.47	-1.22	-0.68	-1.38	-1.13	-0.96
	i = 3	1.34	1.17	1.27	1.19	0.94	1.34	0.88	1.09	0.67
	i = 4	0.47	0.50	0.64	0.01	1.03	0.20	0.14	0.20	0.11
	i = 5	-0.80	-0.92	-0.84	-0.56	-1.20	-1.28	-1.46	-1.07	-1.45
	i = 6	0.02	-0.11	-0.21	-0.20	0.22	0.06	-0.29	-0.67	-0.49
	i = 7	0.59	0.75	0.66	0.47	0.55	0.83	0.02	0.01	0.21
	i = 8	-0.31	-0.20	-0.17	-0.63	0.08	-0.34	-0.46	-0.56	-0.49
	i = 9	2.26	1.99	2.01	1.80	2.12	2.19	1.77	1.45	1.87
	i = 10	-1.36	-1.25	-1.31	-1.68	-1.62	-1.50	-1.49	-1.77	-2.16

Following the calculation steps above (according to [AIAG MSA], page 198) up to the F-tests yields the following table (so-called ANOVA table):

Influence quantity		SS		DF		MS		F	F_{crit}	
Parts	SS_P	88.3619	DF_P	9	MS_P	9.81799	F_P	492.291	2.456	significant
Appraisers	SS_A	3.1673	DF_A	2	MS_A	1.58363	F_A	79.406	3.555	significant
Interaction	SS_{AP}	0.3590	DF_{AP}	18	MS_{AP}	0.01994	F_{AP}	0.434	1.778	not significant
Measuring equipment	SS_E	2.7589	DF_E	60	MS_E	0.04598				
Total	TSS	94.6471	DF_{TSS}	89						

For this example, the test of significance with confidence level 95% ($\alpha = 0.05$) shows that the interaction between appraisers and parts is insignificant. Thus, the calculation of variance components according to the above equations is inappropriate and can be omitted here.

Modification of the calculation model for insignificant interactions

The influence quantity *interaction* is removed from the model which requires a recalculation of the variables in the table above according to modified equations. SS_E and SS_{AP} are combined according to

$$SS_E^* = SS_E + SS_{AP}$$

with

$$DF_E^* = DF_E + DF_{AP}$$

degrees of freedom. MS_E and MS_{AP} are replaced by

$$MS_E^* = \frac{SS_E^*}{DF_E^*}.$$

Next, the statistics for the F-test are calculated according to

$$F_P = \frac{MS_P}{MS_E^*} \quad \text{with} \quad F_{P\,krit} = F_{DF_P;\,DF_E^*;\,1-\alpha}$$

$$F_A = \frac{MS_A}{MS_E^*} \quad \text{with} \quad F_{A\,krit} = F_{DF_A;\,DF_E^*;\,1-\alpha}.$$

Finally, the estimators for the variance components σ^2, ω^2, γ^2 and τ^2 are recalculated according to the following modified equations (MS_E and MS_{AP} formally replaced by MS_E^*):

$$\hat{\sigma} = PV = \sqrt{\frac{MS_P - MS_E^*}{k\,r}}$$ **P**art **v**ariation

$$\hat{\omega} = AV = \sqrt{\frac{MS_A - MS_E^*}{n\,r}}$$ **A**ppraiser **v**ariation (reproducibility)

$$\hat{\gamma} = INT = \sqrt{\frac{MS_E^* - MS_E^*}{r}} = 0$$ **Int**eraction appraiser – part

$$\hat{\tau} = EV = \sqrt{MS_E^*}$$ **E**quipment **v**ariation (repeatability)

$$GRR = \sqrt{EV^2 + AV^2}$$ **G**auge **r**epeatability & **r**eproducibility

$$TV = \sqrt{GRR^2 + PV^2}$$ **T**otal **v**ariation (including part variation)

Example (continued)

The following ANOVA table results for the data from [AIAG MSA]:

Influence quantity		SS		DF		MS		F	F_{crit}	
Parts	SS_P	88.3619	DF_P	9	MS_P	9.81799	F_P	245.614	2.002	significant
Appraisers	SS_A	3.1673	DF_A	2	MS_A	1.58363	F_A	39.617	3.114	significant
Measuring equipment	SS_E*	3.1179	DF_E*	78	MS_E*	0.03997				
Total	TSS	94.6471	DF_{TSS}	89						

The remaining influence quantities are shown to be significant so that the variance components σ^2, ω^2, γ^2 and τ^2 are calculated. The results are usually given as

- **S**tandard **d**eviation **SD**: according to equations above (PV, AV, etc.)
- **Var**iance **Var**: **Var** = SD^2
- Variance Var related to total variance TV^2: **%Var** = Var / TV^2·100% = SD^2 / TV^2·100%
- **S**tudy **v**ariation **SV** (process spread): **SV** = 6·SD
- Study variation SV related to total variation 6·TV: **%SV** = SV / (6·TV)·100% = SD / TV·100%·
- Study variation SV related to tolerance T: **%T** = SV / T·100% = 6·SD / T·100%

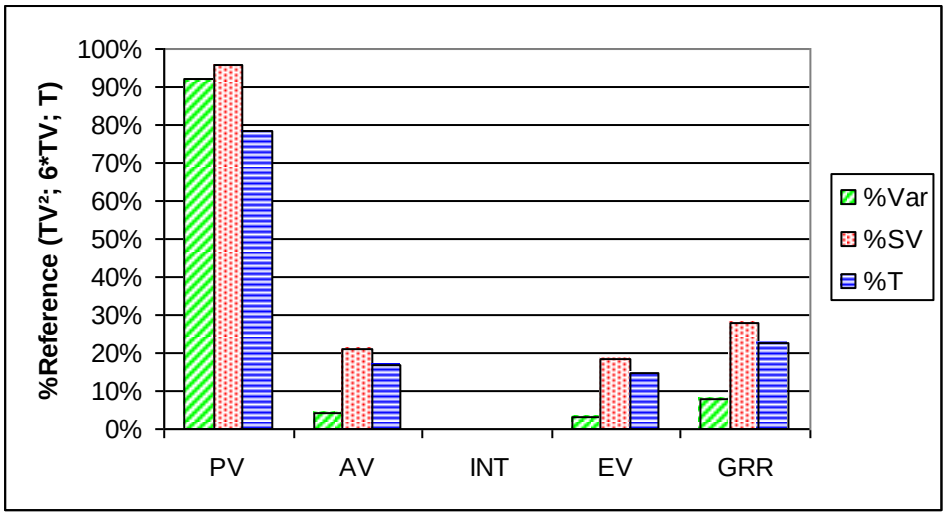
and summarized in a table:

Influence quantity		Standard deviation (estimator)	Variance (estimator)	Percentage of total variance [1]	Study variation (99.73%) [2]	Percentage of total variance	Percentage of tolerance
		SD	Var	%Var	SV	%SV	%T
Parts (part variation)	**PV**	1.04233	1.08645	92.24%	6.25396	96.04%	78.17%
Appraisers (appraiser variation)	**AV**	0.22684	0.05146	4.37%	1.36103	20.90%	17.01%
Interaction	**INT**	0	0	0%	0	0%	0%
Equipment (equipment variation)	**EV**	0.19993	0.03997	3.39%	1.19960	18.42%	14.99%
GRR	**GRR**	0.30237	0.09143	7.76%	1.81423	27.86%	22.68%
Total (total variation)	**TV**	1.08530	1.17788	100.00%	6.51180	100.00%	81.40%

NOTE 1: The sum of the %SV percentages does not add up to 100% since the individual SV components are represented by the (6-fold) standard deviations SD which do not add up arithmetically but geometrically to the total variation TV (square root of the sum of squares of the individual SD components). For the same reason, the %T percentages do not add up to the total percentage of tolerance.

NOTE 2: Tolerance T = 8 is used in the example.

Summary of the percentaged results as a diagram:



Because of 10% < %GRR ≤ 30% the measurement process is only conditionally capable in this example.

[1] In English literature (and in the software Minitab®) also named *%contribution.*

[2] AIAG MSA, edition 3, here (inconsistently) uses the range 99% and thus the factor 5.15 (instead of 6.00). This is meaningless for the percentaged results %SV (factor is used in numerator and denominator) but not for %T. This is corrected in AIAG MSA, edition 4. However, the definitions of PV, AV, and EV as sixfold standard deviations in case of ANOVA continue to be inconsistent with ARM where PV, AV and EV correspond to single standard deviations. Thus, in contrast to AIAG MSA, these quantities are uniformly defined as single standard deviations throughout the present booklet. So uniform formulas are applicable for tolerance-related quantities calculated according to ANOVA or ARM (e. g. %GRR = 6 * GRR / T), i.e. the factor 6 always has to be taken into account.

Analysis of procedure 3

The formalism above can be applied to procedure 3. The number of appraisers has to be k = 1. This leads to several simplifications (omission of $SS_A = 0$ and $SS_{AP} = 0$).

D.3 Analysis using the "average range method (ARM)"

Analyses using the so-called average range method (ARM) are no longer up-to-date and are generally not recommended. A considerable disadvantage is, amongst others, that interactions between appraiser and part cannot be considered. Thus, using ARM should be limited to exceptional cases and agreed with the customer, if necessary.

The basic idea of ARM is equal to ANOVA, i.e. the decomposition of the variation into components that are allocated to the influence quantities *parts, appraisers* and *measuring equipment*.

Scenario

r measurements are done at n parts by k appraisers. The measured values x_{ijm} are documented.

i	=	1, ..., n;	n	-	Number of parts
j	=	1, ..., k;	k	-	Number of appraisers
m	=	1, ..., r;	r	-	Number of measurements per part and appraiser

Calculations

$R_{ij} = \mathrm{Max}(x_{ijm}) - \mathrm{Min}(x_{ijm})$	Range of all measurement results of appraiser j for part i
$\overline{R}_j = \frac{1}{n}\sum_{i=1}^{n} R_{ij}$	Mean value of all ranges R_{ij} of appraiser j
$\overline{\overline{R}} = \frac{1}{k}\sum_{j=1}^{k} \overline{R}_j$	Mean value of all average ranges $\overline{R}_j$
$EV = K_1 \cdot \overline{\overline{R}}$	**E**quipment **v**ariation (repeatability)
$\overline{x}_j = \frac{1}{n \cdot r}\sum_{i=1}^{n}\sum_{m=1}^{r} x_{ijm} = \frac{1}{n \cdot r} \cdot x_{\bullet j \bullet}$	Mean value of all measurement results of appraiser j
$R_{\overline{x}} = \mathrm{Max}(\overline{x}_j) - \mathrm{Min}(\overline{x}_j)$	Range of all mean values $\overline{x}_j$
$AV = \sqrt{(K_2 \cdot R_{\overline{x}})^2 - \frac{EV^2}{n \cdot r}}$	**A**ppraiser **v**ariation (reproducibility)
$GRR = \sqrt{EV^2 + AV^2}$	**G**auge **r**epeatability & **r**eproducibility
$\overline{x}_i = \frac{1}{k \cdot r}\sum_{j=1}^{k}\sum_{m=1}^{r} x_{ijm} = \frac{1}{k \cdot r} \cdot x_{i \bullet \bullet}$	Mean value of all measurement results for part i
$R_P = \mathrm{Max}(\overline{x}_i) - \mathrm{Min}(\overline{x}_i)$	Range of all mean values $\overline{x}_i$
$PV = K_3 \cdot R_P$	**P**art **v**ariation
$TV = \sqrt{GRR^2 + PV^2}$	**T**otal **v**ariation

The factors K_1, K_2 and K_3 are determined depending on the number of parts n, the number of appraisers k and the number of measurements r (see also Appendix D.4).

Example

Analysis of data according to [AIAG MSA], page 118:

Measured values x_{ijm}		Appraiser j = 1 (appraiser A)				Appraiser j = 2 (appraiser B)				Appraiser j = 3 (appraiser C)				Mean values
		Measurement			Ranges	Measurement			Ranges	Measurement			Ranges	
		m = 1	m = 2	m = 3	R_{ij}	m = 1	m = 2	m = 3	R_{ij}	m = 1	m = 2	m = 3	R_{ij}	$\bar{x}_i$
Part	i = 1	0.29	0.41	0.64	**0.35**	0.08	0.25	0.07	**0.18**	0.04	-0.11	-0.15	**0.19**	**0.169**
	i = 2	-0.56	-0.68	-0.58	**0.12**	-0.47	-1.22	-0.68	**0.75**	-1.38	-1.13	-0.96	**0.42**	**-0.851**
	i = 3	1.34	1.17	1.27	**0.17**	1.19	0.94	1.34	**0.40**	0.88	1.09	0.67	**0.42**	**1.099**
	i = 4	0.47	0.50	0.64	**0.17**	0.01	1.03	0.20	**1.02**	0.14	0.20	0.11	**0.09**	**0.367**
	i = 5	-0.80	-0.92	-0.84	**0.12**	-0.56	-1.20	-1.28	**0.72**	-1.46	-1.07	-1.45	**0.39**	**-1.064**
	i = 6	0.02	-0.11	-0.21	**0.23**	-0.20	0.22	0.06	**0.42**	-0.29	-0.67	-0.49	**0.38**	**-0.186**
	i = 7	0.59	0.75	0.66	**0.16**	0.47	0.55	0.83	**0.36**	0.02	0.01	0.21	**0.20**	**0.454**
	i = 8	-0.31	-0.20	-0.17	**0.14**	-0.63	0.08	-0.34	**0.71**	-0.46	-0.56	-0.49	**0.10**	**-0.342**
	i = 9	2.26	1.99	2.01	**0.27**	1.80	2.12	2.19	**0.39**	1.77	1.45	1.87	**0.42**	**1.940**
	i = 10	-1.36	-1.25	-1.31	**0.11**	-1.68	-1.62	-1.50	**0.18**	-1.49	-1.77	-2.16	**0.67**	**-1.571**
Mean values $\bar{R}_j$					**0.184**				**0.513**				**0.328**	
Mean $\bar{\bar{R}}$					**0.34167**									
Mean values $\bar{x}_j$		**0.190**				**0.068**				**-0.254**				
Range $R_{\bar{x}}$		**0.44467**												
Range R_P														**3.511**

For n = 10 parts, k = 3 appraisers and m = 3 measurements per appraiser and part, the following K-factors apply:

$K_1 = 0.5908$ $K_2 = 0.5231$ $K_3 = 0.3146$

Summary of results:

Influence quantities		Standard deviation (estimator) SD	Variance (estimator) Var	Percentage of total variance %Var	Study variation (99.73%) SV	Percentage of total variance %SV	Percentage of tolerance %T
Parts (part variation)	**PV**	1.10445	1.21982	92.88%	6.62672	96.37%	82.83%
Appraisers (appraiser variation)	**AV**	0.22968	0.05275	4.02%	1.37810	20.04%	17.23%
Equipment (equipment variation)	**EV**	0.20186	0.04075	3.10%	1.21118	17.61%	15.14%
GRR	**GRR**	0.30578	0.09350	7.12%	1.83469	26.68%	22.93%
Total (total variation)	**TV**	1.14600	1.31332	100.00%	6.87601	100.00%	85.95%

Explanations: See corresponding table in Appendix D.2, page 51.

D.4 K-factors for ARM

The unknown standard deviation σ of a normally distributed population is usually estimated by the standard deviation s of a sample $x_1, x_2, ..., x_m$ that is taken from this population, i.e. $\hat{\sigma} = s$.

However, it is also possible to estimate σ using the range $R = x_{max} - x_{min}$ of this sample. It seems logical that R increases with increasing sample size m: the more values in a sample, the higher the probability that it contains very large and very small values from the "tails" of the normal distribution.

If samples of size m are taken repeatedly from a population with standard deviation σ, then an average range $\bar{R}_m = d_2 \cdot \sigma$ will be obtained.

If the number of samples is very large, $\hat{\sigma} = \overline{R}_m / d_2$ is an estimate for σ. The constant d_2 can be determined using the distribution of standardized ranges.

The matter becomes more complicated if the number of samples g is very small: $\hat{\sigma} = \overline{R}_{g;m} / d_2^*$. In this case, d_2^* can be determined by means of an approximate distribution only.

d_2 is the limiting value of d_2^* for an indefinite number of samples. Values for d_2^* can be found in tables (e. g. [AIAG MSA], page 203). Since d_2^* approximates the limiting value d_2 very quickly, d_2^* is usually tabulated for values up to g = 20 only. In case of a larger number of samples, d_2 is used instead.

