TECHNICAL REPORT

ISO/TR 12888

First edition
2011-08-01

Selected illustrations of gauge repeatability and reproducibility studies

Illustrations choisies d'études de répétabilité et de reproductibilité par calibre

Reference number
ISO/TR 12888:2011(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 12888 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 7, *Applications of statistical and related techniques for the implementation of Six Sigma*.

Introduction

The Six Sigma [1] and international statistical standards communities share a philosophy of continuous improvement and many analytical tools. The Six Sigma community tends to adopt a pragmatic approach driven by time and resource constraints. The statistical standards community arrives at rigorous documents through long-term international consensus. The disparities in time pressures, mathematical rigor, and statistical software usage have inhibited exchanges, synergy, and mutual appreciation between the two groups.

The present document takes one specific statistical tool (gauge repeatability and reproducibility, also known as GRR), develops the topic somewhat generically (in the spirit of International Standards), then illustrates it through the use of four detailed and distinct applications. The generic description focuses on the commonalities across studies designed to assess the variability of testing equipment and measurement systems. The annexes presenting four illustrations follow the basic framework but also identify the nuances and peculiarities in the specific applications. Each illustration offers at least one “wrinkle” to the problem, which is generally the case for real Six Sigma applications. It is thus hoped that practitioners can identify with at least one of the four illustrations, if only to remind them of the basic material on GRR that was encountered during their Six Sigma training. Each of the four illustrations is developed and analysed using statistical software of current vintage. The explanations throughout are devoid of mathematical detail — such material can be readily obtained from many references on GRR (such as those given in the Bibliography).

1) Six Sigma is a trade mark of Motorola, Inc.

Selected illustrations of gauge repeatability and reproducibility studies

1 Scope

This Technical Report describes the measurement process where the characteristic(s) being measured is a continuous variable. Measurement processes where the characteristic(s) of interest is an attribute (i.e. pass/fail) are not treated in this document.

This Technical Report provides examples of simple measurement systems and gives usable results as used in industry where there are two major factors contributing to the variation of the measurement results, such as variation between operators or appraisers and within operators or appraisers.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1
measurement system
collection of operations, procedures, devices and other equipment, software, and personnel used to assign a value to the characteristic being measured

NOTE This includes the complete process used to obtain measurements.

[IWA 1:2005, 3.1.9]

2.2
discrimination
ability of the measurement system to identify the infinitesimal change of the characteristic being measured

2.3
precision
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 3534-2:2006, 3.3.4]

2.4
repeatability
precision under repeatability conditions

NOTE Repeatability can be expressed quantitatively in terms of the dispersion characteristics of the results.

[ISO 3534-2:2006, 3.3.5]

2.5
repeatability conditions
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 3534-2:2006, 3.3.6]

2.6
gauge reproducibility
reproducibility which represents the variation that occurs when different appraisers measure the same part with the same equipment

NOTE 1 This term should be used only in a GRR (gauge repeatability and reproducibility) study.

NOTE 2 This definition for reproducibility differs from those in ISO 3534-2, ISO 5725-1 and ISO/IEC Guide 99. The definition is that used in the software related to GRR calculation and other industry standards.

NOTE 3 The computer software output in the annexes given as “reproducibility” means “gauge reproducibility” as defined here.

3 Symbols and abbreviated terms

The symbols and abbreviated terms used in this Technical Report are as follows:

ANOVA analysis of variance

DF degree of freedom

DOE design of experiments

F F-test statistic (coefficient of determination)

GRR gauge repeatability and reproducibility

MS mean of squares

MSA measurement system analysis

NDC number of distinct categories

P p-value (probability of obtaining a test statistic)

REML restricted estimation maximum likelihood

RF reference figure

SD standard deviation

SS sum of squares

SV study variation

%P/T precision to tolerance ratio, in percent

%R&R repeatability and reproducibility, in percent relative to the reference figure

R reproducibility

U upper specification limit

L lower specification limit

σ standard deviation

σ_{MS} standard deviation of measurement system

σ_r repeatability standard deviation

σ_R reproducibility standard deviation

σ_{P} standard deviation of the manufacturing process without measurement error

4 Generic description of GRR studies

4.1 Overview of the structure of GRR studies

This Technical Report provides general guidelines on the design, conduct and analysis of GRR studies and illustrates the steps with four distinct applications given in Annexes A through D. Each of these four examples follows the basic structure given in Table 1.

Table 1 — Basic steps in GRR studies

1	State the overall objectives of GRR
2	Describe the measurement process
3	Select a GRR studies method
4	Design a sampling plan for GRR studies
5	Analyse the result
6	Provide a conclusion with suggestions

The steps given in Table 1 apply to the design and analysis of GRR in general, although this Technical Report focuses on two-factor GRR studies. Each of the six steps is explained in general in 4.2 to 4.7. Specific explanations of the substance of these steps are provided in the examples in Annexes A through D.

4.2 Overall objectives of GRR

GRR studies are often used in Six Sigma projects. The primary motivation for GRR studies should be clearly stated and agreed upon by all parties involved in the design, analysis and implications of the GRR studies effort. The main purpose of GRR studies is to identify the capability of a measurement system and to judge whether it is acceptable for a given monitored process. GRR studies determine how much of the observed process variation is due to measurement system variation.

GRR studies are conducted for a variety of reasons, which include but are not limited to the following conditions:

a) the measurement system exhibits big variation under normal maintenance conditions;

b) the measurement equipment has been modified or upgraded, such as replacement of important part(s);

c) the measurement equipment or measurement system is new;

d) different measurement systems need to be compared;

e) GRR studies are required by quality management standards such as ISO/TS 16949.

4.3 Measurement process description

This Technical Report focuses on measurement process where the characteristic(s) being measured is a continuous variable. Measurement processes where the characteristic(s) of interest is an attribute (i.e. pass/fail) are not treated in this document.

The measurement process should be clearly described before conducting GRR, including the name of the equipment, its resolution, the quality characteristic(s) to be measured, measurement conditions, etc. Instruments should be properly calibrated. For information, see ISO 10012.

If necessary, measurement process mapping may be required to identify factors that may affect the observations. There may be many possible factors identified; however, only two-factor examples are provided in this document. In some cases, there may be multiple characteristics of interest to be measured. For the purpose of this document, only a single variable quality characteristic is considered in each example.

4.4 GRR studies methodology

In typical GRR studies, different appraisers (operators) are often used to capture variability within a given measurement system because in many cases the appraiser is a significant factor affecting the measurement data. However, in automated measurement processes appraisers are not involved in the measurement process. In such situations, changing fixtures or software or calibrating the equipment may be considered as changing the measurement system and potentially impacting its reproducibility.

For two-factor cases, the data collection model for GRR can be either a crossed or a nested design. Crossed design is similar to a full factorial design in DOE. The same subgroup of parts are measured by all appraisers for one round and then they are measured again (in a second round or more). If the subgroup size of parts (generally 10 to 20) is n, the number of the appraisers (at least 2) is a, the number of rounds (how many times one appraiser repeats measurement, at least 2) is b, and the total data set is $n \times a \times b$. Nested design also yields an $n \times a \times b$ data set, but the difference is that different appraisers measure different subgroups of parts of same subgroup size with repetitions. Each subgroup cannot be measured by another appraiser. That is, subgroups of parts are nested within appraisers. Crossed design and nested design are illustrated in Figures 1 and 2, respectively.

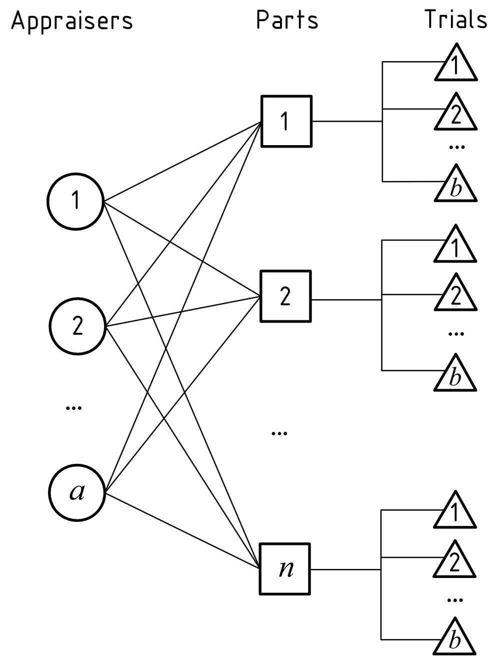


Figure 1 — Crossed design in GRR studies

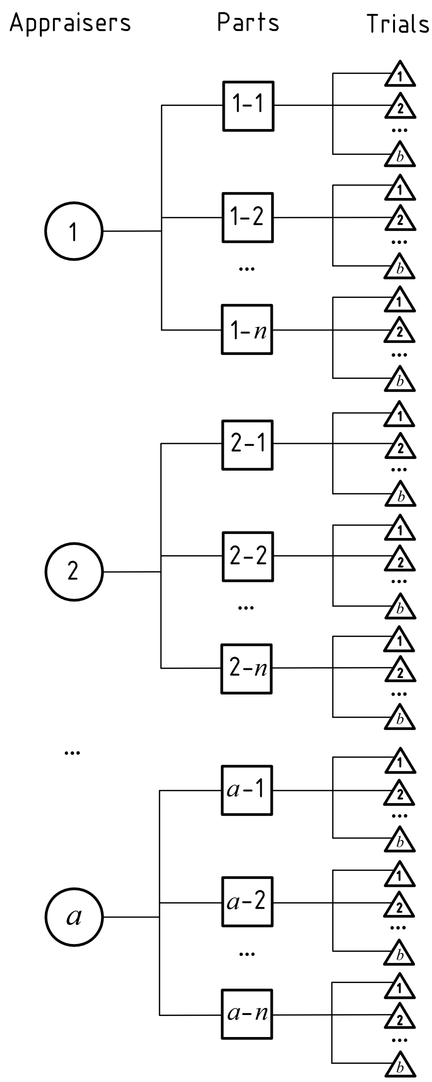


Figure 2 — Nested design in GRR studies

Crossed design assumes that the parts being measured by appraisers are undamaged and can be measured repeatedly during the measurement process. However, in some conditions, once the measurement is obtained for a particular part, that part is no longer available for additional measurements with the same appraiser or different appraisers; it is then appropriate to accept nested design. For destructive measurement, if homogenous samples are available, nested design may be a good choice.