Procedure 2: Constants K_1, K_2 and K_3

The **standard deviation EV** (*equipment variation*) is determined from a total of 30 ranges (10 parts x 3 appraisers) by means of double samples (2 measurements per part and appraiser), i.e. g = 30 samples each consisting of m = 2 measured values. d_2^* for g = 30 is not contained in the table so that the limiting value d_2 = 1.12838 is used:

$$EV = \frac{\overline{R}_{30;2}}{d_2} = \frac{\overline{\overline{R}}}{1.12838} = 0.8862 \cdot \overline{\overline{R}}, \qquad \text{i.e. } K_1 = 0.8862 \text{ for 2 measurements.}$$

If each appraiser carries out m = 3 instead of m = 2 measurements per part, the limiting value d_2 = 1.69257 is used:

$$EV = \frac{\overline{R}_{30;3}}{d_2} = \frac{\overline{\overline{R}}}{1.69257} = 0.5908 \cdot \overline{\overline{R}}, \qquad \text{i.e. } K_1 = 0.5908 \text{ for 3 measurements.}$$

The **standard deviation AV** (*appraiser variation*) is determined from the range of the 3 mean values of each appraiser, i.e. g = 1 sample consisting of m = 3 mean values. The tabular value is d_2^* = 1.91155:

$$AV = \frac{R_{1;3}}{d_2^*} = \frac{R}{1.91155} = 0.5231 \cdot \overline{\overline{R}}, \qquad \text{i.e. } K_2 = 0.5231 \text{ for 3 appraisers,}$$

or with the tabular value d_2^* = 1.41421 for only 2 appraisers and 2 mean values:

$$AV = \frac{R_{1;2}}{d_2^*} = \frac{R}{1.41421} = 0.7071 \cdot \overline{\overline{R}}, \qquad \text{i.e. } K_2 = 0.7071 \text{ for 2 appraisers.}$$

NOTE: The equation for AV according to [AIAG MSA] additionally considers a correction term that contains EV.

The **standard deviation PV** (*part variation*) is determined from the range R_p of the 10 mean values of each part, i.e. g = 1 sample consisting of m = 10 mean values. The tabular value is d_2^* = 3.17905:

$$PV = \frac{R_{1;10}}{d_2^*} = \frac{R_p}{3.17905} = 0.3146 \cdot R_p, \qquad \text{i.e. } K_3 = 0.3146 \text{ for 10 parts.}$$

In case of deviating numbers (measurements, appraisers, parts) the values of the K-factors (K_1, K_2, K_3) must be adapted accordingly. Otherwise, the analysis according to ARM leads to incorrect results.

Procedure 3: Constants K_1 and K_3

In contrast to procedure 2, the **standard deviation EV** is determined from a total of 25 ranges (25 parts) from double samples (2 measurements per part), i.e. g = 25 samples each consisting of m = 2 measured values. d_2^* for g = 25 is not contained in the table. Instead, the value d_2 = 1.12838 is used:

$$EV = \frac{\overline{R}_{25;2}}{d_2} = \frac{\overline{R}}{1.12838} = 0.8862 \cdot \overline{R}, \qquad \text{i.e. } K_1 = 0.8862 \text{ for 2 measurements.}$$

The **standard deviation PV** is determined from the range R_p of the 25 mean values of each part, i.e. g = 1 sample consisting of m = 25 mean values. The tabular value is d_2^* = 3.99377 (not included in [AIAG MSA], table on page 203):

$$PV = \frac{R_{1;25}}{d_2^*} = \frac{R_p}{3.99377} = 0.2504 \cdot R_p, \qquad \text{i.e. } K_3 = 0.2504 \text{ for 25 parts.}$$

In case of deviating numbers (measurements, parts) the values of the K-factors (K_1, K_3) must be adapted accordingly. Otherwise, the analysis according to ARM leads to incorrect results.

D.5 Approach in case of an insufficient number of measuring objects

The "reliability" of a statistical result is defined quantitatively by the confidence interval at a certain confidence level 1 – α, i.e. the width of the interval in which the measurement results are to be expected with a probability of 1 – α: the smaller the width the "more reliable" the statistical result.

The width of the confidence interval is determined by the number of degrees of freedom, i.e. mainly by the number of individual components (e. g. the measurement results) which are used to calculate the statistical result (e. g. the variance).

For capability studies of measuring equipment, the number of measurement results is determined by the test scenario, i.e. the number of parts n, the number of appraisers k and the number r of measurements per part and appraiser. The test scenario for a certain test is usually specified. Thus, the confidence interval for the variance is also specified.

The variance consists of components that are allocated to certain influence quantities and calculated using the analysis of variance (ANOVA, see Appendix D.2). The following table summarizes the degrees of freedom for four common test scenarios:

		Procedure 2 2 measurement series Test scenario n = 10 k = 3 r = 2	Procedure 2 3 measurement series Test scenario n = 10 k = 3 r = 3	Procedure 3 2 measurement series Test scenario n = 25 k = 1 r = 2	Procedure 3 3 measurement series Test scenario n = 25 k = 1 r = 3
Influence quantity	Degrees of freedom	Degrees of freedom	Degrees of freedom	Degrees of freedom	Degrees of freedom
Parts	n – 1	9	9	24	24
Appraisers	k – 1	2	2	0	0
Interaction	(n – 1)·(k – 1)	18	18	0	0
Measuring equipment	nk·(r – 1)	30	60	25	50
Total	nkr – 1	59	89	49	74

The influence quantity *"measuring equipment"* is of special importance for the procedures 2 and 3. If, in exceptional cases, less than the number of parts n specified for the respective test scenario is available, the basic approach includes changing the parameters r and/or k, so that the same number of degrees of freedom and thus the same confidence interval is reached for the influence quantity *"measuring equipment"* as for the specified test scenario. The number r of measurements is usually changed.

The influence quantity *"measuring equipment"* has $f = n \cdot k \cdot (r-1)$ degrees of freedom for n parts, k appraisers and r measurements per part and appraiser. If only $n' < n$ parts are available, the number of measurements per part and appraiser r' must be adapted, so that the resulting number of degrees of freedom $f' = n' \cdot k \cdot (r'-1)$ does not underrun the intended number of degrees of freedom f.

A pragmatic solution is to require a minimum value for the corresponding products.

According to [CDQ 0301] the following rules apply.

- Procedure 2: $n \cdot k \cdot r \geq 60$
- Procedure 3: $n \cdot k \geq 50$

In addition, the number of parts must not be less than n=5.

D.6 Procedure 1 – procedure 2 and 3: Inconsistent classification into capability categories

Classifications of measurement processes into the categories “capable”, “conditionally capable” or “not capable” according to procedure 1 are not fully consistent with classifications according to procedures 2 and 3. This may lead to (possibly technically unsubstantiated) problems when capability can be proven according to procedure 1 but not according to procedures 2 or 3.

Procedures 2 and 3 must ensure a reliable and consistent classification into capability categories also in a (theoretically possible) limit case of ideal measuring objects, i.e. for production parts from a (nearly) ideal production process without (significant) variation of characteristics. In this case, the measuring objects can be considered identical concerning their characteristics and it is statistically irrelevant for the measurement results whether each of n different measuring objects is measured once, or n/2 measuring objects are measured twice, or a single measuring object is measured n times. If these measuring objects are used for both a type-1 and a type-3 study, the observed variation s of the measured values is exclusively caused by the measuring equipment in either case. Thus, both procedures yield statistically identical results concerning the variation. Consequently, the analyses both should lead to the same classification of the measuring equipment into the same capability category. However, this is not the case.

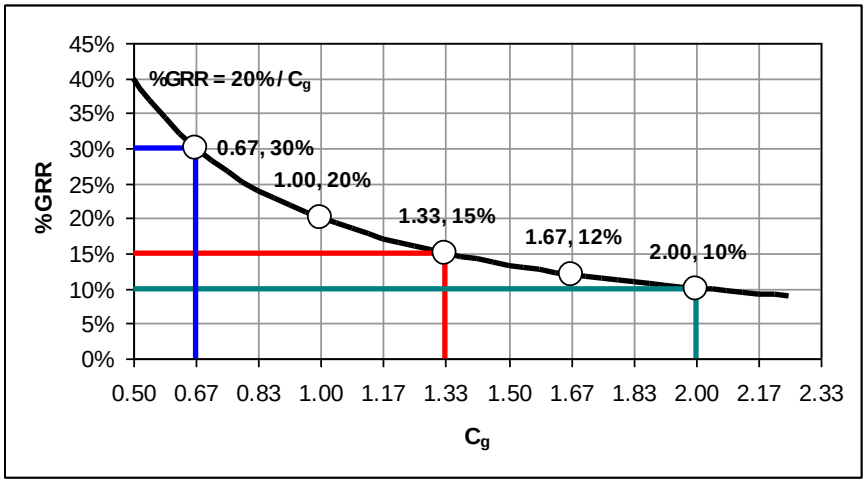
C_g and %GRR are defined according to

$$C_g = \frac{0{,}2 \cdot T}{6 \cdot s} \quad \text{and} \quad \%GRR = \frac{6 \cdot s}{T} \cdot 100\% \text{, respectively.}$$

Both equations solved for 6 s / T and equalized gives

$$\%GRR = \frac{20\%}{C_g}.$$

The diagram shows this relationship and the specified limits for C_g and %GRR:



This leads to the following allocations with partly contradicting capability classifications:

C_g		%GRR	
0.50	not capable	40%	not capable
0.67	**not capable**	30%	**conditionally capable** (upper limit)
0.80	conditionally capable (lower limit)	25%	conditionally capable
1.00	conditionally capable	20%	conditionally capable
1.33	**capable** (lower limit)	15%	**conditionally capable**
1.67	**capable**	12%	**conditionally capable**
2.00	capable	10%	capable (upper limit)
4.00	capable	5%	capable

There are no physical or technical but historical reasons for these contradictions since the procedures were developed independently of each other. Procedure 1 is based on company guidelines and [VDA 5], procedure 2 on [AIAG MSA], procedure 3 on company guidelines and (as a special case of procedure 2) on [AIAG MSA]. The classification of quantitative results into the capability categories “capable”, “conditionally capable” and “not capable” is based on experience and was done arbitrarily and without adapting the procedures to each other. Thus, there is no physically or technically justifiable method that can remove these contradictions. Either procedure 1 will conflict with numerous company guidelines (including RB-internal guidelines and [VDA 5]) or procedure 2 (and thus procedure 3) will conflict with (the recommendations of) the international guideline [AIAG MSA] used by numerous companies.

Up to now, there is no standardized specification for the calculation of C_g (and C_{gk}) that is obligatory for all companies. The equations mainly differ in the factors they include. However, consistency with [AIAG MSA] cannot be achieved using constant factors for mathematical reasons. Essentially, there are two options.

Option 1: Adaptation of the capability classification of C_g to the classification of %GRR

not capable:	***C_g < 0.67***	%GRR > 30%
conditionally capable:	***0.67 ≤ C_g < 2.00***	30% ≥ %GRR > 10%
capable:	***C_g ≥ 2.00***	%GRR ≤ 10%

This option conforms to [AIAG MSA], whose binding character – compared to company guidelines and [VDA 5] – is generally considered stronger. In practice, however, it leads to the majority of measuring systems being classified “conditionally capable” when using procedure 1.

Option 2: Adaptation of the capability classification of %GRR to the classification of C_g

not capable:	C_g < 0.80	***%GRR > 25%***
conditionally capable:	0.80 ≤ C_g < 1.33	***25% ≥ %GRR > 15%***
capable:	C_g ≥ 1.33	***%GRR ≤ 15%***

Regarding technical relevance, this alternative is supposed to be the better adaptation. However, it does not conform to [AIAG MSA]. In particular, it can be expected that increasing the limiting value for “capable” to %GRR = 15% is seen skeptically by the customers, while decreasing the limiting value for “not capable” to %GRR = 25% is seen skeptically by RB.

Note

If the application of these (or similar) options is considered, it will always have to be agreed upon with the customer.

E Amendments and Notes on Procedure 4 (Linearity)

E.1 Procedure according to AIAG MSA

Description of the procedure

- **Preparation:** g ≥ 5 serial parts are selected which adequately cover the operating range (measurement range) of the measuring equipment (e. g. equidistant arrangement). A reference value x_i is determined for each part by measurements with a sufficiently small measurement uncertainty.
- **Conducting the measurements:** Each of these g reference parts is measured at least 12 times (m ≥ 12) by the designated appraiser with the measuring equipment to be examined. The measured values ξ_{ik} (Greek letter "ksi") are documented. ξ_{ik} is the measured value no. k that was measured at reference part no. i.

Analysis

g·m pairs of values (x_i; ξ_{ik}) have to be analyzed where i = 1 ... g and k = 1 ... m. The mathematical representation in [AIAG MSA] was shown to be insufficiently clearly interpretable in some aspects. This problem is removed in the following reanalysis using a mathematically clear nomenclature and notation while full compliance with [AIAG MSA] is maintained. In particular, all sums are represented as double sums over the index i of the reference values and the index k of the measured values [3] (instead of single sums without indexes in [AIAG MSA]).

- Determine the measurement errors
$$y_{ik} = \xi_{ik} - x_i ,$$
i.e. the deviations (residuals) of each measured value ξ_{ik} from the respective reference value x_i
- Plot the measurement errors y_{ik} versus the reference values x_i

  NOTE: Usually the mean values of each group i (systematic measurement errors, bias) are also plotted:
$$\bar{y}_i = \frac{1}{m}\sum_{k=1}^{m} y_{ik} .$$
- Calculate the regression line $\hat{y}(x_i) = b + a \cdot x_i$

  Slope:
$$a = \frac{\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i \cdot y_{ik}) - \frac{1}{g \cdot m}\left(\sum_{i=1}^{g}\sum_{k=1}^{m} x_i\right)\cdot\left(\sum_{i=1}^{g}\sum_{k=1}^{m} y_{ik}\right)}{\sum_{i=1}^{g}\sum_{k=1}^{m} x_i^{\,2} - \frac{1}{g \cdot m}\left(\sum_{i=1}^{g}\sum_{k=1}^{m} x_i\right)^2}$$

  Intercept:
$$b = \frac{1}{g \cdot m}\left(\sum_{i=1}^{g}\sum_{k=1}^{m} y_{ik} - a \cdot \sum_{i=1}^{g}\sum_{k=1}^{m} x_i\right)$$
- Calculate the confidence limits for confidence level 1 - α

  Mean value of reference values:
$$\bar{x} = \frac{1}{g \cdot m}\sum_{i=1}^{g}\sum_{k=1}^{m} x_i = \frac{1}{g}\sum_{i=1}^{g} x_i$$

  Variation of measurement errors around regression line:
$$s = \sqrt{\frac{\sum_{i=1}^{g}\sum_{k=1}^{m} y_{ik}^2 - b \cdot \sum_{i=1}^{g}\sum_{k=1}^{m} y_{ik} - a \cdot \sum_{i=1}^{g}\sum_{k=1}^{m} x_i \cdot y_{ik}}{g \cdot m - 2}} = \sqrt{\frac{\sum_{i=1}^{g}\sum_{k=1}^{m} y_{ik} \cdot (y_{ik} - b - a \cdot x_i)}{g \cdot m - 2}}$$

  NOTE: Text books usually give this equation according to
$$s = \sqrt{\frac{\sum_{i=1}^{g}\sum_{k=1}^{m} (y_{ik} - \hat{y}(x_i))^2}{g \cdot m - 2}} = \sqrt{\frac{\sum_{i=1}^{g}\sum_{k=1}^{m} (y_{ik} - b - a \cdot x_i)^2}{g \cdot m - 2}} .$$

  It is not directly transparent that both representations are identical. However, the identity can be proven by substituting the formulas for slope a and intercept b and algebraically rearranging the equation.

[3] For sums whose argument is not dependent on the index of the sum, the following applies: $\sum_{k=1}^{m} x_i = m \cdot x_i$

Lower limit of confidence interval:

$$LCI(x_0) = b + a \cdot x_0 - t_{g \cdot m-2;\,1-\alpha/2} \cdot \sqrt{\frac{1}{g \cdot m} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i - \bar{x})^2}} \cdot s$$

Upper limit of confidence interval:

$$UCI(x_0) = b + a \cdot x_0 + t_{g \cdot m-2;\,1-\alpha/2} \cdot \sqrt{\frac{1}{g \cdot m} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i - \bar{x})^2}} \cdot s$$

NOTE 1: The confidence limits are usually calculated for the confidence level 95% (α = 5%). Deviations from this convention should be agreed upon with the customer.

*NOTE 2: The calculated confidence limits apply to the **mean value** of the expected measurement errors at an arbitrary point x_0. Differing equations have to be used to calculate the confidence limits for the corresponding **individual values** (here not relevant).*

- Plot the regression line and the confidence limits.
- If necessary, statistical t-tests for the significance of the slope and the intercept of the regression line must be applied [AIAG MSA].

  Slope:

  $$t_a = \frac{|a|}{s} \cdot \sqrt{\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i - \bar{x})^2}$$

  The slope a is statistically insignificant (i.e. negligible) if $t_a \le t_{g \cdot m-2;\,1-\alpha/2}$ is fulfilled. [4]

  Intercept:

  $$t_b = \frac{|b|}{s} \cdot \sqrt{\frac{\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i - \bar{x})^2}{\frac{1}{g \cdot m}\sum_{i=1}^{g}\sum_{k=1}^{m}(x_i - \bar{x})^2 + \bar{x}^2}}$$

  The intercept b is statistically insignificant if $t_b \le t_{g \cdot m-2;\,1-\alpha/2}$ is fulfilled. [4]

- If necessary, a statistical F-test for the compatibility of the linear model with the measured data must be applied (recommendation of [AIAG MSA] which refers to the literature at this point).

  Linear model:

  $$F_{LM} = \frac{\frac{1}{g-2} \cdot \sum_{i=1}^{g}\sum_{k=1}^{m}(\bar{y}_i - b - a \cdot x_i)^2}{\frac{1}{g \cdot (m-1)} \cdot \sum_{i=1}^{g}\sum_{k=1}^{m}(y_{ik} - \bar{y}_i)^2}$$

  Deviations from linear behavior are statistically insignificant if $F_{LM} \le F_{g-2;\,g \cdot (m-1);\,1-\alpha}$ is fulfilled. [4]

Due to the complexity of the equations the analysis can only be done with a computer in practice.

Capability criterion

The zero line of the deviations y_{ik} must be completely within the confidence limits. This is equivalent to the requirement that the slope and the intercept of the regression line are not significantly different from zero (t-test [AIAG MSA]).

NOTE: The applicability of the capability criterion requires measured data which comply with the linear model.

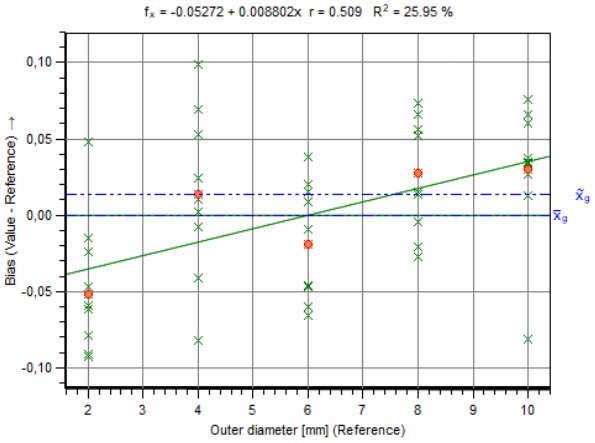
[4] The values of the quantiles $t_{g \cdot m-2;\,1-\alpha/2}$ and $F_{g-2;\,g \cdot (m-1);\,1-\alpha}$ can be found in tables (e. g. [Booklet 3]) or determined
using a suitable software (e. g. MS EXCEL: **=TINV**(*α; gm-2*) or **=FINV**(*α; g-2; g(m-1)*)

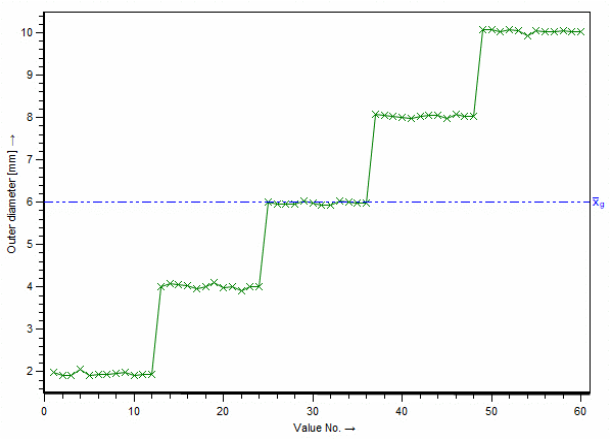
BOSCH	Measurement System Analysis Procedure 4 (Linearity)	Sheet 1 / 1

Area	:	QMM	Operation	:		Characteristic	:	Outer diameter
Group/Dptm.	:	QMM7	Machine	:	Z-Mike 1220 Gold	Char. No.	:	1
Workshop/sect.	:	W780	Machine No.	:	1205292	Nominal value	:	6.000
Product	:		Test station	:		Lower allowance	:	
Part	:	Reference	Gage	:	JMO0014W008	Upper allowance	:	
Article number	:		Gage No.	:	67027855300001	Tolerance	:	
Change status	:		Gage Manuf.	:	Z-Mike	Unit	:	mm
			Resolution	:	0.0001			

Comment	:	
Standard:	Standard No.:	Standard/Ref. value:

i	$x_{g,Ref}$	$x_{A;1}$	$x_{A;2}$	$x_{A;3}$	$x_{A;4}$	$x_{A;5}$	$x_{A;6}$	$x_{A;7}$	$x_{A;8}$	$x_{A;9}$	$x_{A;10}$	$x_{A;11}$	$x_{A;12}$	$x_{g,j}$	s_j
1	2.0010	1.9770	1.9220	1.9080	2.0490	1.9100	1.9420	1.9390	1.9540	1.9860	1.9080	1.9495	1.9495	1.9495	0.04038
2	4.0030	4.0170	4.0720	4.0560	4.0270	3.9620	4.0130	4.1020	3.9950	4.0050	3.9210	4.0170	4.0170	4.0170	0.04744
3	5.9990	6.0080	5.9520	5.9530	5.9520	6.0190	5.9900	5.9390	5.9330	6.0370	6.0140	5.9797	5.9797	5.9797	0,03437
4	8.0010	8.0740	8.0570	8.0150	7.9970	7.9740	8.0160	8.0530	8.0530	7.9800	8.0670	8.0286	8.0286	8.0286	0.03333
5	10.002	10.068	10.0620	10.032	10.078	10.037	9.921	10.036	10.015	10.029	10.039	10.0317	10.0317	10.0317	0.03934





Equation of the straight line:	f(x) = -0.05272 + 0.008802 x	r = 0.509	R^2 = 25.953%

Test of the slope		
α = 5 %		H_0: The slope of the regression line is equal to 0 H_1: The slope of the regression line is not equal to 0
Upper critical value (α = 5 %)	= 2.00	Test result is significant (α ≤ 0.1 %)
Upper critical value (α = 1 %)	= 2.66	4.50877 ***
Upper critical value (α = 0,1 %)	= 3.47	
H1***: The slope of the regression line is not equal to 0		

Test of the intercept		
α = 5 %		H_0: The intercept of the regression line is equal to 0 H_1: The intercept of the regression line is not equal to 0
Upper critical value (α = 5 %)	= 2.00	Test result is significant (α ≤ 0.1 %)
Upper critical value (α = 1 %)	= 2.66	4.07076 ***
Upper critical value (α = 0,1 %)	= 3.47	
H1***: The intercept of the regression line is not equal to 0		

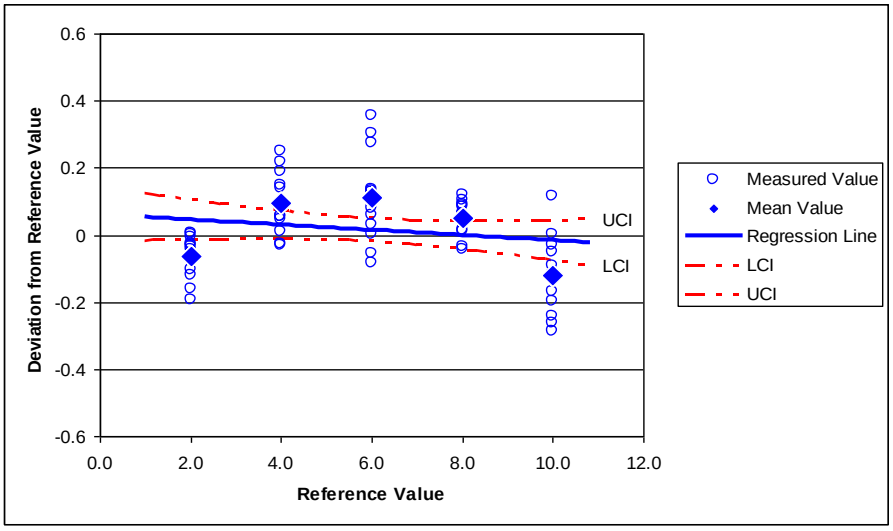
Repeatability	EV: 0.0033129 ≤ 0.042772 ≤ 0.048304	%EV = ---
		Bi = 0.051500
Requirements are not fulfilled (min, t-T)		
Bosch 2018 — Procedure, Linearity		

2020-04-06 - SOCOS

Examples and notes on the procedure according to AIAG-MSA

Practical experience with this procedure has shown that even a strongly ***non-linear*** (e. g. parabolic) ***behavior*** of measuring equipment is not always reliably detected.

Reference values x_i		Measured values ξ_{ik}											
		1	2	3	4	5	6	7	8	9	10	11	12
1	**2.001**	1.960	1.928	2.006	1.809	1.971	1.996	1.896	2.004	1.965	1.841	1.973	1.883
2	**4.003**	3.971	4.255	4.057	4.221	4.082	4.153	4.056	3.977	4.012	4.144	4.049	4.193
3	**5.999**	6.058	5.914	6.079	6.003	5.943	6.137	6.127	6.132	6.029	6.356	6.272	6.304
4	**8.001**	8.054	8.122	7.958	8.103	8.085	8.017	8.089	8.092	8.064	8.012	7.967	8.064
5	**10.002**	9.953	9.715	10.004	10.116	9.886	9.759	9.911	9.973	9.885	9.741	9.805	9.832



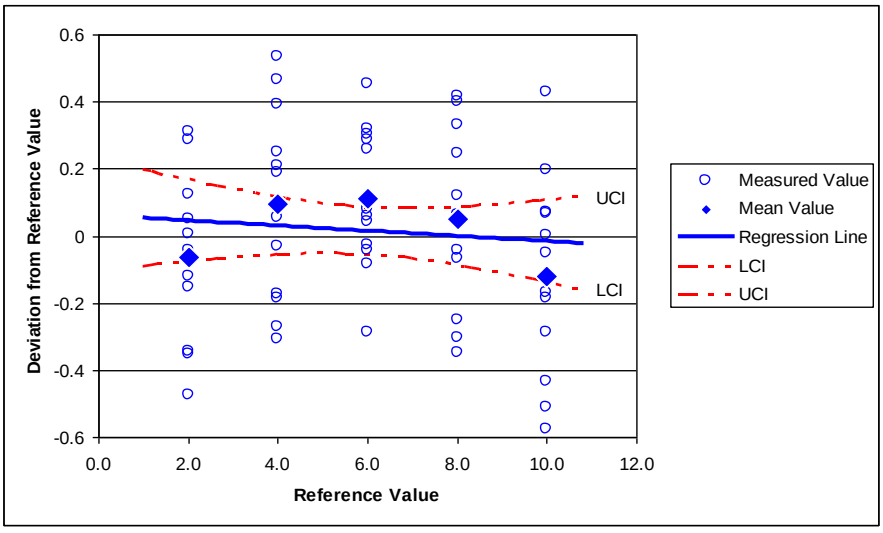
Results of the statistical tests for confidence level 95%:

Slope:	t_a =	**1.271**	≤	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**2.002**	Deviation from 0 is **not significant**
Intercept:	t_b =	**1.519**	≤	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**2.002**	Deviation from 0 is **not significant**
Linear model:	F_{LM} =	**16.055**	>	$F_{g-2;\ g\cdot(m-1);\ 1-\alpha}$ =	**2.773**	Deviation is **significant**

Only the test for compatibility of the measured values with the linear model (which is merely recommended by [AIAG MSA]) shows a significant incompatibility.

Increasing variation of the measured values unjustifiably favors meeting the capability criterion (particularly if the mean values do not change) which is due to the increasing width of the confidence interval.

Reference values x_i		Measured values ξ_{ik}											
		1	2	3	4	5	6	7	8	9	10	11	12
1	**2.001**	1.960	1.928	2.006	2.124	2.286	2.311	2.054	1.847	1.650	1.526	1.658	1.883
2	**4.003**	3.971	4.255	4.057	4.536	4.397	4.468	4.214	3.820	3.697	3.829	3.734	4.193
3	**5.999**	6.058	5.914	6.079	6.318	6.258	6.452	6.285	5.975	5.714	6.041	5.957	6.304
4	**8.001**	8.054	8.122	7.958	8.418	8.400	8.332	8.247	7.935	7.749	7.697	7.652	8.064
5	**10.002**	9.953	9.715	10.004	10.431	10.201	10.074	10.069	9.816	9.570	9.426	9.490	9.832



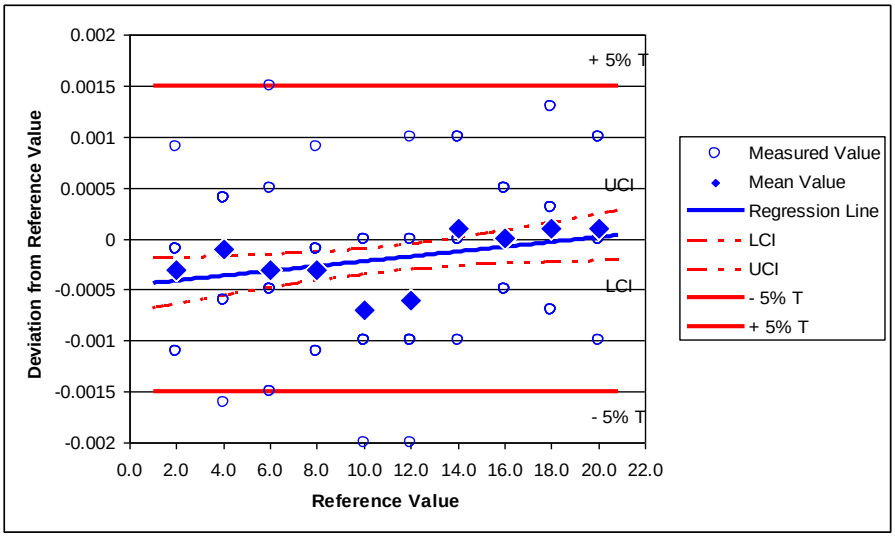
Results of the statistical tests for confidence level 95%:

Slope:	t_a =	**0.618**	≤	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**2.002**	Deviation from 0 is **not significant**
Intercept:	t_b =	**0.739**	≤	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**2.002**	Deviation from 0 is **not significant**
Linear model:	F_{LM} =	**2.275**	≤	$F_{g-2;\ g\cdot(m-1);\ 1-\alpha}$ =	**2.773**	Deviation is **not significant**

In contrast to the previous example, the incompatibility of the linear model with the measured data is not identified so that the measuring equipment would have to be classified as *"capable"* concerning its linearity.

The procedure can also be ***too sensitive*** so that measuring equipment with technically excellent characteristics concerning their linearity is statistically classified as *"not capable"*.

Reference values x_i		Measured values ξ_{ik}											
		1	2	3	4	5	6	7	8	9	10	11	12
1	**2.0011**	2.001	2.000	2.001	2.000	2.001	2.001	2.000	2.001	2.002	2.001	2.001	2.001
2	**4.0006**	4.001	4.000	3.999	4.000	4.001	4.001	4.001	4.000	4.001	4.001	4.001	4.001
3	**6.0005**	6.000	6.001	6.000	6.000	5.999	6.001	6.002	6.000	5.999	6.000	6.000	6.000
4	**8.0011**	8.001	8.001	8.002	8.001	8.000	8.001	8.001	8.000	8.001	8.000	8.001	8.001
5	**10.0010**	10.001	10.001	10.000	9.999	10.000	10.000	10.001	10.000	10.001	10.000	10.000	10.000
6	**12.0010**	12.001	12.000	12.000	12.002	12.000	12.001	12.001	11.999	12.000	12.000	12.000	12.000
7	**14.0010**	14.002	14.001	14.002	14.001	14.000	14.001	14.001	14.002	14.001	14.000	14.001	14.001
8	**16.0005**	16.000	16.000	16.001	16.000	16.001	16.001	16.000	16.000	16.001	16.001	16.001	16.001
9	**18.0007**	18.001	18.002	18.000	18.001	18.000	18.001	18.002	18.000	18.001	18.000	18.001	18.001
10	**20.0000**	20.001	20.000	20.000	19.999	20.000	19.999	20.000	20.001	20.000	20.001	20.000	20.000

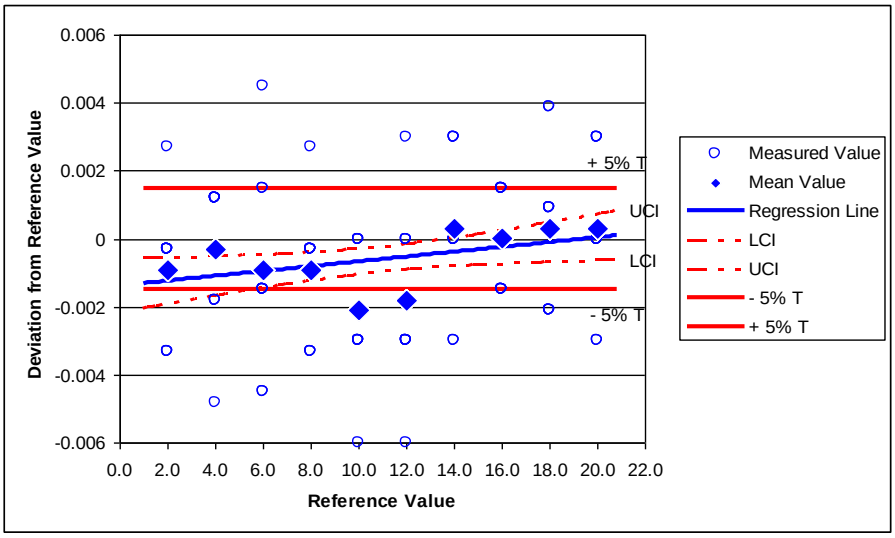


Results of the statistical tests for confidence level 95%:

Slope:	t_a =	**2.186**	>	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**1.980**	Deviation from 0 is **significant**
Intercept:	t_b =	**3.429**	>	$t_{g\cdot m-2;\ 1-\alpha/2}$ =	**1.980**	Deviation from 0 is **significant**
Linear model:	F_{LM} =	**1.991**	≤	$F_{g-2;\ g\cdot(m-1);\ 1-\alpha}$ =	**2.024**	Deviation is **not significant**

These results are surprising from a technical point of view. Background is that the statistical tests assess only on a relative basis, i.e. independent of the absolute values of the measurement errors y_{ik} and thus independent of their technical relevance.