To estimate repeatability and reproducibility, several methods can be used, as illustrated in commercial statistical software packages. The three most commonly encountered methods are described below and will be illustrated in the annexes.

The range method is based on the estimation of repeatability standard deviation (σ_r), using the range of multiple observations of one appraiser measuring the same part with the same equipment, and of reproducibility standard deviation (σ_R), using the difference of the average of different appraisers. The ANOVA method is based on the estimation of repeatability and reproducibility standard deviation using variance component analysis. For a two-factor crossed design, the advantage of ANOVA is that it can estimate the interaction between appraisers and parts. Many commercial statistical software packages provide alternatives for the two methods. The REML method estimates repeatability and reproducibility by optimizing the likelihood of the observations. This more sophisticated method is useful when other methods lead to negative estimates of variance components.

In this Technical Report,

— σ_{MS} denotes the standard deviation of measurement system error, where σ_{MS} is the square root of the sum of σ_r^2 and σ_R^2,

— $6\sigma_{\text{MS}}$ (some companies use $5{,}15\sigma_{\text{MS}}$) is the value of GRR (precision),

— σ_{P} is the standard deviation of the manufacturing process without measurement error.

Thus, the observed total variance is the sum of σ_{P}^2 and σ_{MS}^2. In practice, two indicators are used to measure GRR relative to process spread and tolerance, %R&R and %P/T, where

$$\%\text{R\&R} = \frac{\sigma_{\text{MS}}}{\sqrt{\sigma_{\text{MS}}^2 + \sigma_{\text{P}}^2}} \times 100\ \%$$

$$\%\text{P/T} = \frac{6\sigma_{\text{MS}}}{U - L} \times 100\ \%$$

4.5 Sampling plan for GRR studies

The sampling plan is very important for GRR studies. Poor design can lead to a situation where the true variation in the measurement process is underestimated or overestimated, and this will result in an overly optimistic or pessimistic conclusion regarding measurement system capability.

Different designs adopt different tables for collecting measurement data. Tables 2 and 3 provide templates for the basic layout for crossed and nested design respectively with 3 operators, 3 repetitions and 10 items measured by each operator. The main difference between the two layouts is the “Item No.” column (representing the parts being measured). For crossed design, three appraisers share the same “Item No.” column, which means the same part subgroup is measured by different appraisers. However, for nested design, there is a different “Item No.” column for each appraiser, which means one part is measured by only one appraiser.

Table 2 — Layout of a generic crossed GRR design

Item No.	GRR studies											
	Appraiser A				Appraiser B				Appraiser C			
	Trial 1	Trial 2	Trial 3	Range	Trial 1	Trial 2	Trial 3	Range	Trial 1	Trial 2	Trial 3	Range
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												

Table 3 — Layout of a generic nested GRR design

GRR studies														
Appraiser A					Appraiser B					Appraiser C				
Item No.	Trial 1	Trial 2	Trial 3	Range	Item No.	Trial 1	Trial 2	Trial 3	Range	Item No.	Trial 1	Trial 2	Trial 3	Range
A1					B1					C1				
A2					B2					C2				
A3					B3					C3				
A4					B4					C4				
A5					B5					C5				
A6					B6					C6				
A7					B7					C7				
A8					B8					C8				
A9					B9					C9				
A10					B10					C10				

In the sampling plan for GRR studies, the subgroup size of parts, the number of the appraisers and the number of rounds should be determined. Generally speaking, three to five appraisers are selected to measure more than ten parts with two or three trials. Note that the selected samples must come from the production process and represent the entire production variance. (In situations where it is difficult to get ten parts or more, although GRR can be estimated with few parts, the uncertainty of part variability can be large and thus %R&R can be unreliable. In this case, if the process standard deviation is known, it is strongly recommended to use the known standard deviation instead of using the process standard deviation estimated from few samples.)

In the process of measurement for GRR, randomization is a very important consideration. Randomization means that the parts should be measured by the operator in a random order. During the experiment, the Hawthorne effect should be avoided because appraisers with a higher degree of attention may lead to a poor estimation of the measurement process variation.

4.6 Data analysis (numerically and graphically)

After choosing the most appropriate sampling plan for the experiment and collecting the measurement values, an appropriate analysis method needs to be selected for the interpretation of the results. Although the range method can be easily done in a spreadsheet and used to inspect whether outliers exist or not, this does not take into account the interaction between appraisers and parts, which results in the underestimation of measurement error. Hence, in this Technical Report, the analysis of variance (ANOVA) method and the restricted maximum likelihood (REML) method based on a variance component analysis are used to identify variance components for each possible variation. For GRR studies with two factors, variation factors include parts, appraisers, interaction between parts and appraisers, and repeatability (pure error). In the GRR report, the capability indicators of the measurement system (%R&R and %P/T) can be obtained directly from the analysis table.

In addition to the GRR report, GRR study graphics are useful. These can be used to visually identify the main variation sources. For a range chart by operator and part, if all ranges are randomly distributed within upper and lower control limits, the measurement process is in control. If the range chart signals that the process is out of control, special causes may exist. These should be identified and the corresponding part remeasured.

4.7 Conclusions and suggestions

4.7.1 The purpose of GRR studies is to determine if the variability of a measurement system is small relative to the variability of the monitored process.

Common guidelines such as MSA indicate that if the %R&R and %P/T are both less than 10 %, the measurement system is acceptable. If they are in the range 10 % to 30 %, the measurement system may be acceptable depending on the importance of the application, cost of the gauge, cost of repairs, etc. If one or two indicators are over 30 %, the measurement system is considered inadequate and immediate corrective actions are called for.

For more information on acceptability, see MSA or the best practices available in different industries (chemical, automotive, financial, electronics, etc.).

Either high repeatability or high reproducibility can result in high %R&R.

4.7.2 If the repeatability is high, follow the following steps.

— Go back to the range chart by appraiser to ensure that there is no point out of upper and lower control limits.

— Check whether there is an excessive within-part variation by studying the profile of the part and collecting appropriate data.

— Check whether the gauge is sufficiently rigid.

— Check whether the location at which the measurement is to be taken is clearly defined and understood properly by the appraisers. If it is not, this may result in high reproducibility as well.

— Check whether the instrument requires maintenance.

— Check whether a fixture of some sort is needed to help the appraiser use the gauge more consistently.

• If none of the above reasons are found to be valid, brainstorm and find out whether the gauge is suitable for the intended measurement. Unsuitable gauges may also lead to high repeatability.

4.7.3 If the reproducibility is high, the following possible actions are recommended.

— Find out whether all the appraisers are adequately trained in the measurement method.

— Check whether or not the calibrations on the gauge dial are clear.

— Check whether or not the location at which the measurement to be taken is clearly defined and understood properly by the appraisers.

— Check whether or not a fixture of some sort is needed to help the appraiser use the gauge more consistently.

— Check whether or not two or more key elements within the measurement system work under the same conditions.

— If none of the above reasons are found to be valid, brainstorm and find out whether the gauge is suitable for the intended measurement.

After certain actions have been taken for improving the measurement system, GRR studies should be done again to validate whether the improved measurement system is acceptable or not.

5 Description of Annexes A through D

5.1 Comparison of the examples

Four distinct examples of GRR studies are illustrated in Annexes A to D. Each of these examples follows the same general template given in Table 1 and follows a version of the standard design given in Tables 2 and 3.

5.2 Example summaries

Table 4 summarizes the examples detailed in the annexes and indicates aspects of the analyses that were unique to that experiment.

Table 4 — Example summaries found by annex

Annex	Experiment	GRR-specific aspects
A	Cell phones	This example includes characterization of a testing process (not a piece of equipment); automation of equipment (i.e. no operator effect); multiple responses; and crossed, balanced design.
B	Load cell	In this GRR study, the gauge is a load cell used to measure the metal cap placement force in the assembly of a semiconductor device. Range and ANOVA methods of analysis are used for comparison. The traditional crossed study design is used.
C	Motor shaft radial run-out of the axis	A motor manufacturer produces motors for air conditioners. Dial indicators are used to measure the radial run-out of the axis. 2-factor crossed case.
D	Rip-off force of A/B cover of charges	In order to improve the process capability of rip-off force of chargers or precisely determine if chargers are good or bad due to the rip-off force, a destructive measurement process with one rip-off test machine. The REML method is used.

Annex A
(informative)

GRR for automated testing of RF performance of cell phones

A.1 Measurement data type

Multiple responses (95 responses), all responses are continuous (variable data type).