NOTE: Substituting $\varepsilon \cdot y_{ik}$ for y_{ik} in the above equations where the factor ε represents an arbitrary positive number, results in the slope $\varepsilon \cdot a$, the intercept $\varepsilon \cdot b$, the variation $\varepsilon \cdot s$ and the confidence limits $\varepsilon \cdot LCI$ und $\varepsilon \cdot UCI$, i.e. these quantities also decrease ($0 < \varepsilon < 1$) or increase ($\varepsilon > 1$) by the factor ε. Thus, the diagrams for $\varepsilon \cdot y_{ik}$ und y_{ik} appear identical if a y-axis is used to plot the results for $\varepsilon \cdot y_{ik}$ whose scale is stretched or compressed by the factor $1/\varepsilon$. Example for $\varepsilon = 3$:



In contrast, the test statistics t_a, t_b and F_{LM} for the statistical tests remain unchanged, i.e. they are independent of ε and always yield the same results.

For the same reason, the procedure can be ***too insensitive*** so that measuring equipment with technically insufficient characteristics concerning their linearity are statistically classified as *"capable"* (e. g. in case of high variation, see page 61, 2nd example).

E.2 Multiple use of procedure 1: Additional considerations

The linearity study according to [AIAG MSA] does not include a criterion for the assessment of the technical relevance of statistical results. However, the alternative analysis of the measurements at each reference part i according to ***procedure 1*** includes the tolerance T of the characteristic to be measured in the analysis. The parameters C_g and C_{gk} comprise additional tolerance-related (and thus technically relevant) criteria for the variations s_i of the measured values and the systematic measurement errors $\bar{y}_i$:

$$\%SDev_i = \frac{s_i}{T} \cdot 100\% \le 2.5\%,$$

$$\%Bias_i = \frac{|\bar{y}_i|}{T} \cdot 100\% = \frac{|\bar{\xi}_i - x_i|}{T} \cdot 100\% \le \left(0.1 - 4 \cdot \frac{s_i}{T}\right) \cdot 100\% = 10\% - 4 \cdot \%SDev_i .$$

NOTE: The first criterion results from solving the equation for C_g for s_i / T and inserting the condition $C_g \ge 1.33 = 4/3$; the second criterion results from solving the equation for C_{gk} for $|\bar{\xi}_i - x_i|$ / T and inserting the condition $C_{gk} \ge 1.33 = 4/3$; see chapter 4.1 for the equations.

Using the second criterion with the limiting values $\%SDev_i = 0$ and $\%SDev_i = 2.5\%$ shows that systematic measurement errors in the range

$$0 \le \%Bias_i \le 10\%$$

are acceptable under the condition that the corresponding variations of the measured values are in the range

$$0 \le \%SDev_i \le 2.5\% - \frac{\%Bias_i}{4} = \%SDev_{max\,i} .$$

Thus, 10% bias is only acceptable for 0% variation and 2.5% variation only for 0% bias.

Example: From the measured values in the table on page 62, the parameters C_g and C_{gk} are estimated using T = 0.003 mm; %Bias and %SDev are determined; results that violate the criteria are highlighted:

Reference value		Mean of measured values	Standard deviation of measured values	C_g	C_{gk}	%Bias	%SDev	$\%SDev_{max}$	%SDev > $\%SDev_{max}$
1	**2.0011**	2.0008	0.00057	1.75	1.57	1.00%	1.91%	2.25%	
2	**4.0006**	4.0005	0.00064	1.56	1.51	0.33%	2.13%	2.42%	
3	**6.0005**	6.0002	0.00083	***1.20***	***1.08***	1.00%	***2.77%***	2.25%	**X**
4	**8.0011**	8.0008	0.00057	1.75	1.57	1.00%	1.91%	2.25%	
5	**10.0010**	10.0003	0.00061	1.64	***1.26***	2.33%	***2.04%***	1.92%	**X**
6	**12.0010**	12.0004	0.00076	***1.31***	***1.05***	2.00%	***2.54%***	2.00%	**X**
7	**14.0010**	14.0011	0.00067	1.50	1.45	0.33%	2.22%	2.42%	
8	**16.0005**	16.0005	0.00048	2.10	2.10	0.00%	1.59%	2.50%	
9	**18.0007**	18.0008	0.00071	1.40	1.35	0.33%	2.38%	2.42%	
10	**20.0000**	20.0001	0.00067	1.50	1.45	0.33%	2.22%	2.42%	

*NOTE: It is important to note that **definitive** analyses according to procedure 1 require m ≥ 25 measurements per reference part i.*

Thus, the multiple use of procedure 1 is similarly meaningful like the procedure according to [AIAG MSA]. In contrast to the linearity study according to [AIAG MSA], the technical relevance of the result is generally ensured since the tolerance T is included in the analysis.

Procedure 1 and previous alternative analysis according to QA information 02/2004

According to the column %Bias of the table above, the systematic measurement errors are at most 2.33% of the tolerance of the characteristic to be measured. Thus, they fulfill the criterion for unconditional capability in terms of the old QA information 02/2004, i.e. all systematic measurement errors are within 5% of the tolerance (see also diagram on top of page 62). In contrast to this, the capability according to procedure 1 is not always verified.

This discrepancy is due to the missing analysis of the variation of the measured values when using the procedure according to the QA information 02/2004. Obviously the simple proof is insufficient that the systematic measurement errors are completely within a given range. In order to achieve consistency, the analysis must be extended by the additional criterion for the variation of the measured values. This corresponds directly to the multiple use of procedure 1.

NOTE: The classification "capable" (0% ≤ $\%Bias_i$ ≤ 5%) and "conditionally capable" (5% < $\%Bias_i$ ≤ 10%) by means of the systematic measurement error (as intended by the QA information 02/2004) is not reasonable together with the additional criterion for the measurement variation (0% ≤ $\%SDev_i$ ≤ 2.50% – $\%Bias_i$ /4). Satisfying both criteria merely corresponds to satisfying the criterion $C_{gk} \ge 1.33$, i.e. arbitrary $C_{gk} \ge 1.33$ are allowed.

Contradicting capability results concerning the same characteristic of measuring equipment are not acceptable. Thus, the QA information 02/2004 must not be applied any longer.

F Procedure 5: Factors for the Calculation of Control Limits

To calculate the control limits for stability charts, the factors u_p, B'_{Eun}, B'_{Eob} and E'_E are needed depending on the type of chart (see chapter 4.5.1).

The factors u_p are dependent on the confidence level 1 - α as well as on the type of limit – one-sided or two-sided. They are calculated as quantiles of the standardized normal distribution. The following table contains u_p for the calculation of one- and two-sided control limits of $\bar{x}$-charts for the confidence levels 99.73% (α = 0.0027) and 99% (α = 0.01).

α = 0.0027 (two-sided)	**α = 0.01 (two-sided)**	**α = 0.0027 (one-sided)**	**α = 0.01 (one-sided)**
$u_p = -u_{\alpha/2} = u_{1-(\alpha/2)}$		$u_p = -u_{\alpha} = u_{1-\alpha}$	
3.000	2.576	2.782	2.326

The factors B'_{Eun}, B'_{Eob} and E'_E additionally depend on the sample size n. They are calculated from the quantiles of the χ^2-distribution (B'_{Eun}, B'_{Eob}) or the standardized normal distribution (E'_E).

Two-sided control limits:

$$B'_{Eun} = \sqrt{\frac{\chi^2_{n-1;\,\alpha/2}}{n-1}} \qquad B'_{Eob} = \sqrt{\frac{\chi^2_{n-1;\,1-(\alpha/2)}}{n-1}} \qquad E'_E = u_{\sqrt[n]{1-(\alpha/2)}}$$

One-sided control limits:

$$B'_{Eun} = \sqrt{\frac{\chi^2_{n-1;\,\alpha}}{n-1}} \qquad B'_{Eob} = \sqrt{\frac{\chi^2_{n-1;\,1-\alpha}}{n-1}} \qquad E'_E = u_{\sqrt[n]{1-\alpha}}$$

The following table contains the factors B'_{Eun}, B'_{Eob} and E'_E for the calculation of one- and two-sided control limits for s-charts or individual value charts for several sample sizes n and the confidence levels 99.73% (α = 0.0027) and 99% (α = 0.01).

n	**α = 0.0027 (two-sided)**			**α = 0.01 (two-sided)**			**α = 0.0027 (one-sided)**			**α = 0.01 (one-sided)**		
	B'_{Eun}	B'_{Eob}	E'_E	B'_{Eun}	B'_{Eob}	E'_E	B'_{Eun}	B'_{Eob}	E'_E	B'_{Eun}	B'_{Eob}	E'_E
2	0.002	3.205	3.205	0.006	2.807	2.807	0.003	3.000	3.000	0.013	2.576	2.575
3	0.037	2.571	3.320	0.071	2.302	2.935	0.052	2.432	3.121	0.100	2.146	2.712
4	0.100	2.283	3.399	0.155	2.069	3.023	0.126	2.172	3.205	0.196	1.945	2.806
5	0.163	2.110	3.460	0.227	1.927	3.090	0.194	2.016	3.269	0.273	1.822	2.877
6	0.218	1.991	3.509	0.287	1.830	3.143	0.252	1.908	3.320	0.333	1.737	2.934
7	0.266	1.903	3.550	0.336	1.758	3.188	0.300	1.829	3.363	0.381	1.674	2.981
8	0.306	1.835	3.585	0.376	1.702	3.227	0.341	1.767	3.399	0.421	1.625	3.022
9	0.341	1.780	3.615	0.410	1.657	3.260	0.376	1.717	3.431	0.454	1.585	3.057
10	0.371	1.735	3.642	0.439	1.619	3.290	0.405	1.675	3.460	0.482	1.552	3.089
11	0.398	1.697	3.667	0.464	1.587	3.317	0.431	1.640	3.485	0.506	1.523	3.117
12	0.422	1.664	3.689	0.486	1.560	3.341	0.454	1.610	3.509	0.527	1.499	3.143
13	0.443	1.635	3.709	0.506	1.536	3.363	0.475	1.584	3.530	0.545	1.478	3.166
14	0.461	1.609	3.728	0.524	1.515	3.383	0.493	1.561	3.549	0.562	1.459	3.187
15	0.479	1.587	3.745	0.539	1.496	3.402	0.509	1.540	3.567	0.577	1.443	3.207
20	0.545	1.502	3.817	0.600	1.425	3.480	0.573	1.462	3.642	0.634	1.380	3.289
25	0.591	1.446	3.872	0.642	1.378	3.539	0.617	1.411	3.699	0.673	1.338	3.351

Factors for other values of α and n are calculated using the equations above. The quantiles $\chi^2_{f;p}$ and u_p can be found in tables or calculated using e. g. the EXCEL worksheet function *CHIINV(1-p;f)* or *STANDNORMINV(p)*. The degrees of freedom f and the probabilities p have to be included according to the equations above.

NOTE: The quantile $\chi^2_{f;p}$ is calculated using the EXCEL worksheet function according to CHIINV(1-p;f), i.e. 1 – p always has to be used instead of p; example: $\chi^2_{n-1;\,\alpha/2} = \text{CHIINV}(1-\alpha/2;\, n-1)$.

G Parameter “Kappa” (κ)

The degree of agreement of ratings is assessed quantitatively by means of the parameter “kappa” (κ) [ISO 14468].

The basic element for the assessment of agreement of all ratings is the rating pair that is made of two individual ratings. Depending on the selection, combination and analysis of these paired ratings, *Cohen's kappa* and *Fleiss' kappa* are distinguished.

- ***Cohen's kappa*** was developed to quantitatively assess the degree of agreement of rating results from two trials (two appraisers with one test run each or one appraiser with two test runs). [AIAG MSA] applies this approach to multiple raters each with multiple trials by forming and analyzing defined pairings. Thus, the question is which pairings are significant for the result.
- ***Fleiss' kappa*** is a systematic and consequent generalization since all theoretically possible pair-wise combinations of individual ratings are taken into account.

  NOTE: The present booklet only covers Fleiss' kappa. Complementing documentation on the topics “Cross-table method” and “Analysis according to [AIAG MSA]” is available at C/QMM and on the C/QMM intranet pages.

The ratings as a whole usually contain a certain number of agreements that are caused randomly – i.e. they are not based on objective decisions. It is an important characteristic of the parameter “kappa” that the number of random agreements is estimated and eliminated so that only the objective non-random agreement is assessed.

NOTE: One would get only random agreements, for example, if the appraiser would have to make decisions blindfolded in visual inspections.

G.1 Mathematical background

NOTE: This chapter mainly addresses readers who need more detailed information concerning the determination of the parameter kappa (κ).

Nomenclature and definitions

N_O	Number of test **o**bjects
N_A ; n_A *)	Number of **a**ppraisers (raters)
N_T ; n_T *)	Number of **t**rials (test runs)
N_R	Number *) of **r**atings per test object
$N = N_O \cdot N_R$	Total number *) of ratings
N_C	Number rating **c**ategories
n_{ik}	Number*) of allocations of test object i to rating category k; i = 1, ..., N_O; k = 1, ..., N_C
$\sum_{k=1}^{N_C} n_{ik} = N_R$	Total number*) of allocations of test object i to all N_C rating categories; i is arbitrary (number *) of ratings per test object)

*)*Number that is included in the analysis depending on the respective criterion to be analyzed.*

For every test object i (i = 1, ..., N_O), N_A appraisers provide a total of $N_A \cdot N_T$ ratings in N_T trials. Additionally, a reference value is allocated to each test object i .

NOTE: During analysis reference values are treated like ratings of one appraiser with one trial.

For the analyses based on different criteria (e. g. repeatability, reproducibility) different suitable subsets $n_A \le N_A$ and $n_T \le N_T$ are used, i.e. ratings of certain appraisers and trials and, if necessary, reference ratings. Thus, depending on the respective criterion $n_A \cdot n_T = N_R \le N_A \cdot N_T$ or in case of inclusion of the reference ratings $n_A \cdot n_T + 1 = N_R \le N_A \cdot N_T + 1$ ratings are analyzed for each test object i. n_{ik} of these N_R ratings allocate test object i to category k. For each test object i a total of N_C categories is available (k = 1, ..., N_C). The number of allocations n_{ik} of test object i to category k is in the range $0 \le n_{ik} \le N_R$.

Measure for observed agreement

[Fleiss] defines the rating pair formed out of two arbitrary individual ratings as the basic element for the assessment of the agreement of all N_R ratings for an individual test object i.

NOTE: Statistically, it is assumed that every individual rating is made by a randomly selected appraiser, i.e. the statistical independence of the rating results is presumed. This includes amongst others that it is impossible to identify a certain appraiser or trial based on the order of the combined individual ratings.

The first of these N_R ratings can be combined with the remaining (N_R-1) ratings to get (N_R-1) rating pairs. In the same way the second of these N_R ratings can be combined with the remaining (N_R-1) ratings to get another (N_R-1) rating pairs so that there is a total of 2•(N_R-1) rating pairs. All N_R ratings for an individual test object i can be combined to get

$$N_R \cdot (N_R - 1)$$

rating pairs.

NOTE: Combinations of different elements (e. g. "ab", "xy") and combinations of the same elements in different order (e. g. "xy", "yx") both are treated as independent combinations.