One response, slope P1-RX1, was analysed completely. Slope P1-RX1 is one of many standardized audio-radio frequency measurements made on cell phones to ensure the audio capability of the handsets. These measurements are usually conducted after the final assembly of the phone and involve verifying the audio stimulus signal over the complete radio frequency (RF) transmit-receive path. They fall into categories such as output power, frequency response, total harmonic distortion, and many others.

Results on all the other variables are summarized.

A.2 Gauge used for measurement

Gauge name: Bank of automated radio performance testers

On each cell phone, 95 radio performance parameters are measured automatically, with different resolutions and different tolerances depending on the parameter being measured.

A.3 Measurement process description

Cell phones, once manufactured, are tested for their radio performance before moving on to a configuration line where software and data (specific to telecommunication carriers that are the service providers for the radios) are loaded. Then the radios are packed and sent to storage or shipped to customers. Banks of automated testing equipment called "testers" are installed at the end of manufacturing lines to measure performance characteristics of the cell phones, in particular RF test parameters. To ensure that any cell phone can be tested by any tester without impacting on the results, a GRR study was undertaken.

Figure A.1 describes the testing process. All cell phones coming out of the manufacturing line are tested individually by any one of the testers making up the bank of testers, depending on availability of the testers. An operator supervises the bank of testers as units are loaded automatically. Measurements taken by each tester are automatically stored in the computer for that bank and a "pass" or "fail" signal is displayed and registered for each cell phone depending on the results. A cell phone is declared "failed" if any of the 95 parameters measured automatically on that cell phone does not meet its specifications. A "failed" cell phone cannot proceed to the configuration line until it has been duly repaired and re-tested successfully.

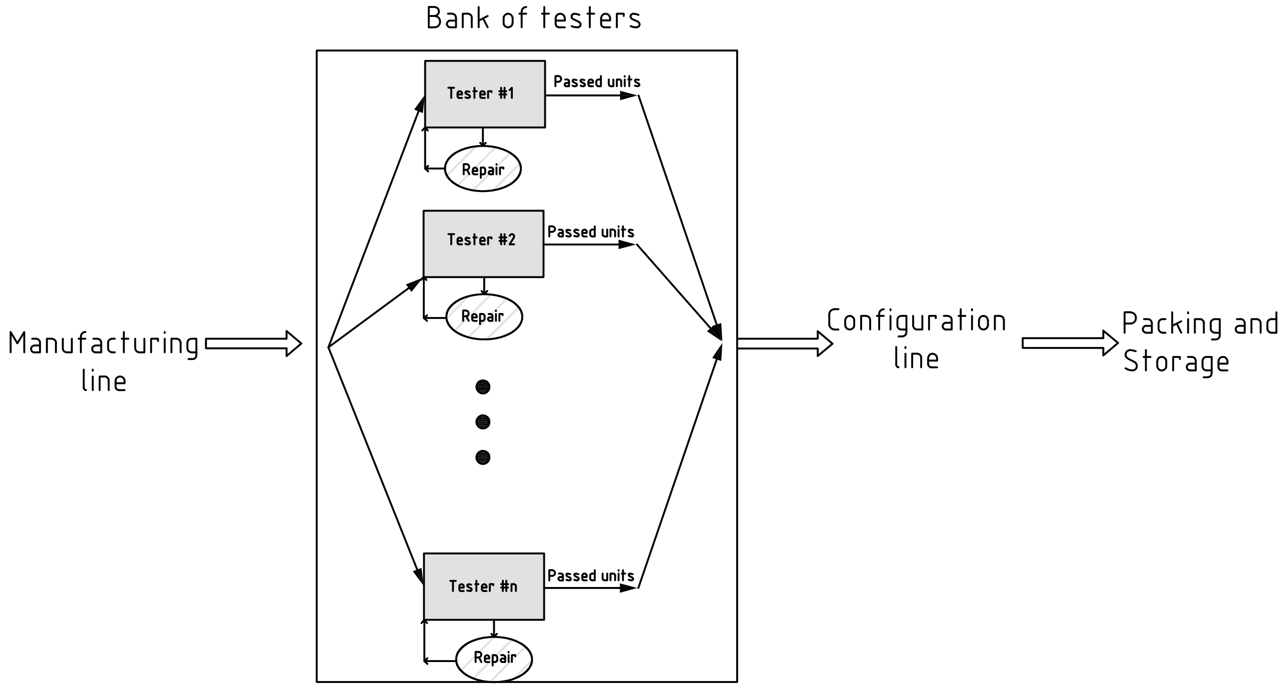


Figure A.1 — Testing process for cell phones

A.4 Possible sources of measurement system variation

Variability in the measurements comes from two sources:

— inherent variability within a given tester;

— slight differences in the hardware or manufacturing of the testers may have an impact on the measurements taken by the testers on the units, as well as the calibration process of the testers.

Therefore, the sources of possible variations for this measurement system are: testers, cell phones, pure error generated by repeated measurements, and interaction effects between testers and cell phones.

A.5 Sampling plan

The factors chosen to be varied within the experiment were determined by the sources of variability identified above. These were

— A: the variation within a given tester;

— B: the testers themselves (also called appraisers).

Three cell phones were selected randomly from the population of units manufactured the previous day. The cell phones were considered random units from a large population of similar devices. They are referred to as units #1, #2 and #3 in this example.

These three units constitute a small sample size of a population of units, which means that caution should be exercised in the interpretation of the results and in drawing a conclusion.

Four testers were randomly selected from the bank of ten testers. They are referred to as TNS 080, TNS 082, TNS 083 and TNS 088. Three replicate measurements were taken on each parameter for each cell phone.

The factors used in the experiment and their associated levels are given in Table A.1.

Table A.1 — Factors and associated levels

Factor	Random/fixed	Number of levels
A: Units	random	3
B: Testers	random	4

This experiment is a typical 2-factor crossed experiment for a GRR study. The model selected is a crossed design, where each unit is measured by each tester several times. The design is balanced, with each unit being measured the same number of times by each tester.

The number of repetitions to estimate the repeatability of this testing bank was set at three.

This led to 36 series of measurements for each of the 95 parameters measured by the testers (3 units × 4 testers × 3 repetitions).

A.6 Measurement data for one parameter

The analysis focuses on a single parameter, namely the slope P1-RX1. That parameter has a target of 14,5 dB ± 1dB. The measurement data are shown in Table A.2. They were keyed into a software program called Minitab 15™[2)].

Table A.2 — Results of the experiment

Tester	Unit slope	P1-RX1	Tester	Unit slope	P1-RX1
TNS 080	1	14,230	TNS 082	1	14,570
TNS 080	1	14,000	TNS 082	1	14,800
TNS 080	1	13,820	TNS 082	1	14,960
TNS 080	2	14,130	TNS 082	2	15,350
TNS 080	2	14,590	TNS 082	2	15,310
TNS 080	2	14,200	TNS 082	2	15,030
TNS 080	3	14,180	TNS 082	3	14,960
TNS 080	3	14,295	TNS 082	3	15,055
TNS 080	3	14,010	TNS 082	3	14,995
TNS 083	1	14,620	TNS 084	1	14,720
TNS 083	1	14,520	TNS 084	1	14,460
TNS 083	1	14,410	TNS 084	1	14,460
TNS 083	2	15,130	TNS 084	2	15,130
TNS 083	2	14,990	TNS 084	2	14,940
TNS 083	2	14,960	TNS 084	2	14,830
TNS 083	3	14,875	TNS 084	3	14,925
TNS 083	3	14,755	TNS 084	3	14,700
TNS 083	3	14,685	TNS 084	3	14,625

2) Minitab is the trade name of a product supplied by Minitab, Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