For all N_O test objects there is a total of

$$N_O \cdot N_R \cdot (N_R - 1)$$

possible rating pairs.

NOTE: Combinations of elements of different test objects are not considered.

In order to assess the agreement, only those rating pairs consisting of *consistent* individual ratings are used. n_{ik} represents the number of allocations (ratings) of a certain test object i to a certain category k. As explained earlier, n_{ik} ratings can be combined to get n_{ik}•(n_{ik}-1) rating pairs. Consequently, in total over all N_C categories there are

$$\sum_{k=1}^{N_C} n_{ik}(n_{ik} - 1)$$

consistent rating pairs which are allocated to an individual test object i. For all N_O test objects there is a total of

$$\sum_{i=1}^{N_O} \sum_{k=1}^{N_C} n_{ik}(n_{ik} - 1)$$

consistent rating pairs.

According to [Fleiss] the portion of actually **obs**erved *consistent* rating pairs out of all *possible* rating pairs is defined as a measure for the degree of agreement of N_O•N_R individual ratings which are available in total for N_O test objects:

$$P_{Obs} = \frac{\sum_{i=1}^{N_O} \sum_{k=1}^{N_C} n_{ik} \cdot (n_{ik} - 1)}{N_O \cdot N_R \cdot (N_R - 1)} .$$

Measure for random agreement

In case of random results, the test objects are allocated randomly to the N_C categories (e. g. if appraisers had to decide blindfolded in visual inspections).

If complete randomness is assumed, the N_O•N_R individual ratings are seen as N_O•N_R equivalent random experiments for the determination of relative frequencies of the random occurrances of the individual categories k. "*Equivalent*" means that it is irrelevant which appraiser allocates a rating to which test object in which trial. Only the total number of allocations to each category k is important. The relative frequencies determined this way are used as estimates for the probabilities p_k, with which the individual categories k can be expected in a random experiment:

$$\sum_{i=1}^{N_O} \frac{n_{ik}}{N_O \cdot N_R} = p_k .$$

According to probability theory, the probability for a test object being randomly allocated to category k in a first trial and being randomly allocated to category k' in a second trial is calculated as the product p_k•$p_{k'}$ of the individual probabilities p_k and $p_{k'}$.

For Fleiss' kappa, only rating pairs of *consistent* individual ratings are relevant, i.e. k = k':

$$p_k \cdot p_k = p_k^2 .$$

Accordingly, the following applies for the portion of pair-wise agreements, which has to be randomly **exp**ected in total over all N_C categories:

$$P_{Exp} = \sum_{k=1}^{N_C} p_k^2 = \sum_{k=1}^{N_C} \left(\sum_{i=1}^{N_O} \frac{n_{ik}}{N_O \cdot N_R}\right)^2 .$$

Parameter κ (kappa)

The parameter κ (kappa) can be interpreted as the portion of observed non-random agreements related to the portion of possible non-random agreements. The portion of observed *non-random* agreements is determined according to $P_{Obs} - P_{Exp}$, the portion of possible, *non-random* agreements according to $1 - P_{Exp}$. Thus the parameter κ (kappa) is calculated according to

$$\kappa = \frac{P_{Obs} - P_{Exp}}{1 - P_{Exp}} = \frac{\text{Observed non-random agreements}}{\text{Possible non-random agreements}}$$

Further background information for the parameter κ (kappa) can be found in the literature [Fleiss-2].

G.2 Manual analysis using data from AIAG MSA as an example

The analysis is explained using the test scenario and test data from [AIAG MSA] as an example.

Scenario

N_A = 3 appraisers (named A, B and C), N_T = 3 trials (no. 1 – 3), N_O = 50 reference parts as test objects (no. 1 – 50), N_C = 2 rating categories (0 - not OK, 1 - OK).

The test results (ratings) were documented in a table (see columns A-1 to C-3 of the evaluation diagrams on pages 69 and 70). The reference ratings (i.e. the "correct" ratings) are also included.

Parameters to be determined

To determine the different parameters kappa (κ) (see page 29), the test results (ratings) and the discrete reference values (reference ratings) have to be included according to the following table:

Test (Comparison)	Columns to be included in analysis:										Intermediate results in column group:												Result
	Reference	Test results (ratings)									A x A	B x B	C x C	A x B x C	A-1 x Ref	A-2 x Ref	A-3 x Ref	B-1 x Ref	B-2 x Ref	B-3 x Ref	C-1 x Ref	C-2 x Ref	C-3 x Ref	
		A-1	A-2	A-3	B-1	B-2	B-3	C-1	C-2	C-3														
Within appraisers: Agreement of ratings of one appraiser		X	X	X							X													$\kappa_{A \times A}$
					X	X	X					X												$\kappa_{B \times B}$
								X	X	X			X											$\kappa_{C \times C}$
Between appraisers: Agreement of ratings of all appraisers		X	X	X	X	X	X	X	X	X				X										$\kappa_{A \times B \times C}$
Each individual appraiser versus reference: Agreement of ratings of one appraiser with reference	X	X													X									$\kappa_{A\text{-}1 \times Ref}$
	X		X													X								$\kappa_{A\text{-}2 \times Ref}$
	X			X													X							$\kappa_{A\text{-}3 \times Ref}$
	X	X	X	X											X									$\kappa_{A \times Ref}$ = Mean $\kappa_{A\text{-}1 \times Ref}$, $\kappa_{A\text{-}2 \times Ref}$, $\kappa_{A\text{-}3 \times Ref}$
	X				X													X						$\kappa_{B\text{-}1 \times Ref}$
	X					X													X					$\kappa_{B\text{-}2 \times Ref}$
	X						X													X				$\kappa_{B\text{-}3 \times Ref}$
	X				X	X	X												X					$\kappa_{B \times Ref}$ = Mean $\kappa_{B\text{-}1 \times Ref}$, $\kappa_{B\text{-}2 \times Ref}$, $\kappa_{B\text{-}3 \times Ref}$
	X							X													X			$\kappa_{C\text{-}1 \times Ref}$
	X								X													X		$\kappa_{C\text{-}2 \times Ref}$
	X									X													X	$\kappa_{C\text{-}3 \times Ref}$
	X							X	X	X												X		$\kappa_{C \times Ref}$ = Mean $\kappa_{C\text{-}1 \times Ref}$, $\kappa_{C\text{-}2 \times Ref}$, $\kappa_{C\text{-}3 \times Ref}$
All appraisers together versus reference: Agreement of ratings of all appraisers with reference	X	X	X	X	X	X	X	X	X	X					X									$\kappa_{A \times B \times C \times Ref}$ = Mean $\kappa_{A \times Ref}$, $\kappa_{B \times Ref}$, $\kappa_{C \times Ref}$

Procedure using partial analysis "between appraisers" (κ_{AxBxC}) as an example

Calculation step	**Equation** (see Appendix G.1).	**Result**
1. *0*-ratings in the columns *A-1* to *C-3* are counted line-by-line, the results are entered in the corresponding line of the column ***n_{ik}, k=1***	n_{i1}	
2. *1*-ratings in the columns *A-1* to *C-3* are counted line-by-line, the results are entered in the corresponding line of the column ***n_{ik}, k=2***	n_{i2}	
3. Column ***n_{ik}, k=1*** is summed up	$\sum_{i=1}^{N_O=50} n_{i1} = n_1$	148
4. Column ***n_{ik}, k=2*** is summed up	$\sum_{i=1}^{N_O=50} n_{i2} = n_2$	302
5. The results of steps 3 and 4 are added	$\sum_{k=1}^{N_C=2} n_k = n_1 + n_2 = N$	450
6. The result from step 3 is divided by the result from step 5	$\frac{n_1}{N} = p_1$	0.3289
7. The result from step 4 is divided by the result from step 5	$\frac{n_2}{N} = p_2$	0.6711
8. The result from step 6 is multiplied by itself	$p_1 \cdot p_1 = {p_1}^2$	0.1082
9. The result from step 7 is multiplied by itself	$p_2 \cdot p_2 = {p_2}^2$	0.4504
10. The results of steps 8 and 9 are added	$\sum_{k=1}^{N_C=2} {p_k}^2 = {p_1}^2 + {p_2}^2 = P_{Exp}$	**0.5586**
11. From the counting results in column ***n_{ik}, k=1*** the factors $n_{ik}\cdot(n_{ik}-1)$ are calculated line-by-line and entered in the corresponding line of column ***$n_{ik}\cdot(n_{ik}-1)$, k=1***	$n_{i1} \cdot (n_{i1} - 1)$	
12. From the counting results in column ***n_{ik}, k=2*** the factors $n_{ik}\cdot(n_{ik}-1)$ are calculated line-by-line and entered in the corresponding line of column ***$n_{ik}\cdot(n_{ik}-1)$, k=2***	$n_{i2} \cdot (n_{i2} - 1)$	
13. All values in the columns ***$n_{ik}\cdot(n_{ik}-1)$, k=1*** and ***$n_{ik}\cdot(n_{ik}-1)$, k=2*** are summed up	$\sum_{k=1}^{N_C=2}\sum_{i=1}^{N_O=50} n_{ik}(n_{ik}-1) = \sum_{k=1}^{N_C=2} {n_k}^* = n^*$	3272
14. An arbitrary line of the columns ***n_{ik}, k=1*** and ***n_{ik}, k=2*** is summed up	$\sum_{k=1}^{N_C=2} n_{ik} = N_R$ *(for an arbitrary i)*	9
15. From the result in step 14, the factor $N_R\cdot(N_R-1)$ is calculated: *9 x (9 – 1) = 9 x 8 = 72*	$N_R \cdot (N_R - 1) = {N_R}^*$	72
16. The result from step 15 is multiplied by the number of test objects N_O: *50 x 72 = 3600*	$N_O \cdot N_R \cdot (N_R - 1) = N_O \cdot {N_R}^* = N^*$	3600
17. The result from step 13 is divided by the result from step 16: *3272 / 3600 = 0.9089*	$\sum_{k=1}^{N_C=2}\sum_{i=1}^{N_O=50} \frac{n_{ik}(n_{ik}-1)}{N_O \cdot N_R(N_R-1)} = \frac{n^*}{N^*} = P_{Obs}$	**0.9089**
18. Kappa is calculated with the results from steps 10 and 17 according to formula	$\frac{P_{Obs} - P_{Exp}}{1 - P_{Exp}} = \kappa$	**0.7936**

Categories:	**0 - Not OK** **1 - OK**	**A x B x C**

Test Object No.	Reference	Appraiser – Trial								
		A-1	A-2	A-3	B-1	B-2	B-3	C-1	C-2	C-3
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	1	1	1	0	1	1	0	1	0	0
7	1	1	1	1	1	1	1	1	0	1
8	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1
12	0	0	0	0	0	0	0	0	1	0
13	1	1	1	1	1	1	1	1	1	1
14	1	1	1	0	1	1	1	1	0	0
15	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1
21	1	1	1	0	1	0	1	0	1	0
22	0	0	0	1	0	1	0	1	1	0
23	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1
25	0	0	0	0	0	0	0	0	0	0
26	0	0	1	0	0	0	0	0	0	1
27	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1
30	0	0	0	0	0	0	1	0	0	0
31	1	1	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1	1
34	0	0	0	1	0	0	1	0	1	1
35	1	1	1	1	1	1	1	1	1	1
36	1	1	1	0	1	1	1	1	0	1
37	0	0	0	0	0	0	0	0	0	0
38	1	1	1	1	1	1	1	1	1	1
39	0	0	0	0	0	0	0	0	0	0
40	1	1	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1	1	1
42	0	0	0	0	0	0	0	0	0	0
43	1	1	0	1	1	1	1	1	1	0
44	1	1	1	1	1	1	1	1	1	1
45	0	0	0	0	0	0	0	0	0	0
46	1	1	1	1	1	1	1	1	1	1
47	1	1	1	1	1	1	1	1	1	1
48	0	0	0	0	0	0	0	0	0	0
49	1	1	1	1	1	1	1	1	1	1
50	0	0	0	0	0	0	0	0	0	0

	n_{ik}			$n_{ik}*(n_{ik}-1)$	
	k=1	k=2		k=1	k=2
Rating:	**0**	**1**	**Rating:**	**0**	**1**
	0	9		0	72
	0	9		0	72
	9	0		72	0
	9	0		72	0
	9	0		72	0
	4	5		12	20
	1	8		0	56
	0	9		0	72
	9	0		72	0
	0	9		0	72
	0	9		0	72
	8	1		56	0
	0	9		0	72
	3	6		6	30
	0	9		0	72
	0	9		0	72
	0	9		0	72
	0	9		0	72
	0	9		0	72
	4	5		12	20
	5	4		20	12
	0	9		0	72
	0	9		0	72
	9	0		72	0
	7	2		42	2
	0	9		0	72
	0	9		0	72
	0	9		0	72
	8	1		56	0
	0	9		0	72
	0	9		0	72
	0	9		0	72
	5	4		20	12
	0	9		0	72
	2	7		2	42
	9	0		72	0
	0	9		0	72
	9	0		72	0
	0	9		0	72
	0	9		0	72
	9	0		72	0
	2	7		2	42
	0	9		0	72
	9	0		72	0
	0	9		0	72
	0	9		0	72
	9	0		72	0
	0	9		0	72
	9	0		72	0

$\sum_{i=1}^{N_O} n_{ik} = n_k$	**148**	**302**	$\sum_{i=1}^{N_O} n_{ik}(n_{ik}-1) = n_k^*$	**1020** **2252**
$\sum_{k=1}^{N_C} n_k = N$	**450**		$\sum_{k=1}^{N_C} n_k^* = n^*$	**3272**
$\frac{n_k}{N} = p_k$	**0.3289**	**0.6711**	$\sum_{k=1}^{N_C} n_{ik} = N_R$	**9**
$p_k \cdot p_k = p_k^2$	**0.1082**	**0.4504**	$N_R \cdot (N_R - 1) = N_R^*$	**72**
			N_O	**50**
			$N_O \cdot N_R^* = N^*$	**3600**
$\sum_{k=1}^{N_C} p_k^2 = P_{Exp}$	**0.5586**		$\frac{n^*}{N^*} = P_{Obs}$	**0.9089**
			$\frac{P_{Obs} - P_{Exp}}{1 - P_{Exp}} = \kappa$	**0.7936**

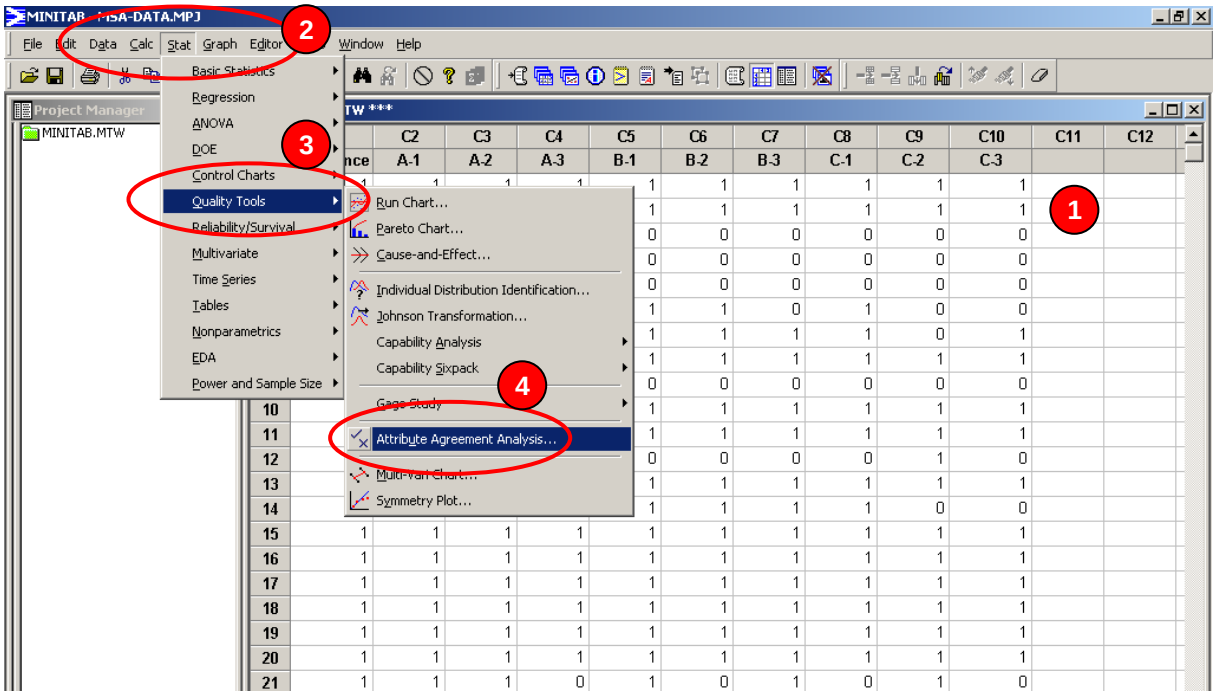
Number of test objects	N_O =	50
Number of appraisers	N_A =	3
Number of trials per appraiser	N_T =	3
Number of categories	N_C =	2