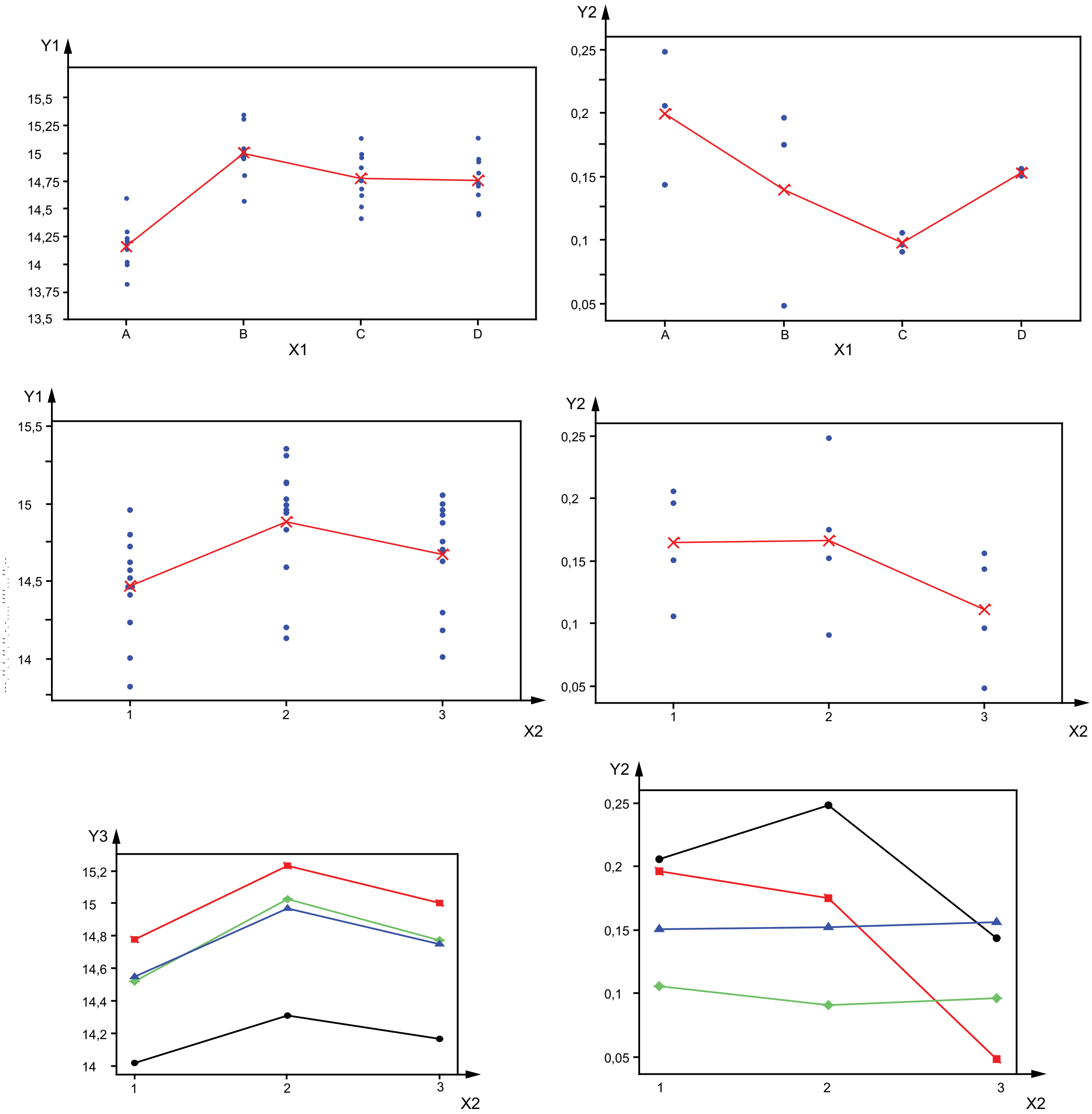
A.7 Statistical method used for GRR studies

Method: ANOVA

A.8 Statistical analysis

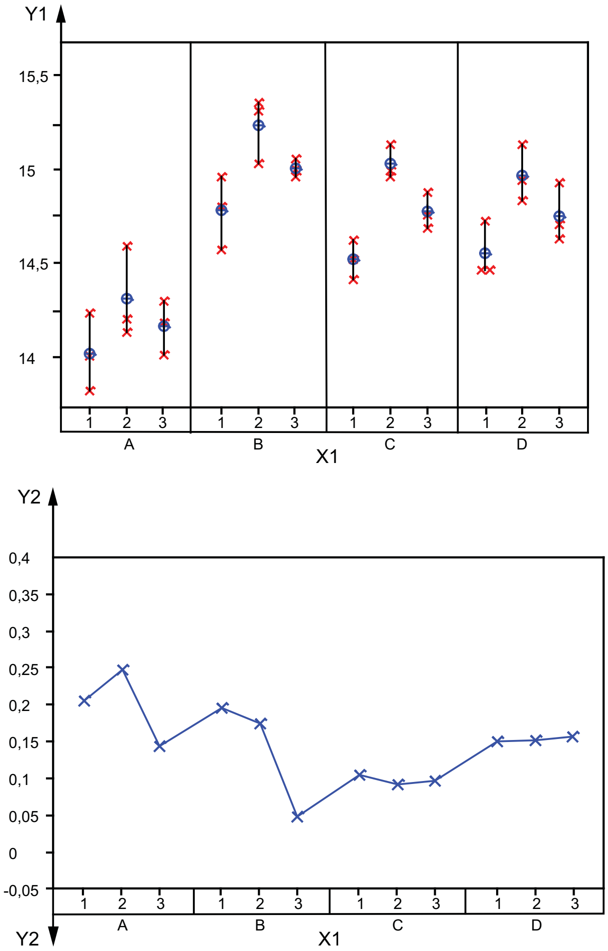
Figure A.2 and Table A.3 show that the major source of variation for this measurement system is reproducibility.

Software outputs from the Minitab 15 ANOVA method are shown below.



a) GRR mean plots **b) GRR standard deviation plots**

Figure A.2 — GRR graphical results for slope P1-RX1; 14,5 ± 1,0 dB max. (*continued*)



c) Variability gauge

Key

X1 tester
X2 unit slope
Y1 slope P1-RX1
Y2 standard deviation
Y3 mean
A TNS 080
B TNS 082
C TNS 083
D TNS 084

Tester

TNS 080
TNS 082
TNS 083
TNS 084

Figure A.2 — GRR graphical results for slope P1-RX1; 14,5 ± 1,0 dB max.

Table A.3 — Two-way ANOVA table with interaction

Source	DF	SS	MS	F	P
Unit	2	1,050 02	0,525 01	80,461	0,000
Station	3	3,485 85	1 161 95	178,077	0,000
Unit × Station	6	0,039 15	0,006 52	0,269	0,946
Repeatability	24	0,582 88	0,024 29		
Total	35	5,157 90			
Alpha to remove interaction term = 0,25					

As shown in Table A.4, the measurement system variance takes up 77,83 % in the total variance. Repeatability takes up 10,94 %, while reproducibility takes up 66,89 %, and the interaction between the tester and the unit is non-significant.

From Table A.5, it can be seen that for this measurement system, %R&R = 88 % and NDC = 1.

Table A.4 — GRR table

Source	VarComp	% Contribution (of VarComp)
Total GRR	0,147 536	77,83
Repeatability	0,020 734	10,94
Reproducibility	0,126 802	66,89
Station	0,126 802	66,89
Part-To-Part	0,042 023	22,17
Total Variation	0,189 559	100,00

Table A.5 — GRR statistics

Process tolerance = 2				
Source	**StdDev (SD)**	**Study Var (6 × SD)**	**% Study Var**	**% Tolerance (SV/Toler)**
Total GRR	0,384 104	2,304 63	88,22	115,23
Repeatability	0,143 995	0,863 97	33,07	43,20
Reproducibility	0,356 092	2,136 55	81,79	106,83
Station	0,356 092	2,136 55	81,79	106,83
Part-To-Part	0,204 995	1,229 97	47,08	61,50
Total Variation	0,435 384	2,612 30	100,00	130,62
NDC = 1				

A.9 Conclusions

%R&R = 88 %, which is greatly above the commonly used 30 % threshold. Because the sample size of units is very small (3), the part variation estimate of Table A.4 has a large uncertainty associated with it, thus potentially impacting notably the total variability and the GRR ratio in either direction (higher or lower). Yet, looking at the %P/T (115 %) also indicates that the measurement variability of the testers is too large with respect to the tolerance of 2 dB associated with that parameter. So this measurement system needs to be improved with respect to the parameter “slope P1-RX1”.

A.10 Improvement actions or suggestions

Because the major source of variation of the GRR in this example is reproducibility error, this implies that the testers are likely to be different. They need to be calibrated to bring the reproducibility to an acceptable number.

Once actions have been taken to calibrate the testers, another GRR needs to be run. A recommendation is made to use more than three units in the study and to target a minimum of ten units to ensure a more robust estimate of the part variability and thus of the total variability.

A.11 Analysis of all the parameters measured

Detailed results of the GRR analysis on one of the 95 parameters measured by the automatic testers were collected. The univariate GRR analysis indicated that the testers needed to be calibrated for the parameter selected, namely the slope P1-RX1.

Conducting univariate GRR analyses on each of the 95 parameters gives the results for %R&R and %P/T ratio shown in Table A.6. Missing data in the %P/T column are indicative of a one-sided specification. Missing data in the GRR ratio indicate no variability at all in all measurements of the three units.

Table A.6 — Gauge statistics on a subset of 30 parameters

Parameter	%R&R	%P/T
Power A, Vdc[a] Test Point; 23,20 to 26,10 Vdc.	2,4	2,7
Power B, Vdc[a] Test Point; 23,20 to 26,10 Vdc.	2,4	2,7
Power C, Vdc[a] Test Point; 23,20 to 26,10 Vdc.	2,4	2,7
RX1 Optical TP[b] Voltage; 0,552 ± 0,002 Vdc.	0	0
Flatness P1-RX1; ±0,70 dB max.	100,0	159,9
Gain P1-RX1; 56,2 dB min.	36,4	
Slope P1-RX1; 14,5 dB ± 1,0 dB max.	88,2	115,3
Test Point Loss Fwd[c] P1; −20 dB ± 0,5 dB max.	84,8	27,3
Return Loss Fwd P1; −16 dB max.	37,7	
Return Loss Rtn[d] P1; −16 dB max.	86,1	
Flatness P3-RX1; ±0,70 dB max.	99,2	143,3
Gain P3-RX1; 56,2 dB min.	26,3	
Slope P3-RX1; 14,5 dB ± 1,0 dB max.	86,6	107,3
Test Point Loss Fwd P3; −20 dB ± 0,5 dB max.	93,5	151,6
Return Loss Fwd[c] P3; −16 dB max.	100,0	
Return Loss Rtn[d] P3; −16 dB max.	19,9	
Flatness P4-RX1; ±0,70 dB max.	100	171,0
Gain P4-RX1; 56,2 dB min.	34,8	
Slope P4-RX1; 14,5 dB ± 1,0 dB max.	91,3	75,1
Test Point Loss Fwd[c] P4; −20 dB ± 0,5 dB max.	69,6	131,4
Return Loss Fwd[c] P4; −16 dB max.	100,0	
Return Loss Fwd[c] P4; −16 dB max.	68,3	
Flatness P6-RX1; ±0,70 dB max.	100,0	134,4
Gain P6-RX1; 56,2 dB min.	25,0	
Slope P6-RX1; 14,5 dB ± 1,0 dB max.	83,8	86,0
Test Point Loss Fwd[c] P6; −20 dB ± 0,5 dB max.	96,4	76,9
Return Loss Fwd[c] P6; −16 dB max.	38,6	
Return Loss Fwd[c] P6; −16 dB max.	8,5	
RX2 Optical TP[b] Voltage; 0,552 ± 0,002 Vdc		0
Flatness P1-RX2; ±0,70 dB max.	100,0	172,6

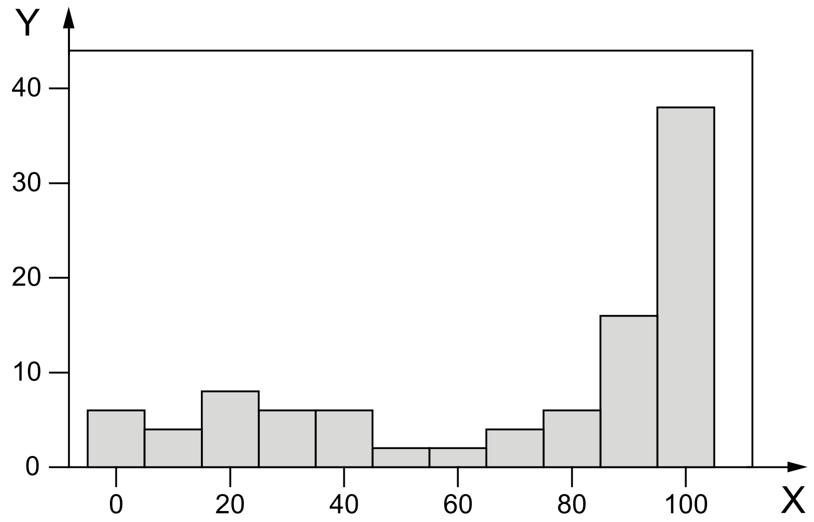
[a] Vdc = dc voltage.

[b] TP = test point.

[c] Fwd = forward link.

[d] Rtn = return link.

Figures A.3 and A.4 provide graphical representations of the distributions of %R&R and %P/T.

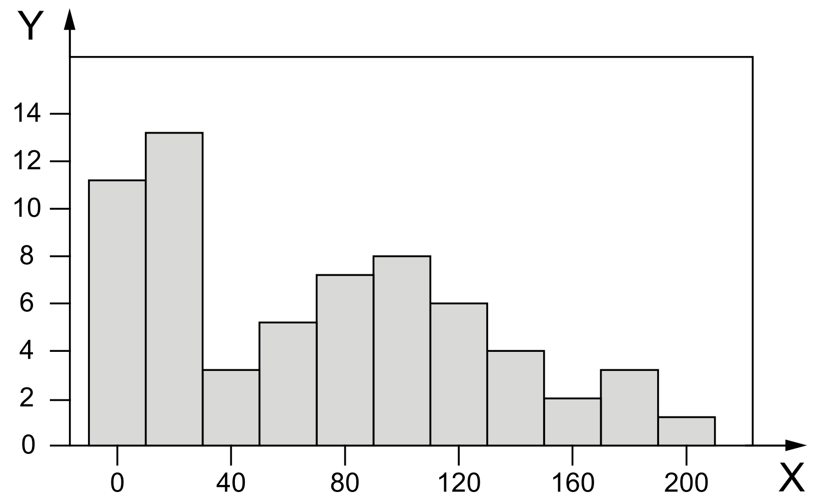


Key

X %R&R

Y frequency

Figure A.3 — Histogram of gauge %R&R across all 95 parameters measured



Key

X %P/T

Y frequency

Figure A.4 – Histogram of %P/T across all parameters measured with two-sided specifications

A.12 Conclusions pertinent to all parameters

Many parameters have quite high values for their %R&R or %P/T, or for both ratios.

Since there are only three units for this GRR study, the historical standard deviation of the manufacturing process, if available, should have been used for the calculations rather than the estimated variability from the three units from our sample. In this study, the manufacturing process was relatively new and that historical value was not available. One strong recommendation is to monitor the manufacturing process and capture that value, which will be available if needed in the follow-up GRR study.

Another recommendation to address the many deficiencies of the measurement system as identified by this study is to tackle the top worst characteristics from a Pareto chart and then run individual control charts using the long method (individual moving range) to identify potential problems such as stability, drift, etc.

It is also recommended to confirm calibration of the testers for all parameters and redo the GRR analysis with a minimum of ten units.

Annex B
(informative)

RF metal-ceramics semiconductor package assembly load cell GRR study

B.1 Measurement data type

Continuous (variable data type) – in gram-force (1 gf $\approx$ 9,81 $\times$ 10^{-3} N $\approx$ 9,81 mN)

B.2 Gauge used for measurement

Gauge name: Load cell, which is used to measure the force of applying the cap onto the base of the device.

Gauge resolution: smallest unit of measurement = 1 gram-force.

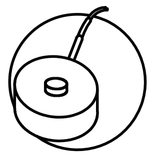


Figure B.1 — Example of a button-type compression load cell

B.3 Measurement process description

This is an example of a GRR study of a semiconductor package assembly, comparing the results obtained using the average-range method with the results from the ANOVA method. The example is from the cap and cure process, one of the assembly processes in radio-frequency (RF) metal-ceramics package assembly. This process is carried out to seal the device hermetically to protect it from moisture absorption, which could lead to corrosion and reliability issues.

The gauge in the measurement process being studied is a load cell, which is used to measure the force of applying the cap onto the base of the device. A load cell is a transducer which converts force into a measurable electrical output.

The measurement is done by placing the load cell in the work-holder of the cap placement machine as if it is an actual metal-ceramics package going through the capping process. The force applied to the load cell by the cap placement arm is then read from a digital display.

B.4 Possible sources of measurement system variation

The possible sources of variation of this measurement system are the following:

— the inherent variation of the cap placement machine at each force setting (in this study, the variation due to the cap placement machine is confounded with the variation of the measurement process and gauge);

— the placement position of the load cell by the operator;

— the inherent variation of the load cell (e.g. due to strain gauges in the load cell).

B.5 Sampling plan

In this GRR study, 10 different force settings were selected for the cap placement machine being studied, representing the distribution of force settings used in the cap and cure assembly process of various RF metal-ceramics packages.

Three production operators certified to run the cap and cure assembly process were selected randomly from three different shifts.

This experiment is a typical two-factor crossed experiment for a GRR study. The model selected is a crossed design, in which each force setting is measured three times by each operator using the load cell being studied. The measurements by operator were carried out in random order. The design is balanced, with each force setting being measured the same number of times by each operator. This led to a total of 90 gf measurements for this GRR study (10 force settings × 3 operators × 3 trials).

B.6 Measurement data for one parameter

The data collected is shown in Table B.1, the generic crossed GRR table.

Table B.1 — Layout of the crossed GRR design and experiment results

Measured values in gram-force

Item No.	Appraiser A					Appraiser B					Appraiser C				
	Trial 1	Trial 2	Trial 3	Average	Range	Trial 1	Trial 2	Trial 3	Average	Range	Trial 1	Trial 2	Trial 3	Average	Range
n	$x_{A;1}$	$x_{A;2}$	$x_{A;3}$	$\bar{x}_{gj}$	R_{gj}	$x_{B;1}$	$x_{B;2}$	$x_{B;3}$	$\bar{x}_{gj}$	R_{gj}	$x_{C;1}$	$x_{C;2}$	$x_{C;3}$	$\bar{x}_{gj}$	R_{gj}
1	373	375	374	374,0	2	374	370	372	372,0	4	378	380	381	379,7	3
2	391	388	389	389,3	3	385	389	387	387,0	4	392	398	396	395,3	6
3	339	340	342	340,3	3	348	346	350	348,0	4	340	342	343	341,7	3
4	363	365	367	365,0	4	365	368	363	365,3	5	360	365	359	361,3	6
5	401	403	404	402,7	3	402	400	399	400,3	3	405	400	406	403,7	6
6	460	457	459	458,7	3	463	460	465	462,7	5	468	466	464	466,0	4
7	330	332	332	331,3	2	328	327	330	328,3	3	335	336	333	334,7	3
8	414	412	415	413,7	3	411	413	408	410,7	5	410	408	413	410,3	5
9	426	424	428	426,0	4	428	425	424	425,7	4	427	430	428	428,3	3
10	454	453	457	454,7	4	443	447	445	445,0	4	450	453	452	451,7	3
$\bar{x}_A = 395{,}57$						$\bar{x}_B = 394{,}50$					$\bar{x}_C = 397{,}27$				
$\bar{R}_A = 3{,}1$						$\bar{R}_B = 4{,}1$					$\bar{R}_A = 4{,}2$				
Results: $\bar{x}_{diff} = 2{,}77$; $\bar{\bar{R}} = 3{,}8$															
NOTE $\bar{x}_{diff}$ denotes the maximum difference between the averages of several measurement series. $\bar{\bar{R}}$ denotes the average of average range.															