Procedure 7: Test results (ratings) and complete analysis

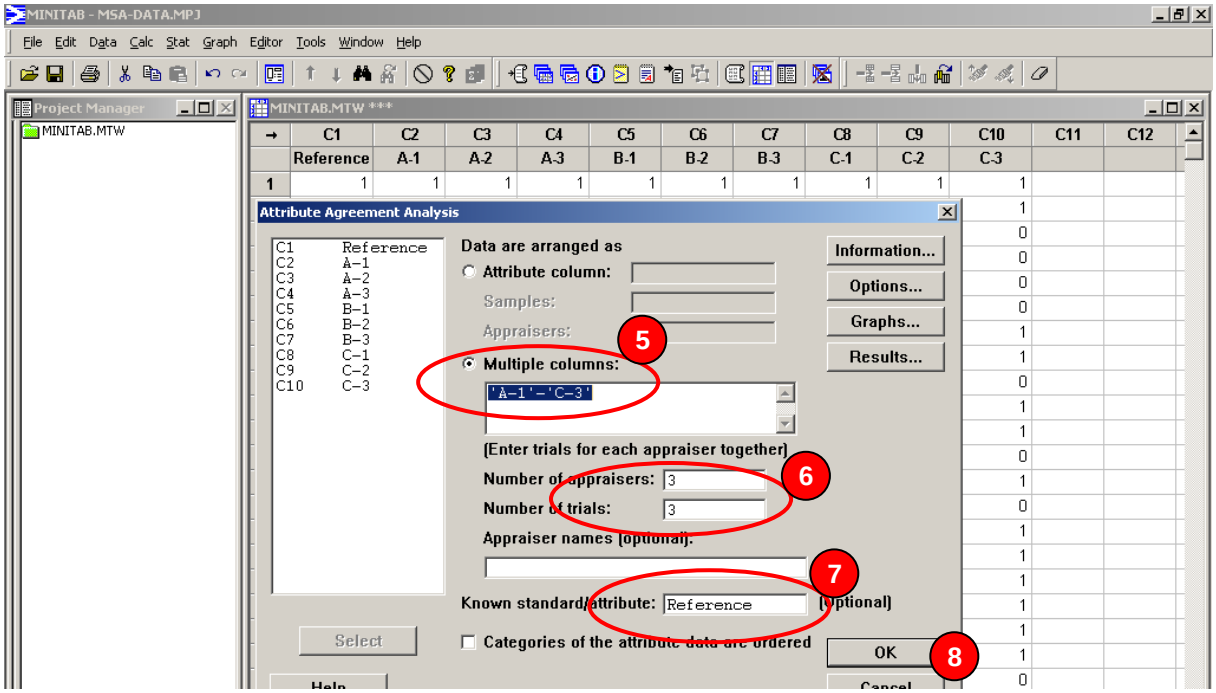
										Number of pair-wise consistent combinations per test object i (i = 1, … N_O)									
Categories: 0 - Not OK 1 - OK										A x A	B x B	C x C	A x B x C	A-1 x Ref	A-2 x Ref	A-3 x Ref	B-1 x Ref	B-2 x Ref	B-3 x Ref	C-1 x Ref	C-2 x Ref	C-3 x Ref
Test Object No.	Reference	Appraiser – Trial A-1	A-2	A-3	B-1	B-2	B-3	C-1	C-2	C-3												
1	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2
2	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2
3	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2
4	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2
5	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2
6	1	1	1	0	1	1	0	1	0	0	2	2	2	32	2	2	0	2	2	0	2	0	0
7	1	1	1	1	1	1	1	1	0	1	6	6	2	56	2	2	2	2	2	2	2	0	2
8	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
9	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
10	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
11	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
12	0	0	0	0	0	0	0	0	1	0	6	6	2	56	2	2	2	2	2	2	2	0	2
13	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
14	1	1	1	0	1	1	1	1	0	0	2	6	2	36	2	2	0	2	2	2	2	0	0
15	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
16	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
17	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
18	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
19	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
20	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
21	1	1	1	0	1	0	1	0	1	0	2	2	2	32	2	2	0	2	0	2	0	2	0
22	0	0	0	1	0	1	0	1	1	0	2	2	2	32	2	2	0	2	0	2	0	0	2
23	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
24	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
25	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
26	0	0	1	0	0	0	0	0	0	1	2	6	2	44	2	0	2	2	2	2	2	2	0
27	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
28	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
29	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
30	0	0	0	0	0	0	1	0	0	0	6	2	6	56	2	2	2	2	2	0	2	2	2
31	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
32	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
33	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
34	0	0	0	1	0	0	1	0	1	1	2	2	2	32	2	2	0	2	2	0	2	0	0
35	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
36	1	1	1	0	1	1	1	1	0	1	2	6	2	44	2	2	0	2	2	2	2	0	2
37	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
38	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
39	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
40	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
41	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
42	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
43	1	1	0	1	1	1	1	1	1	0	2	6	2	44	2	0	2	2	2	2	2	2	0
44	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
45	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
46	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
47	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
48	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
49	1	1	1	1	1	1	1	1	1	1	6	6	6	72	2	2	2	2	2	2	2	2	2
50	0	0	0	0	0	0	0	0	0	0	6	6	6	72	2	2	2	2	2	2	2	2	2
Observed number of pair-wise consistent combinations n*										**268**	**280**	**260**	**3272**	**100**	**96**	**88**	**100**	**96**	**94**	**96**	**86**	**88**
Possible number of pair-wise consistent combinations N*										**300**	**300**	**300**	**3600**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Observed portion of pair-wise consistent combinations P_{Obs} = n* / N*										**0.8933**	**0.9333**	**0.8667**	**0.9089**	**1.0000**	**0.9600**	**0.8800**	**1.0000**	**0.9600**	**0.9400**	**0.9600**	**0.8600**	**0.8800**
Observed number of ratings per category n_k (k = 1, … N_C)										**50** **100**	**47** **103**	**51** **99**	**148** **302**	**32** **68**	**32** **68**	**34** **66**	**32** **68**	**32** **68**	**31** **69**	**32** **68**	**33** **67**	**34** **66**
Total number of ratings N										**150**	**150**	**150**	**450**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Observed portion of ratings per category n_k / N (k = 1, … N_C)										**0.3333** **0.6667**	**0.3133** **0.6867**	**0.34** **0.66**	**0.3289** **0.6711**	**0.32** **0.68**	**0.32** **0.68**	**0.34** **0.66**	**0.32** **0.68**	**0.32** **0.68**	**0.31** **0.69**	**0.32** **0.68**	**0.33** **0.67**	**0.34** **0.66**
Expected portion of randomly consistent ratings $P_{Exp} = \Sigma_k (n_k/N)^2$										**0.5556**	**0.5697**	**0.5512**	**0.5586**	**0.5648**	**0.5648**	**0.5512**	**0.5648**	**0.5648**	**0.5722**	**0.5648**	**0.5578**	**0.5512**
Kappa: $\frac{P_{Obs} - P_{Exp}}{1 - P_{Exp}} = \kappa$										**0.7600**	**0.8451**	**0.7029**	**0.7936**	**1.0000**	**0.9081**	**0.7326**	**1.0000**	**0.9081**	**0.8597**	**0.9081**	**0.6834**	**0.7326**
Kappa: Each appraiser versus reference (mean values)														**0.8802**			**0.9226**			**0.7747**		
Kappa: All appraisers versus reference (mean value)														**0.8592**								

G.3 Analysis using commercial statistics software: Minitab

The following quick guide allows the schematic processing of standard cases. In all other cases, a user training is indispensible, in particular, if default settings are changed, other software features are used, etc.



(1) Prepare a worksheet, fill in reference data and ratings to be analyzed
(2) Select menu *"Stat"*
(3) Select sub-menu *"Quality Tools"*
(4) Select sub-menu *"Attribute Agreement Analysis"*



(5) Check field *"Multiple Columns"* and enter the names of the columns containing the ratings
(6) Click on field *"Number of appraisers"* and enter the number of appraisers
Click on field *"Number of trials"* and enter the number of trials
(7) Click on field *"Known standard/attribute"* and enter the name of the column containing the reference ratings
(8) OK

The results of the analysis are displayed in the so-called *"Session Window"*.

Results of the analysis in the Minitab Session Window

Attribute Agreement Analysis for A-1; A-2; A-3; B-1; B-2; B-3; C-1; C-2; C-3

Within Appraisers

```
Assessment Agreement

Appraiser  # Inspected  # Matched  Percent      95 % CI
1                  50         42    84.00  (70.89; 92.83)
2                  50         45    90.00  (78.19; 96.67)
3                  50         40    80.00  (66.28; 89.97)

# Matched: Appraiser agrees with him/herself across trials.


Fleiss' Kappa Statistics

Appraiser  Response     Kappa   SE Kappa        Z  P(vs > 0)
1          0         0.760000  0.0816497   9.3081     0.0000
           1         0.760000  0.0816497   9.3081     0.0000
2          0         0.845073  0.0816497  10.3500     0.0000
           1         0.845073  0.0816497  10.3500     0.0000
3          0         0.702911  0.0816497   8.6089     0.0000
           1         0.702911  0.0816497   8.6089     0.0000
```


Each Appraiser vs Standard

```
Assessment Agreement

Appraiser  # Inspected  # Matched  Percent      95 % CI
1                  50         42    84.00  (70.89; 92.83)
2                  50         45    90.00  (78.19; 96.67)
3                  50         40    80.00  (66.28; 89.97)

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


Assessment Disagreement

Appraiser  # 1 / 0  Percent  # 0 / 1  Percent  # Mixed  Percent
1                0     0.00        0     0.00        8    16.00
2                0     0.00        0     0.00        5    10.00
3                0     0.00        0     0.00       10    20.00


# 1 / 0:  Assessments across trials = 1 / standard = 0.
# 0 / 1:  Assessments across trials = 0 / standard = 1.
# Mixed: Assessments across trials are not identical.


Fleiss' Kappa Statistics

Appraiser  Response     Kappa   SE Kappa        Z  P(vs > 0)
1          0         0.880236  0.0816497  10.7806     0.0000
           1         0.880236  0.0816497  10.7806     0.0000
2          0         0.922612  0.0816497  11.2996     0.0000
           1         0.922612  0.0816497  11.2996     0.0000
3          0         0.774703  0.0816497   9.4881     0.0000
           1         0.774703  0.0816497   9.4881     0.0000
```

Between Appraisers

```
Assessment Agreement

# Inspected  # Matched  Percent      95 % CI
         50         39    78.00  (64.04; 88.47)

# Matched: All appraisers' assessments agree with each other.


Fleiss' Kappa Statistics

Response     Kappa   SE Kappa        Z  P(vs > 0)
0         0.793606  0.0235702  33.6698     0.0000
1         0.793606  0.0235702  33.6698     0.0000
```

All Appraisers vs Standard

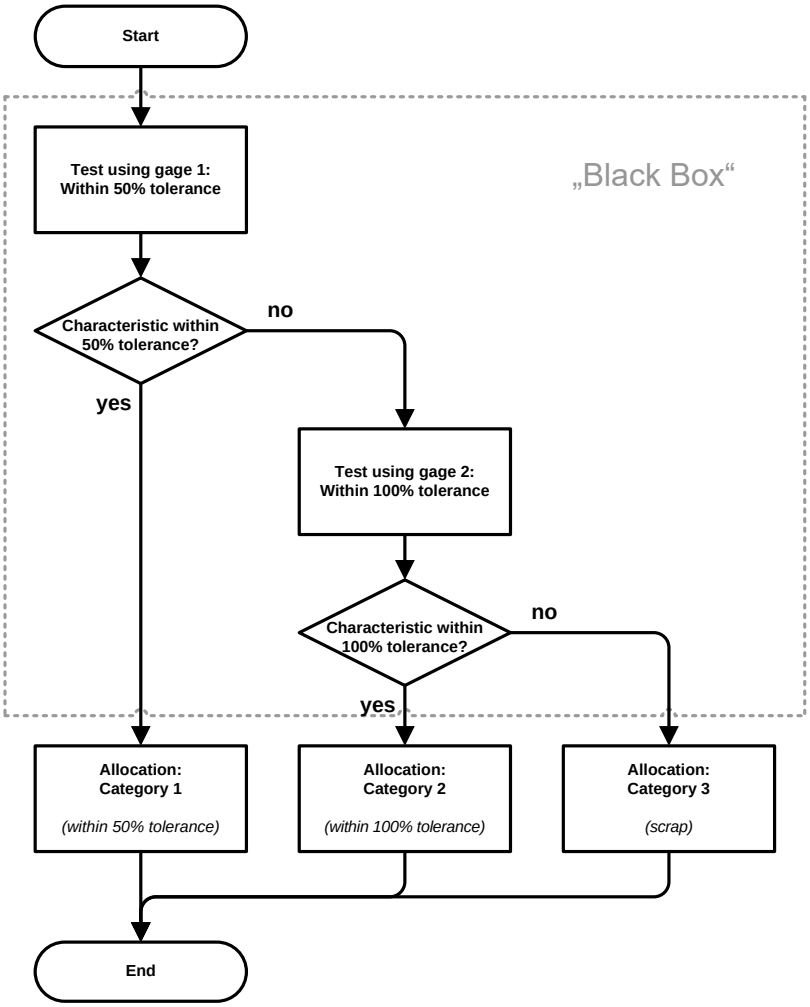
```
Assessment Agreement

# Inspected  # Matched  Percent      95 % CI
         50         39    78.00  (64.04; 88.47)

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


Fleiss' Kappa Statistics

Response     Kappa   SE Kappa        Z  P(vs > 0)
0         0.859184  0.0471405  18.2260     0.0000
1         0.859184  0.0471405  18.2260     0.0000
```

G.4 Single stage and multiple stage test processes: Attribute gage as an example

Test scenario

The test objects must be sorted according to their deviations from the nominal value, i.e. one of the three categories *"Within 50% tolerance"*, *"Within 100% tolerance"* and *"Scrap"* has to be allocated to each test object. To do this, two gages are used to test for compliance with the 50% and 100% tolerance limits.

Test process



Two different approaches are possible when verifying this test process.

- Single stage process: The gage tests are seen as a single "black box" which allocates one of several categories in a single test step. A single test process capability is assigned to the entire "black box".
- Multiple stage process: The gage tests are seen as a serial test process which allocates one of two categories at a time in multiple test steps. An individual test process capability is assigned to each individual process step.

The multiple stage process has been shown to be more reliable in practice when testing manually. The single stage process is mainly appropriate for semi-automatic or fully automatic test processes.