Methods: Both average-range and ANOVA for comparison.

B.7 Statistical analysis

B.7.1 General

Software outputs from Q-DAS solara® ME8 3) are shown below.

The specification tolerance (upper specification minus lower specification) is 160 gf.

The historical standard deviation of this cap placement process is 29,4 gf.

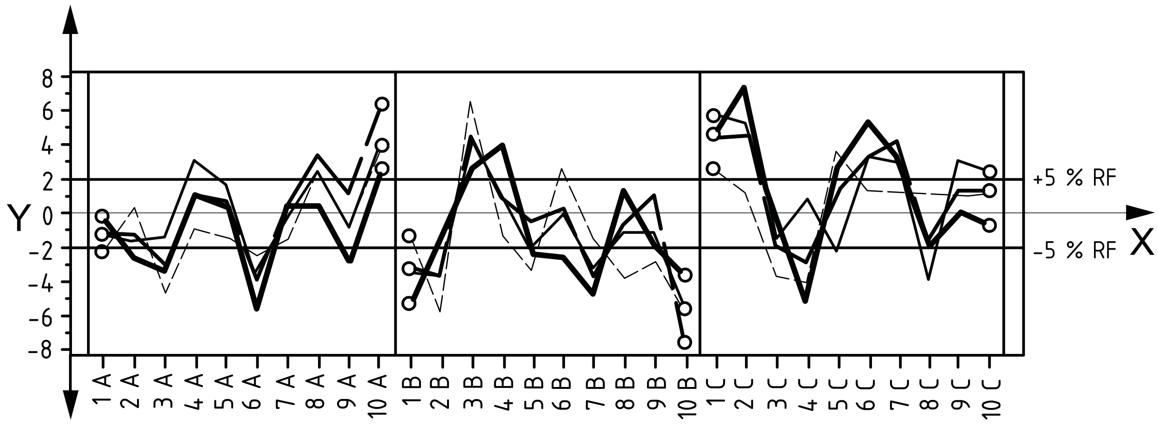
The criteria of acceptance are the following:

— below 10 %, the measurement system is capable;

— between 10 % and 30 %, the measurement system is marginally capable.

B.7.2 Average-range method

Graphical representations are shown in Figures B.2 and B.3.



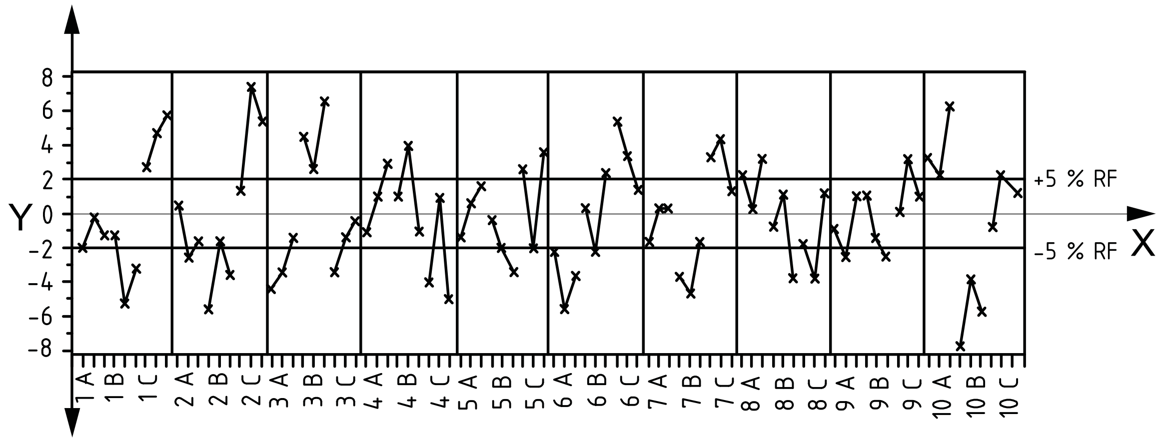
Key

X piece number/operator

Y capping pressure, in gram-force

Figure B.2 — Value chart by appraisers

3) Q-DAS solara® ME8 is a trade mark supplied by Q-DAS® GmbH & Co. KG. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.



Key

X piece number/operator

Y capping pressure, in gram-force

Figure B.3 — Value chart by items

Tables B.2, B.3 and B.4 show the GRR analysis outputs.

Table B.2 — GRR analysis table (Reference Variable: Total Variation)

Reference	Total Variation	TV[a] = 41,291 7	
Component	**Parameter**	**Standard deviation**	**Result**
Part Variation	K_3 = 0,314 6	PV[b] = $K_3 \times R_0$ = 41,207 3	%PV = 99,80 %
Repeatability	K_1 = 0,590 8	EV[c] = $K_1 \times \overline{\overline{R}}$ = 2,245 11	%EV = 5,44 %
Reproducibility	K_2 = 0,523 1	AV[d] = $\sqrt{(K_2 \times \overline{x}_{\text{diff}})^2 - \left[\text{EV}^2 / (n \times r)\right]}$ = 1,388 09 where n is the number of parts and r is the number of trials.	%AV = 3,36 %
GRR		$\sigma_{\text{MS}} = \sqrt{\text{EV}^2 + \text{AV}^2}$ = 2,639 56	%R&R = 6,39 %
Resolution		%RES[e] = 2,42 %	
Repeatability & Reproducibility		%R&R = 6,39 %	
Number of distinct categories		NDC = 22	
Measurement system capable (%R&R, NDC)			
Minimum reference figure for capable measuring system			$\text{TV}_{\text{min(\%R\&R)}}$ = 26,395 6

[a] TV = total variation.

[b] PV = part variation.

[c] EV = repeatability of the equipment's variation measurement system.

[d] AV = reproducibility/appraiser variation.

[e] %RES = resolution of the measurement system.

Table B.3 — GRR analysis table (Reference Variable: Tolerance)

Reference	Tolerance	T[a] = 160	
Component	**Parameter**	**Standard deviation**	**Result**
Part Variation	K_3 = 0,314 6	PV[b] = $K_3 \times R_0$ = 41,207 3	%PV = 154,53 %
Repeatability	K_1 = 0,590 8	EV[c] = $K_1 \times \bar{\bar{R}}$ = 2,245 11	%EV = 8,42 %
Reproducibility	K_2 = 0,523 1	AV[d] = $\sqrt{(K_2 \times \bar{x}_{\text{diff}})^2 - \left[\text{EV}^2 / (n \times r)\right]}$ = 1,388 09 where n is the number of parts and r is the number of trials.	%AV = 5,21 %
GRR		$\sigma_{\text{MS}} = \sqrt{\text{EV}^2 + \text{AV}^2}$ = 2,639 56	%R&R = 9,90 %
Resolution		%RES[e] = 0,63 %	
Repeatability & Reproducibility		%R&R = 9,90 %	
Number of distinct categories		NDC = 22	
Measurement system capable (%R&R, NDC)			
Minimum reference figure for capable measuring system			$T_{\min(\%\text{R\&R})}$ = 125,373

a T = total tolerance.

b PV = part variation.

c EV = repeatability of the equipment's variation measurement system.

d AV = reproducibility/appraiser variation.

e %RES = resolution of the measurement system.

Table B.4 — GRR analysis table (Reference Variable: Process Variation)

Reference	1 × Process Variation $1 \times \sigma_P$ = 29,400 0		
Component	**Parameter**	**Standard deviation**	**Result**
Part Variation	K_3 = 0,314 6	PV[a] = $K_3 \times R_0$ = 41,207 3	%PV = 140,16 %
Repeatability	K_1 = 0,590 8	EV[b] = $K_1 \times \overline{\overline{R}}$ = 2,245 11	%EV = 7,64 %
Reproducibility	K_2 = 0,523 1	AV[c] = $\sqrt{(K_2 \times \overline{x}_{\text{diff}})^2 - \left[\text{EV}^2 / (n \times r)\right]}$ = 1,388 09 where n is the number of parts and r is the number of trials.	%AV = 4,72 %
GRR		$\sigma_{MS} = \sqrt{\text{EV}^2 + \text{AV}^2}$ = 2,639 56	%R&R = 8,98 %
Resolution		%RES[d] = 3,40 %	
Repeatability & Reproducibility		%R&R = 8,98 %	
Number of distinct categories		NDC = 22	
Measurement system capable (%R&R, NDC)			
Minimum reference figure for capable measuring system			$1 \times \sigma_{P\ \min\ (\%R\&R)}$ = 26,395 6

[a] PV = part variation.

[b] EV = repeatability of the equipment's variation measurement system.

[c] AV = reproducibility/appraiser variation.

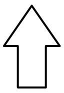
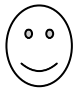
[d] %RES = resolution of the measurement system.

From the results above, the %R&R based on tolerance is calculated as 9,90 % and the %R&R based on process variation is 8,98 %. Based on the criteria for acceptance of below 10 %, the measurement system is deemed to be acceptable. Notice that the operator part interaction is not estimated in the reproducibility component. The absence of this interaction component is implicitly assumed when using the average-range method of analysis.

B.7.3 ANOVA method

For graphical representations, see the average-range method (B.7.2). Tables B.5, B.6 and B.7 show the GRR analysis outputs using the ANOVA method.

Table B.5 — GRR analysis table (Reference Variable: Total Variation)

Reference	Total variation		TV[a] = 43,624 7		
Component	**Variance**	**Standard deviation**	**95 % lower CI**	**95 % upper CI**	**Result**
Repeatability	4,077 78	EV[b] = 2,019 35	1,713 84	2,458 43	%EV = 4,63 %
Reproducibility	0,914 40	AV[c] = 0,956 24	0,000 00	8,704 20	%AV = 2,19 %
Interaction	8,967 08	IA[d] = 2,994 51	1,960 62	4,755 26	%IA = 6,68 %
GRR	13,959 3	σ_{MS} = 3,736 21	3,541 19	9,429 53	%R&R = 8,56 %
Resolution	%RES[e] = 2,29 %				
Repeatability & Reproducibility	%R&R = 8,56 %				
Number of distinct categories	NDC = 16				
	Measurement system capable (RES, %R&R)				
Minimum reference figure for capable measuring system					$TV_{min(R\&R)}$ = 37,362 1

[a] TV = total variation.

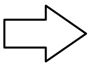
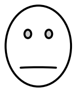
[b] EV = repeatability of the equipment's variation measurement system.

[c] AV = reproducibility/appraiser variation.

[d] IA = interaction variation.

[e] %RES = resolution of the measurement system.

Table B.6 — GRR analysis table (Reference Variable: Tolerance)

Reference	Tolerance		T[a] = 160		
Component	**Variance**	**Standard deviation**	**95 % lower CI**	**95 % upper CI**	**Result**
Repeatability	4,077 78	EV[b] = 2,019 35	1,713 84	2,458 43	%EV = 7,57 %
Reproducibility	0,914 40	AV[c] = 0,956 24	0,000 00	8,704 20	%AV = 3,59 %
Interaction	8,967 08	IA[d] = 2,994 51	1,960 62	4,755 26	%IA = 11,23 %
GRR	13,959 3	σ_{MS} = 3,736 21	3,541 19	9,429 53	%R&R = 14,01 %
Resolution	%RES[e] = 0,63 %				
Repeatability & Reproducibility	%R&R = 14,01 %				
Number of distinct categories	NDC = 16				
	Measurement system capable (RES, %R&R)				
Minimum reference figure for capable measuring system					$T_{min(R\&R)}$ = 224,172

[a] T = tolerance.

[b] EV = repeatability of the equipment's variation measurement system.

[c] AV = reproducibility/appraiser variation.

[d] IA = interaction variation.

[e] %RES = resolution of the measurement system.

Table B.7 — GRR analysis table (Reference Variable: Process variation)

Reference	1 × Process Variation $1 \times \sigma_P$ = 29,400 0				
Component	**Variance**	**Standard deviation**	**95 % lower CI**	**95 % upper CI**	**Result**
Repeatability	4,077 78	EV[a] = 2,019 35	1,713 84	2,458 43	%EV = 6,87 %
Reproducibility	0,914 40	AV[b] = 0,956 24	0,000 00	8,704 20	%AV = 3,25 %
Interaction	8,967 08	IA[c] = 2,994 51	1,960 62	4,755 26	%IA = 10,19 %
GRR	13,959 3	σ_{MS} = 3,736 21	3,541 19	9,429 53	%R&R=12,71 %
Resolution	%RES[d] = 3,40 %				
Repeatability & Reproducibility	%R&R = 12,71 %				
Number of distinct categories	NDC = 16				
⇨ Measurement system capable (RES, %R&R)					
Minimum reference figure for capable measuring system	$1 \times \sigma_{P\ min(\%R\&R)}$ = 37,362 1				

[a] EV = repeatability of the equipment's variation measurement system.

[b] AV = reproducibility/appraiser variation.

[c] IA = interaction variation.

[d] %RES = resolution of the measurement system.

From the results shown above, the %R&R based on tolerance is 14,01 % and the %R&R based on process variation is 12,71 %. Both these metrics are higher than the criteria for acceptance of 10 %, which indicates that corrective or improvement actions should be taken on the load cell to improve its measurement precision. This is the opposite of the results calculated using the range method.

B.8 Conclusions

Upon comparing the variance component analysis between the two methods, the range method and the ANOVA method give different results. This applies to the calculation method as well as for the different reference variables (total variation, tolerance, and process variation).

— The average-range method estimated a higher repeatability (estimated at 2,25) compared to that of the ANOVA method (estimated at 2,02).

— The average-range method does not provide the operator-part interaction component of variance. As such, the estimated GRR (2,64) is much smaller than that estimated by the ANOVA method (3,74), which is able to estimate this interaction component (2,99). It is mainly due to this significantly large operator-part interaction, which could be estimated by the ANOVA method, that the %R&R metrics go above 10 % in the results from the ANOVA method.

In summary, from this example it can be concluded that the ANOVA method is superior compared to the average-range method in estimating the precision of a measurement system. The average-range method is not able to estimate the operator-part interaction, and therefore tends to underestimate the %R&R, making the measurement system look more capable than it really is, and thus masking opportunities for improvement of the measurement system.

In order to get comparable results from the capability study for measurement systems, it is important to stipulate the calculation method as well as the reference variable to be used. Different calculation results from different measuring systems can be compared only using the same calculation method and the same reference variable.

In this example, the average range method results in a smaller %R&R value than the ANOVA method and thus in a presumably better measurement system. However, this should not be generalized, because in other examples the ANOVA method might result in a smaller %R&R value than the average range method.

B.9 Improvement actions or suggestions

The following suggestions for improvement to the GRR were made:

— design and fabricate a placement jig for the load cell in the work-holder so that the placement of the load cell is consistently at the same position for every measurement;

— study the sources of inherent variation of the cap placement machine and reduce it by improving or redesigning the placement arm mechanism.

Annex C
(informative)

GRR for motor shaft radial run-out of the axis

C.1 Measurement data type

Variable data

C.2 Gauge used for measurement

Gauge name: Dial indicator

Resolution: 0,01 mm

C.3 Measurement process description

A motor manufacturer produces motors for air conditioners. For motors, the radial run-out of the axis is one of the most important quality characteristics. The tolerance of radial run-out acceptable by the clients is less than 0,03 mm. If the radial run-out of the axis is greater than 0,03 mm, it may cause the motor to make a loud noise which is unacceptable for customers. Dial indicators are used to measure the radial run-out of the axis. The product chosen for GRR study was the model YSK30-6A motor.

C.4 Possible sources of measurement system variation

In the measurement process, generally one dial indicator is used. The measurement results can be different for different testers (operators) even though the same gauge is used. Also, for one tester using the gauge to measure one motor, the measurement results vary for repeated measurements. It is hard to know if there is any interaction between the operators and the motors to be measured.

Therefore, the sources for possible variations of this measurement system are: operators, pure error generated by repeated measurements, and interaction effects between operators and motors.

C.5 Sampling plan

Two testers (Carol and Bill) were selected randomly from a qualified group of operators.

Ten motors were randomly selected from the current production process and labelled 1 to 10, ensuring that the label could not be seen by the two operators in order to avoid the Hawthorne effect.

Each operator measured the 10 motors in random order for 3 rounds or trials. The measurement process was conducted as follows: first, Carol was asked to measure the 10 motors in a random order, then Bill was asked to do it in another random order. That was the first round for the measurement. Then Carol and Bill measured a second round and a third round without knowing that they were actually measuring the same 10 parts.

This is a typical 2-factor crossed case for a GRR study.

C.6 Measurement data

The measurement data are shown in Table C.1.

Table C.1 — Results of the experiment

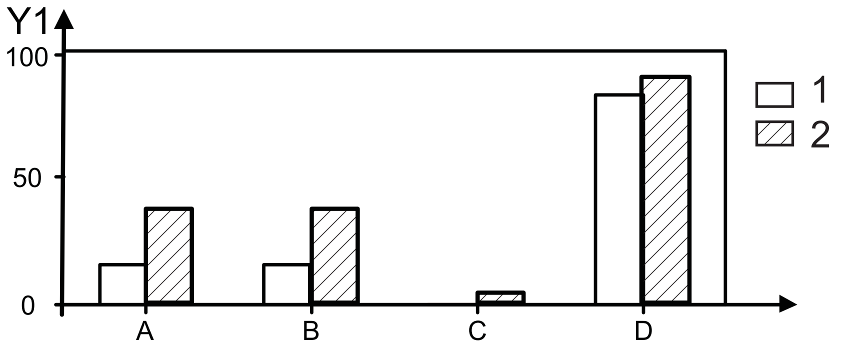
Operater	Part	Round	Data	Operater	Parts	Round	Data
Carol	1	1	0,025	Bill	1	1	0,020
Carol	1	2	0,020	Bill	1	2	0,015
Carol	1	3	0,020	Bill	1	3	0,020
Carol	2	1	0,030	Bill	2	1	0,025
Carol	2	2	0,045	Bill	2	2	0,040
Carol	2	3	0,030	Bill	2	3	0,030
Carol	3	1	0,015	Bill	3	1	0,020
Carol	3	2	0,015	Bill	3	2	0,015
Carol	3	3	0,015	Bill	3	3	0,020
Carol	4	1	0,010	Bill	4	1	0,010
Carol	4	2	0,010	Bill	4	2	0,010
Carol	4	3	0,010	Bill	4	3	0,010
Carol	5	1	0,040	Bill	5	1	0,040
Carol	5	2	0,040	Bill	5	2	0,030
Carol	5	3	0,040	Bill	5	3	0,040
Carol	6	1	0,045	Bill	6	1	0,030
Carol	6	2	0,045	Bill	6	2	0,040
Carol	6	3	0,045	Bill	6	3	0,040
Carol	7	1	0,010	Bill	7	1	0,010
Carol	7	2	0,020	Bill	7	2	0,015
Carol	7	3	0,010	Bill	7	3	0,015
Carol	8	1	0,010	Bill	8	1	0,020
Carol	8	2	0,010	Bill	8	2	0,010
Carol	8	3	0,010	Bill	8	3	0,015
Carol	9	1	0,025	Bill	9	1	0,020
Carol	9	2	0,025	Bill	9	2	0,030
Carol	9	3	0,020	Bill	9	3	0,020
Carol	10	1	0,045	Bill	10	1	0,030
Carol	10	2	0,030	Bill	10	2	0,025
Carol	10	3	0,030	Bill	10	3	0,040