Table of symbols

%AV	AV related to a reference value (e. g. tolerance)
%EV	EV related to a reference value (e. g. tolerance)
%GRR	GRR related to a reference value (e. g. tolerance)
%PV	PV related to a reference value (e. g. tolerance)
α	Significance level
AV	**A**ppraiser **v**ariation (reproducibility)
B'_{Eob} , B'_{Eun}	Factors required for calculation of control limits of s-stability charts
C_g	Potential capability index (does not include the systematic measurement error) C - **c**apability; g - **g**auge
C_{gk}	Critical capability index (includes the systematic measurement error) k - Japanese: **k**atayori (English: systematic error)
d	Average width of uncertainty ranges with non-uniform test results (procedure 6)
d1 , d2	Widths of uncertainty ranges with non-uniform test results (procedure 6)
E'_E	Factor required for calculation of control limits of individual value charts
EV	**E**quipment **v**ariation (repeatability)
f	Number of degrees of freedom
GRR	**G**auge **r**epeatability and **r**eproducibility; total variation of the measurement process only, i.e. without portions from serial part variation (spread of measuring objects)
i	Index of measured values and/or parts (measuring objects) in a sample ($1 \leq i \leq n$)
k	Number of appraisers
K_1 , K_2 , K_3	Factors required for calculation of EV, AV and PV using the average range method (ARM)
LCL	Lower control limit of $\bar{x}$-stability charts
LCL_s	Lower control limit of s-stability charts
LSL	Lower specification limit
LSL^*	Natural (i.e. physical) lower limit
LSL_0	Lower acceptance limit for measured values z in case of an one-sided lower specification limit LSL (an upper limiting value such as USL or USL* does not exist)
μ	Mean value (expected value) of a population
n	Sample size: Number of measurements and/or parts (measuring objects) in a sample
ndc	**N**umber of **d**istinct **c**ategories which can be distinguished by the measurement process within the spread of measuring objects
PV	**P**art **v**ariation
r	Number of measurement series (number of measurements per measuring object)
$R_{A,i}$	Range of measured values of appraiser A at part no. i
$\overline{R_A}$	Mean value of ranges $R_{A,i}$
$R_{B,i}$	Range of measured values of appraiser B at part no. i
$\overline{R_B}$	Mean value of ranges $R_{B,i}$
$R_{C,i}$	Range of measured values of appraiser C at part no. i
$\overline{R_C}$	Mean value of ranges $R_{C,i}$
R_i	Range of measured values at part no. i (procedure 3)
R_p	Range of mean values $\bar{x}_i$
$\bar{R}$	Mean value of ranges R_i (procedure 3)
$\bar{\bar{R}}$	Mean value of average ranges $\overline{R_A}$, $\overline{R_B}$, $\overline{R_C}$
$R_{\bar{x}}$	Range of mean values $\overline{x_A}$, $\overline{x_B}$, $\overline{x_C}$

s	Standard deviation of measured values x_i
T	Tolerance (of the characteristic to be measured)
T^*	(Positive) difference of a specification limit and a natural limiting value
$t_{f;\,p}$	Quantile of a t-distribution with probability p for f degrees of freedom
TV	**T**otal **v**ariation of measurement process and parts (measuring objects)
U	Measurement uncertainty assigned to a measurement result
U_{cal}	Uncertainty of calibration
UCL	Upper control limit of $\bar{x}$-stability charts
UCL_s	Upper control limit of s-stability charts
u_p	Factor required for calculation of control limits of $\bar{x}$-stability charts
USL	Upper specification limit
USL^*	Natural (i.e. physical) upper limit
USL_0	Upper acceptance limit for measured values z in case of an one-sided upper specification limit USL (a lower limiting value such as LSL or LSL* does not exist)
$\overline{x_A}$	Mean value of measured values of appraiser A
$\overline{x_B}$	Mean value of measured values of appraiser B
$\overline{x_C}$	Mean value of measured values of appraiser C
x_i	Measured value no. i
x_m	Reference value of the reference part (**m**aster)
$\bar{x}$	Mean value of measured values x_i
$\bar{x}_i$	Mean value of measured values at part no. i (measuring object no. i)
z	Measured value determined after completion of the capability study (e. g. during the production process)
z_0	Conventional true value assigned to the measured value z
$\bar{z}$	(Hypothetical) mean value of the measured values z

Symbols differently defined and/or additionally used in individual chapters

Appendix D (procedure 2 and 3): See symbol definitions in the respective sub-sections

Appendix E (procedure 4):

$\%Bias_i$	Systematic measurement error of measured values ξ_{ik} at reference part no. i, related to the tolerance of the characteristic to be measured
$\%SDev_i$	Standard deviation s_i of measured values ξ_{ik} at reference part no. i, related to the tolerance of the characteristic to be measured
$\%SDev_{max\,i}$	Maximum $\%SDev_i$ that can be accepted for reference part no. i
a	Slope of regression line (line of best fit)
b	Intercept of regression line (line of best fit)
ε	Arbitrary positive number
$F_{f_1;\,f_2;\,p}$	Quantile of the F-distribution with probability p for f_1 and f_2 degrees of freedom (f_1 used in the numerator, f_2 used in the denominator)
F_{LM}	Test statistic for statistical significance of the deviation of measured values from the **l**inear **m**odel
g	Number of reference parts
i	Index of reference parts ($1 \leq i \leq g$)
k	Index of measured values at a particular reference part ($1 \leq k \leq m$)
LCI	**L**ower limit of the **c**onfidence **i**nterval for systematic measurement errors
m	Number of measured values per reference part

s	(Residual) variation of measurement errors around the regression line (line of best fit)
s_i	Standard deviation of measured values ξ_{ik} at reference part no. i
t_a	Test statistic for statistical significance of slope a of the regression line (line of best fit)
t_b	Test statistic for statistical significance of intercept b of the regression line (line of best fit)
$t_{f;\,p}$	Quantile of a t-distribution with probability p for f degrees of freedom
UCI	**U**pper limit of the **c**onfidence **i**nterval for systematic measurement errors
x_0	Arbitrary value on x-axis around reference values x_i
x_i	Reference value of reference part no. i
$\bar{\xi}_i$	Mean value of measured values ξ_{ik} at reference part no. i
$\bar{x}$	Mean value of reference values x_i
ξ_{ik}	Measured value no. k at reference part no. i
y_{ik}	Measurement error of measured value no. k at reference part no. i, deviation from reference value x_i
$\bar{y}_i$	Systematic measurement error of all measured values ξ_{ik} at reference part no. i, deviation from reference value x_i

Chapter 5.2, Appendix G (procedure 7):

i	Index of test object ($1 \leq i \leq N_O$)
k	Index of rating category ($1 \leq k \leq N_C$)
κ	Observed non-random agreements related to all possible non-random agreements
κ_{INDEX}	κ for a certain criterion defined by *"INDEX"* (e. g. A x B x C) (e. g. κ_{AxBxC} : κ for the criterion *"agreement of ratings of appraisers A, B, C"*)
N	Total number of ratings
N_O	Total number of test **o**bjects
N_A	Total number of **a**ppraisers
n_A	Number of appraisers that is included in the analysis depending on the particular criterion to be analysed
N_C	Total number of rating **c**ategories
n_{ik}	Number of allocations of test object i to rating category k
N_R	Total number of **r**atings per test object
N_T	Total number of **t**rials (test runs) per appraiser
n_T	Number of trials (test runs) per appraiser that is included in the analysis depending on the particular criterion to be analysed
P_{Exp}	Portion of randomly **exp**ected pair-wise consistent rating pairs from all possible rating pairs
p_k	Estimated value for the probability with which the allocation to rating category k can be expected in a random experiment
P_{Obs}	Portion of actually **obs**erved pair-wise consistent rating pairs from all possible rating pairs

solara.MP® forms:

%IA	IA related to a reference value (e. g. tolerance)
%RE	RE related to a reference value (e. g. tolerance)
|Bi|	Systematic measurement error (**bi**as)
i	Index of measured values in a sample: $1 \leq i \leq n_{tot}$ (procedure 1)
IA	**I**nter**a**ction between appraisers and parts
j	*(see n)*
LSL	Lower specification limit

M	Central position of measured values x_i (procedure 5)
$\hat{\mu}_{INDEX}$	Estimator for the location of a population (INDEX: code for calculation method)
n	Index of parts (measuring objects) in a sample (procedures 2, 3, 6)
n_{eff}	Total number of measured values that comply with the specification limits
n_{tot}	Total number of measured values
RE	**Re**solution of the measuring system
Ref	Continuous reference values of the reference parts (procedure 6)
RF	**R**e**f**erence, e. g. tolerance
R_g	Range of measured values x_i
s_g	Standard deviation of measured values x_i
$s_{g\,j}$	Standard deviation of measured values at part no. j (measuring object no. j)
$\hat{\sigma}_{INDEX}$	Estimator for the variation of a population (INDEX: code for calculation method)
$\bar{s}_g$	Mean of standard deviations of all samples (procedure 5)
$\tilde{s}_g$	Median of standard deviations of all samples (procedure 5)
T_m	Center point between upper and lower specification limit (“tolerance center”)
$T_{min}(\%GRR)$	Minimum reference value (tolerance) required for compliance with the capability criterion for %GRR
$T'_{min}(\%GRR)$	Minimum reference value (tolerance) required for conditional compliance with the capability criterion for %GRR
$T_{min}(C_g)$	Minimum reference value (tolerance) required for compliance with the capability criterion for C_g
$T_{min}(C_{gk})$	Minimum reference value (tolerance) required for compliance with the capability criterion for C_{gk}
$T_{min}(RE)$	Minimum reference value (tolerance) required for compliance with the criterion for RE
USL	Upper specification limit
$x_{A;1}$, $x_{A;2}$	Measured values of the 1st and 2nd measurement series of appraiser A (procedures 2, 3); test results of the 1st and 2nd trial (test run) of appraiser A (procedure 6)
$x_{A;1} \ldots x_{A;12}$	Measured values no. 1 to 12 at the reference part (procedure 4)
$x_{B;1}$, $x_{B;2}$	Measured values of the 1st and 2nd measurement series of appraiser B (procedures 2, 3); test results of 1st and 2nd trial (test run) of appraiser B (procedure 6)
$x_{C;1}$, $x_{C;2}$	Measured values of the 1st and 2nd measurement series of appraiser C (procedures 2, 3);
x_i	Measured value no. i (in a sample)
x_{maxg}, x_{ming}	Maximum and minimum values of the measured values x_i
$\bar{x}_g$	Mean value of the measured values x_i
$\bar{x}_{g\,j}$	Mean value of the measured values at part no. j (measuring object no. j)
$\bar{x}_{g\,Ref}$	Reference value of the reference part (procedure 4)
$\bar{\bar{x}}_g$	Mean value of the mean values $\bar{x}_{g\,j}$
$\tilde{\bar{x}}_{g\,j}$	Median value of the mean values $\bar{x}_{g\,j}$

Definition of terms

NOTE 1: The following definitions of terms were taken from the respective standards cited in this document. Corresponding notes were only adopted in single cases if they were considered directly relevant and/or essential for understanding a standardized term. Otherwise, the respective standard should be referenced for notes and examples.

*NOTE 2: "Editorial notes" are **not** part of the respective standard.*

NOTE 3: The definitions of terms according to [VIM] were used preferably. If terms are not contained in [VIM], the most current definition from the standards [ISO 3534-2], [ISO 3534-1], [ISO 9000], [ISO 10012], [DIN 1319-2] and [DIN 1319-1] were adopted (or listed additionally in some cases). Non-standardized definitions are only used if the listed standards do not provide a definition.

NOTE 4: Terms whose definitions are contained in the following summary are in bold if they are used in definitions of other terms.

accuracy *(Ger. Genauigkeit)*: see **measurement accuracy** [VIM, 2.13]

adjustment of a measuring system *(Ger. Justierung eines Messsystems)*
set of operations carried out on a **measuring system** so that it provides prescribed **indications** corresponding to given **values** of a **quantity** to be measured [VIM, 3.11]

bias *(Ger. Bias der Messung):* see **measurement bias** [VIM, 2.18]

characteristic *(Ger. Merkmal)*
distinguishing feature

NOTE 1: A characteristic can be inherent or assigned.

NOTE 2: A characteristic can be qualitative or quantitative.

NOTE 3: There are various classes of characteristics such as the following:

- *physical (e. g. mechanical, electrical, chemical, biological);*
- *sensory (e. g. relating to smell, touch, taste, sight, hearing);*
- *behavioral (e. g. courtesy, honesty, veracity)*
- *temporal (e. g. punctuality, reliability, availability);*
- *ergonomic (e. g. physiological characteristic or related to human safety);*
- *functional (e. g. maximum speed of an aircraft).*

[ISO 3534-2, 1.1.1]

conformity *(Ger. Konformität)*
fulfilment of a **requirement** [ISO 9000, 3.6.1]

conformity evaluation *(Ger. Konformitätsbewertung)*
systematic examination of the extent to which an **item/entity** fulfils specified **requirements**
[ISO 3534-2, 4.1.1]

continuous characteristic *(Ger. kontinuierliches Merkmal)*
characteristic providing values which are **measured values** of a physical **quantity** (e. g. weight, length, current, temperature); *in analogy to* [CDQ 0301]

EDITORIAL NOTE: A definition based on number theory can be found in DIN 55350, part 12 (German only)

conventional quantity value *(Ger. vereinbarter Wert)*
quantity value attributed by agreement to a **quantity** for a given purpose

NOTE 1: The term "conventional true quantity value" is sometimes used for this concept, but its use is discouraged.

*NOTE 2: Sometimes a conventional quantity value is an estimate of a **true quantity value**.*

*NOTE 3: A conventional quantity value is generally accepted as being associated with a suitably small **measurement uncertainty**, which might be zero.*

[VIM, 2.12]

EDITORIAL NOTE: The term "conventional quantity value" (or briefly "conventional value") obviously replaces the term "conventional true value" which is no longer contained in the current release of [VIM].

conventional true value *(Ger. richtiger Wert)*
value of a quantity or quantitative **characteristic** which, for a given purpose, may be substituted for a **true value**

*NOTE 1: A conventional true value is, in general, regarded as sufficiently close to the **true value** for the difference to be insignificant for the given purpose.*

[ISO 3534-2, 3.2.6]

conventional value *(Ger. vereinbarter Wert)*: see **conventional quantity value** [VIM, 2.12]

discrete characteristic *(Ger. diskretes Merkmal)*
characteristic providing values which are obtained by counting a countable **nominal property** or attribute (e. g. good / bad, pass / fail, red / green / blue); *in analogy to* [CDQ 0301]

EDITORIAL NOTE: A definition based on number theory can be found in DIN 55350, part 12 (German only).

discretized continuous characteristic *(Ger. diskretisiertes kontinuierliches Merkmal)*
characteristic providing discrete values which are obtained by classifying continuous values according to their compliance with a criterion (e. g. **measured value** inside or outside the specification limits)

EDITORIAL NOTE: Standardized definition of term unavailable.

entity *(Ger. Einheit):* see **item** [ISO 3534-2, 1.2.11]

identical test / measurement item *(Ger. identische Untersuchungseinheit)*
sample which is prepared and can be presumed to be identical for the intended purpose [ISO 3534-2, 1.2.34]

independent test / measurement results *(Ger. unabhängige Ergebnisse)*
test results or **measurement results** obtained in a manner that they are not influenced by each other [ISO 3534-2, 3.4.3]

indicating measuring instrument *(Ger. anzeigendes Messgerät)*
measuring instrument providing an output signal carrying information about the **value** of the **quantity** being measured

*NOTE 1: An indicating measuring instrument may provide a record of its **indication**.*

NOTE 2: An output signal may be presented in visual or acoustic form. It may also be transmitted to one or more other devices.

[VIM, 3.3]

indication *(Ger. Anzeige)*
quantity value provided by a **measuring instrument** or a **measuring system** [VIM, 4.1]

inspection *(Ger. Prüfung)*
conformity evaluation by observation and judgement accompanied as appropriate by **measurement**, testing or gauging [ISO 3534-2, 4.1.2]

intermediate measurement precision *(Ger. Vergleichpräzision)*
measurement precision under a set of **intermediate precision conditions** of measurement [VIM, 2.23]

*EDITORIAL NOTE: Cf. "**reproducibility**" according to [ISO 3534-2, 3.3.10]*

intermediate precision condition *(Ger. Vergleichbedingung)*
condition of **measurement**, out of a set of conditions that includes the same **measurement procedure**, same location, and replicate measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes

*NOTE 1: The changes can include new calibrations, calibrators, operators, and **measuring systems**.*

[VIM, 2.22]

*EDITORIAL NOTE: Cf. "**reproducibility conditions**" according to [ISO 3534-2, 3.3.11].*

item *(Ger. Einheit)*
anything that can be described and considered separately [ISO 3534-2, 1.2.11]

kind *(Ger. Art einer Größe, Größenart)*: see **kind of quantity** [VIM, 1.2]

kind of quantity *(Ger. Art einer Größe, Größenart)*
aspect common to mutually comparable **quantities** [VIM, 1.2]

lot *(Ger. Los)*
definite part of a **population** constituted under essentially the same conditions as the population with respect to the sampling process [ISO 3534-2, 1.2.4]

material measure *(Ger. Maßverkörperung)*
measuring instrument reproducing or supplying, in a permanent manner during its use, **quantities** of one or more given **kinds**, each with an assigned **quantity value**

*NOTE 1: The **indication** of a material measure is its assigned quantity value.*

*NOTE 2: A material measure can be a **measurement standard**.*

[VIM, 3.6]

measurand *(Ger. Messgröße)*
quantity intended to be measured [VIM, 2.3]

measured quantity value *(Ger. Messwert)*
quantity value representing a **measurement result** [VIM, 2.10]

measured value *(Ger. Messwert):* see **measured quantity value** [VIM, 2.10]

measurement *(Ger. Messung)*
process of experimentally obtaining one or more **quantity values** that can reasonably be attributed to a **quantity**

*NOTE 1: Measurement does not apply to **nominal properties**.*

NOTE 2: Measurement implies comparison of quantities and includes counting of entities.

*NOTE 3: Measurement presupposes a description of the quantity commensurate with the intended use of a **measurement result**, a **measurement procedure**, and a calibrated **measuring system** operating according to the specified measurement procedure, including the measurement conditions.*

[VIM, 2.1]

measurement accuracy *(Ger. Messgenauigkeit)*
closeness of agreement between a **measured quantity value** and a **true quantity value** of a **measurand** [VIM, 2.13]

measurement bias *(Ger. Bias der Messung)*
estimate of a **systematic measurement error** [VIM, 2.18]

measurement error *(Ger. Messabweichung)*
measured quantity value minus a **reference quantity value** [VIM, 2.16]

measurement method *(Ger. Messmethode)*
generic description of a logical organization of operations used in a **measurement** [VIM, 2.5]

measurement model *(Ger. Modell der Messung)*
mathematical relation among all **quantities** known to involved in a **measurement** [VIM, 2.48]

measurement precision *(Ger. Messpräzision)*
closeness of agreement between **indications** or **measured quantity values** obtained by replicate **measurements** on the same or similar objects under specified conditions [VIM 2.15]

*EDITORIAL NOTE: Cf. "**precision**" according to [ISO 3534-2, 3.3.4].*

measurement principle *(Ger. Messprinzip)*
phenomenon serving as a basis of a **measurement** [VIM, 2.4]

measurement procedure *(Ger. Messverfahren)*
detailed description of a **measurement** according to one or more **measurement principles** and to a given **measurement method** based on a **measurement model** and including any calculation to obtain a **measurement result** [VIM, 2.6]

measurement process *(Ger. Messprozess)*
set of operations to determine the **value of a quantity** [ISO 9000, 3.10.2]

measurement repeatability *(Ger. Wiederholpräzision)*
measurement precision under a set of **repeatability conditions of measurement** [VIM, 2.21]

*EDITORIAL NOTE: Cf. "**repeatability**" according to [ISO 3534-2, 3.3.5].*

measurement result *(Ger. Messergebnis)*
set of **quantity values** being attributed to a **measurand** together with any other available relevant information [VIM, 2.9]

measurement standard *(Ger. Normal)*
realization of the definition of a given **quantity**, with stated **quantity value** and associated **measurement uncertainty**, used as a reference

*NOTE 1: A "realization of the definition of a given quantity" can be provided by a **measuring system**, a **material measure**, or a reference material.*

[VIM, 5.1]

measurement uncertainty *(Ger. Messunsicherheit)*
non-negative parameter characterizing the dispersion of the **quantity values** being attributed to a **measurand**, based on the information used [VIM, 2.26]

measuring equipment *(Ger. Messmittel)*
measuring instrument, software, **measurement standard**, reference material or auxiliary apparatus or combination thereof necessary to realize a **measurement process** [ISO 9000, 3.10.4]

measuring instrument *(Ger. Messgerät)*
device used for making **measurements**, alone or in conjunction with one or more supplementary devices

*NOTE 1: A measuring instrument that can be used alone is a **measuring system**.*

*NOTE 2: A measuring instrument may be an **indicating measuring instrument** or a **material measure**.*

[VIM, 3.1]

measuring object *(Ger. Messobjekt)*
the object being measured in order to determine the value of the **measurand** [DIN 1319-1, 1.2]

measuring system *(Ger. Messsystem)*
set of one or more **measuring instruments** and often other devices, including any reagent and supply, assembled and adapted to give information used to generate **measured quantity values** within specified intervals for **quantities** of specified **kinds**

NOTE: A measuring system may consist of only one measuring instrument.

[VIM, 3.2]

measuring system *(Ger. Messeinrichtung)*
complete set of **measuring instruments** and any other equipment used to carry out a **measurement** [DIN 1319-1, 4.2]

*EDITORIAL NOTE: Cf. "**measuring system**" according to [VIM, 3.2].*

model *(Ger. Modell der Messung):* see **measurement model** [VIM, 2.48]

nominal property *(Ger. Nominalmerkmal)*
property of a phenomenon, body, or substance, where the property has no magnitude [VIM, 1.30]

population *(Ger. Grundgesamtheit)*
totality of **items** under consideration [ISO 3534-2, 1.2.1]

precision *(Ger. Präzision)*: see **measurement precision** [VIM, 2.15]

precision *(Ger. Präzision)*
closeness of agreement between **independent test/measurement results** obtained under stipulated conditions

*NOTE 1: Precision depends only on the distribution of random errors and does not relate to the **true value** or the specified value.*

*NOTE 2: The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the **test results** or **measurement results**. Less precision is reflected by a larger standard deviation.*

*NOTE 3: Quantitative measures of precision depend critically on the stipulated conditions. **Repeatability conditions** and **reproducibility conditions** are particular sets of extreme stipulated conditions.*

[ISO 3534-2, 3.3.4]

*EDITORIAL NOTE: Cf. "**measurement precision**" according to [VIM, 2.15].*

quantity *(Ger. Größe)*
property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference [VIM, 1.1]

quantity value *(Ger. Größenwert)*
number and reference together expressing magnitude of a **quantity** [VIM, 1.19]

random error *(Ger. zufällige Messabweichung)*: see **random measurement error** [VIM, 2.19]

random measurement error *(Ger. zufällige Messabweichung)*
component of **measurement error** that in replicate **measurements** varies in an unpredictable manner

*NOTE 1: A **reference quantity value** for a random measurement error is the average that would ensue from an infinite number of replicate measurements of the same **measurand**.*

NOTE 2: Random measurement errors of a set of replicate measurements form a distribution that can be summarized by its expectation, which is generally assumed to be zero, and its variance.

*NOTE 3: Random measurement error equals measurement error minus **systematic measurement error**.*

[VIM, 2.19]

reference lot *(Ger. Referenzlos)*
lot consisting of **reference parts**

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**lot**" [ISO 3534-2, 1.2.4].*

reference part *(Ger. Referenzteil)*
measuring object or **test object** representing the realization of the definition of a given **quantity** (e. g. a **measurement standard**) or a **nominal property** (e. g. a boundary sample)

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**measurement standard**" [VIM, 5.1].*

reference quantity value *(Ger. Referenzwert)*
quantity value used as a basis for comparison with values of **quantities** of the same **kind**

*NOTE 1: A reference quantity value can be a **true quantity value** of a **measurand**, in which case it is unknown, or a **conventional quantity value**, in which case it is known.*

*NOTE 2: A reference quantity value with associated **measurement uncertainty** is usually provided with reference to*
a) a material, e. g. a certified reference material,
b) a device, e. g. a stabilized laser,
c) a reference measurement procedure,
*d) a comparison of **measurement standards**.*

[VIM, 5.18]

*EDITORIAL NOTE: Due to insufficient standardization, the term "reference quantity value" (or briefly "reference value") is also used in a broader sense in the present booklet, i.e. it is extended to **discrete characteristics**. If the type of the reference value does not become clear out of context, the terms "continuous reference value" or "discrete reference value" (or "reference rating") are used.*

reference value *(Ger. Referenzwert)*: see **reference quantity value** [VIM, 5.18]

repeatability *(Ger. Wiederholpräzision)*: see **measurement repeatability** [VIM, 2.21]

repeatability *(Ger. Wiederholpräzision)*
precision under **repeatability conditions**

*NOTE: Repeatability can be expressed quantitatively in terms of the dispersion **characteristics** of the results.*

[ISO 3534-2, 3.3.5]

*EDITORIAL NOTE: Cf. "**measurement repeatability**" according to [VIM, 2.21].*

repeatability condition of measurement *(Ger. Wiederholbedingung)*
condition of **measurement**, out of a set of conditions that includes the same **measurement procedure**, same operators, same **measuring system**, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time

NOTE 1: A condition of measurement is a repeatability condition only with respect to a specified set of repeatability conditions.

[VIM, 2.20]

repeatability conditions *(Ger. Wiederholbedingungen)*
observation conditions where **independent test/measurement results** are obtained with the same method on **identical test/measurement items** in the same test or measuring facility by the same operator using the same equipment within short intervals of time

NOTE: Repeatability conditions include:
- *the same measurement procedure or test procedure;*
- *the same operator;*
- *the same measuring or test equipment used under the same conditions;*
- *the same location;*
- *repetition over a short period of time.*

[ISO 3534-2, 3.3.6]

reproducibility *(Ger. Vergleichpräzision)*
precision under **reproducibility conditions**

*NOTE 1: Reproducibility can be expressed quantitatively in terms of the dispersion **characteristics** of the results.*

NOTE 2: Results are usually understood to be corrected results.

[ISO 3534-2, 3.3.10]

*EDITORIAL NOTE: Cf. "**intermediate measurement precision**" according to [VIM, 2.23].*

reproducibility conditions *(Ger. Vergleichbedingungen)*
observation conditions where **independent test/measurement results** are obtained with the same method on **identical test/measurement items** in different test or measurement facilities with different operators using different equipment [ISO 3534-2, 3.3.11]

*EDITORIAL NOTE: Cf. "**intermediate precision condition**" according to [VIM, 2.22].*

requirement *(Ger. Anforderung)*
need or expectation that is stated, generally implied or obligatory [ISO 9000, 3.1.2]

resolution *(Ger. Auflösung)*
smallest change in a **quantity** being measured that causes a perceptible change in the corresponding **indication** [VIM, 4.14]

resolution of a displaying device *(Ger. Auflösung eines visuell anzeigenden Messgerätes)*
smallest difference between displayed **indications** that can be meaningfully distinguished [VIM, 4.15]

sample *(Ger. Probe, Stichprobe)*
subset of a **population** made up of one or more **sampling units** [ISO 3534-2, 1.2.17]

sampling unit *(Ger. Auswahleinheit)*
one of the individual parts into which a **population** is divided

*NOTE 1: A sampling unit can contain one or more **items**, for example a box of matches, but one **test result** will obtained for it.*

[ISO 3534-2, 1.2.14]

specification *(Ger. Spezifikation)*
document stating **requirements** [ISO 3534-2, 3.1.1]

stability of a measuring instrument *(Ger. Messbeständigkeit)*
property of a **measuring instrument**, whereby its metrological properties remain constant in time [VIM, 4.19]

statistic *(Ger. Kenngröße)*
completely specified function of random variables [ISO 3534-1, 1.8]

systematic measurement error *(Ger. systematische Messabweichung)*
component of **measurement error** that in replicate **measurements** remains constant or varies in a predictable manner

*NOTE 1: A **reference quantity value** for a systematic measurement error is a **true quantity value**, or a **measured quantity value** of a **measurement standard** of negligible **measurement uncertainty**, or a **conventional quantity value**.*

*NOTE 3: Systematic measurement error equals **measurement error** minus **random measurement error**.*

[VIM, 2.17]

test *(Ger. Ermittlung)*
technical operation that consists of the determination of one or more **characteristics** of a given product, process, or service according to a specified procedure

*NOTE 1: **Measurement** is restricted to the determination of **quantities** whereas test is used in a broader sense in the determination of characteristics by measurement or other means such as quantifying, classifying or detecting the presence or absence of one or more particular characteristics.*

[ISO 3534-2, 3.2.3]

*EDITORIAL NOTE: The terms "test" and "inspection" are often used synonymously in everyday language. Also, they are often translated to other languages using the same term in the target language (e. g. "Prüfung" in German). However, their definitions according to [ISO 9000, 3.8.2 and 3.8.3] are different, i.e. "test" does not (necessarily) include a **conformity evaluation**.*

test equipment *(Ger. Prüfmittel)*
instrument, software, standard (e. g. a boundary samples catalog), reference material or auxiliary apparatus or combination thereof necessary to realize a **test process**

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**measuring equipment**" [ISO 9000, 3.10.4].*

test object *(Ger. Prüfobjekt)*
the object being tested in order to determine the **test result**

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**measuring object**" [DIN 1319-1, 1.2].*

test process *(Ger. Prüfprozess)*
set of operations to determine a **test result**

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**measurement process**" [ISO 9000, 3.10.2].*

test result *(Ger. Ermittlungsergebnis)*
value of a **characteristic** obtained by carrying out a specified test method
[ISO 3534-2, 3.4.1]

EDITORIAL NOTE: Also used as English translation of the German term "Prüfergebnis" (standardized English translation missing).

test statistic *(Ger. Prüfgröße)*
statistic used in conjunction with a **statistical test** [ISO 3534-1, 1.52]

test system *(Ger. Prüfeinrichtung, Prüfsystem)*
complete set of equipment used to carry out a **test**

*EDITORIAL NOTE: Standardized definition of term unavailable; definition in analogy to the term "**measuring system**" [DIN 1319-1, 4.2].*

true quantity value *(Ger. wahrer Wert einer Größe)*
quantity value consistent with the definition of a **quantity** [VIM, 2.11]

true value *(Ger. wahrer Wert)*
value which characterizes a **quantity** or quantitative **characteristic** perfectly defined in the conditions which exist when that quantity or quantitative characteristic is considered

NOTE 1: The true value of a quantity or a quantitative characteristic is a theoretical concept and, in general, cannot be known exactly.

[ISO 3534-2, 3.2.5]

unit *(Ger. Auswahleinheit)*: see **sampling unit** [ISO 3534-2, 1.2.14]

EDITORIAL NOTE: Not to be confused with "measurement unit" (cf. [VIM, 1.9]).

unusual sequence of points *(Ger. ungewöhnliche Punktefolge)*
measurement results or statistical values which show a statistically improbable behavior if plotted versus time in chronological order

EDITORIAL NOTE: Standardized definition of term unavailable.

Literature

[AIAG MSA] AIAG Core Tools, Measurement Systems Analysis (MSA), 4th edition (2010)

[AIAG SPC] AIAG Core Tools, Statistical Process Control (SPC), 2nd edition (2005)

[Booklet 3] Series Quality Management in the Bosch Group, Technical Statistics, Booklet No. 3, Evaluation of Measurement Series

[Booklet 8] Series Quality Management in the Bosch Group, Technical Statistics, Booklet No. 8, Measurement Uncertainty

[Booklet 9] Series Quality Management in the Bosch Group, Technical Statistics, Booklet No. 9, Machine and Process Capability

[CDQ 0301] CDQ 0301, Management of Characteristics
(Central directive, exclusively available in-house RB)

[CDQ 1001] CDQ 1001, Control of Inspection, Measuring and Test Equipment
(Central directive, exclusively available in-house RB)

[EQT] Elementary Quality Assurance Tools, Booklet, Robert Bosch GmbH (2005)

[DIN 1319-1] DIN 1319-1, Fundamentals of metrology, Part 1, Basic terminology

[DIN 1319-2] DIN 1319-2, Fundamentals of metrology, Part 2, Terminology related to measuring equipment

[Fleiss] J. L. Fleiss, Measuring Nominal Scale Agreement Among Many Raters, Psychological Bulletin (1971), Vol. 76, No. 5, pp. 378 – 382

[Fleiss-2] J. L. Fleiss, B. Levin, M. C. Paik, Statistical Methods for Rates and Proportions, 3rd edition (2003), Wiley-Interscience, John Wiley & Sons Inc., ISBN 0-471-52629-0: Chpt. 18, The Measurement of Interrater Agreement, pp. 598 – 626

[GUM] ENV 13005, Guide to the expression of uncertainty in measurement

[ISO 10012] ISO 10012, Measurement management systems – Requirements for measurement processes and measuring equipment

[ISO 14253] ISO 14253-1, Geometrical Product Specification (GPS) – Inspection by measurement of workpieces and measuring equipment – Part 1: Decision rules for proving conformance or non-conformance with specifications

[ISO 14468] ISO/TR 14468, Selected illustrations of attribute agreement analysis

[ISO 3534-1] ISO 3534-1, Statistics – Vocabulary and symbols – Part 1: General statistical terms and terms used in probability

[ISO 3534-2] ISO 3534-2, Statistics – Vocabulary and symbols – Part 2: Applied statistics

[ISO 9000] DIN EN ISO 9000, Quality management systems — Fundamentals and vocabulary (ISO 9000)

[MM] Management System Manual for Quality, Environment, Safety and Security in the Bosch Group *(Central directive, exclusively available in-house RB)*

[Sachs] J. Hedderich, L. Sachs, angewandte Statistik, Springer Verlag Berlin, 16. Auflage, 2018

[VDA 5] VDA, Quality Management in the Automotive Industry, Volume 5, Suitability of Inspection, Measuring and Test Processes, German Association of the Automotive Industry (VDA), ISSN 0943-9412

[VDI 2618] VDI/VDE/DGQ Guideline 2618, Testing instructions for Checking Gauges and Measuring Tools

[VDI 2622] VDI/VDE/DGQ/DKD Guideline 2622, Calibration of measuring equipment for electrical quantities

[VIM] International Vocabulary of Metrology (VIM), German-English Version, ISO/IEC-Guide 99, Editor: DIN Deutsches Institut für Normung, Beuth Verlag Berlin Wien Zürich, ISBN 978-3-410-23472-3

[Wheeler] D. J. Wheeler, R. W. Lyday, Evaluating the Measurement Process, 1st edition (1984), SPC Press Inc.

Index

S

T

U

V

X

2020-04-06 - SOCOS

Robert Bosch GmbH
C/QMM Tilsch
Wiener Strasse 42 - 46
70469 Stuttgart
Germany

Tel. +49 711 811-0
www.bosch.com

Page intentionally left blank

2020-04-06 - SOCOS

Robert Bosch GmbH
C/QMM Tilsch
Wiener Strasse 42 - 46
70469 Stuttgart
Germany
Phone +49 711 811 - 0
www.bosch.com