C.7 Statistical method used for GRR studies

Method: ANOVA

C.8 Statistical analysis

Software: Minitab 14[4)]



Key

A %R&R

B repeatability

C reproducibility

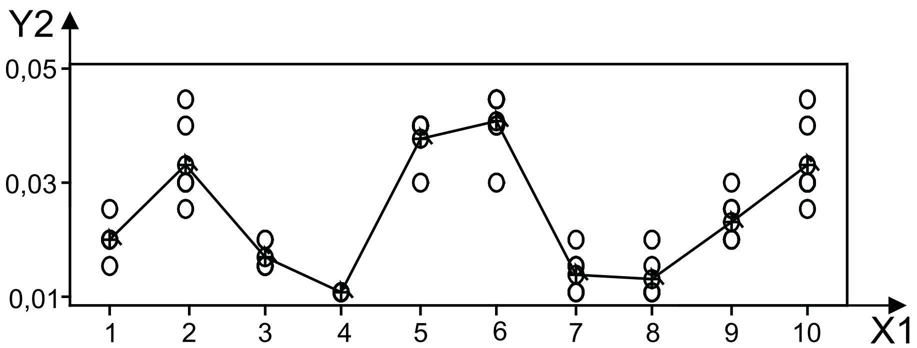
D part-to-part

Y1 percent

1 % contribution

2 % study variation

a) Components of variation



Key

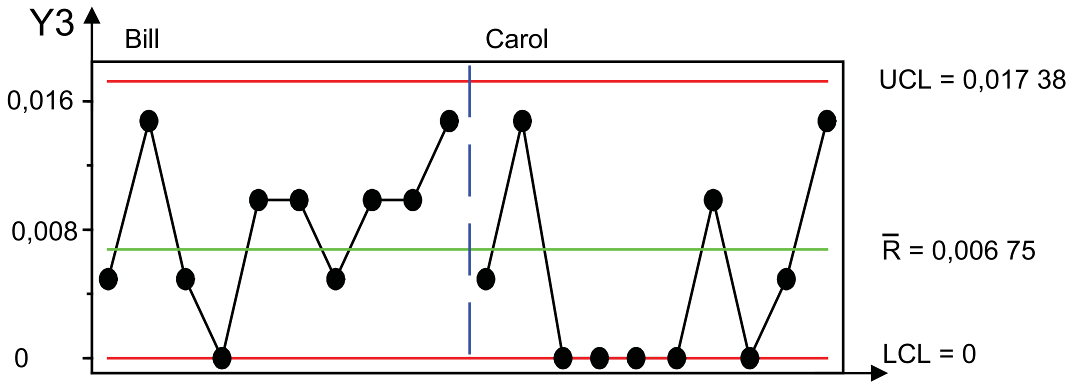
X1 parts

Y1 data

b) Data by parts

Figure C.1 — GRR study for M100 process (*continued*)

4) Minitab is the trade name of a product supplied by Minitab, Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.



Key

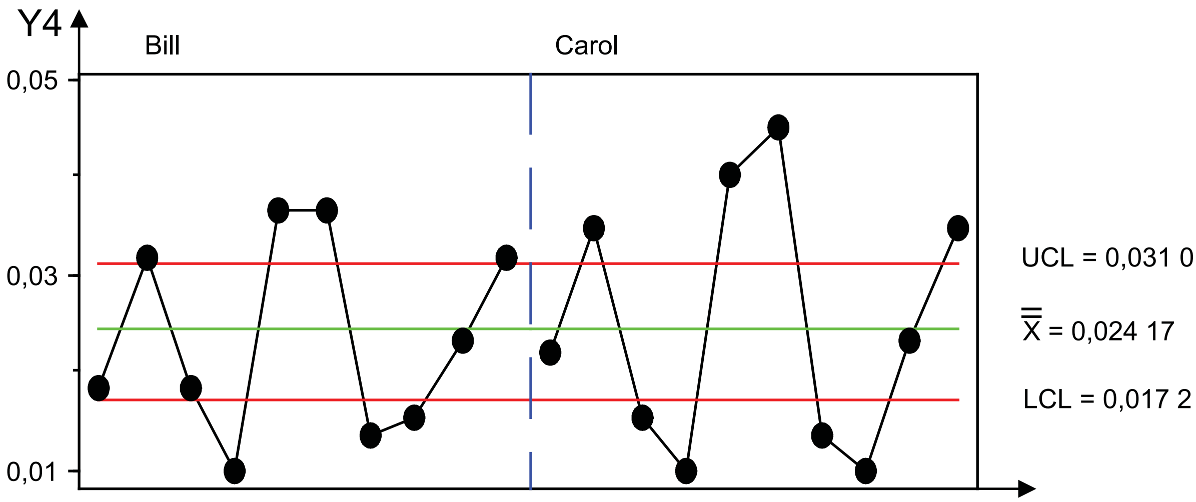
Y3 sample range

c) R chart by operator

Key

Y2 data

d) Data by operator

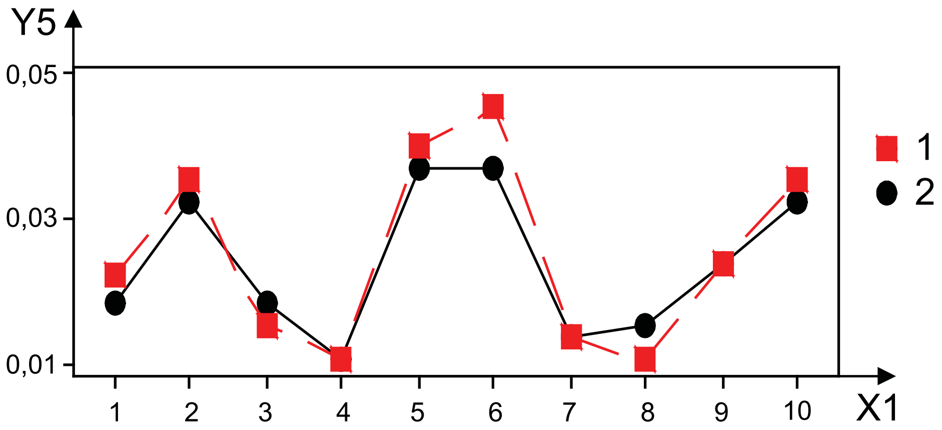


Key

Y4 sample mean

e) Xbar chart by operator

Figure C.1 — GRR study for M100 process (*continued*)



Key

X1 parts

Y5 average

1 Bill

2 Carol

f) Operator * parts interaction

Figure C.1 — GRR study for M100 process

Table C.2 — Two-way ANOVA table with interaction

Source	DF	SS	MS	F	P
Parts	9	0,007 150 0	0,000 783 3	35,546 2	0,000
Operator	1	0,000 026 7	0,000 026 7	1,210 1	0,300
Operator * Parts	9	0,000 198 3	0,000 022 0	0,944 4	0,499
Repeatabliity	40	0,000 933 3	0,000 023 3		
Total	59	0,008 208 3			

Table C.3 — GRR table

Source	VarComp	% Contribution (of VarComp)
Total GRR	0,000 023 5	15,62
Repeatability	0,000 023 3	15,52
Reproducibility	0,000 000 2	0,10
Operator	0,000 000 2	0,10
Operator * Parts	0,000 000 0	0,00
Part-to-part	0,000 126 9	84,38
Total variation	0,000 150 4	100,00

Table C.4 — GRR statistics

Source	StdDev (SD)	Study Var (6 * SD)	% Study Var (% SV)
Total GRR	0,004 846 4	0,029 078 4	39,52
Repeatability	0,004 830 5	0,028 982 8	39,39
Reproducibility	0,000 392 8	0,002 357 0	3,20
Operator	0,000 392 8	0,002 357 0	3,20
Operator * Parts	0,000 000 0	0,000 000 0	0,00
Part-to-part	0,011 264 2	0,067 585 3	91,86
Total variation	0,012 262 6	0,073 575 4	100,00
NDC = 3			

Looking at Figure C.1 and Tables C.2, C.3 and C.4, it can be seen that the major source of variation for this measurement system is repeatability; neither reproducibility nor the interaction between the operator and the motor is significant.

As shown in Table C.3, the measurement system variance takes up 15,62 % of the total variance, of which repeatability takes up 15,52 %. Reproducibility takes up only 0,1 %, which is insignificant, and mainly comes from the difference between the operators. The interaction between the operator and the motor is also insignificant.

It can be concluded from the given results that, for this measurement system, %R&R = 39,52 %, NDC = 3.

C.9 Conclusions

The GRR was 39,52 %, i.e. above the 30 % limit. Therefore, this measurement system was not acceptable and needed to be improved.

C.10 Improvement actions or suggestions

Because the major source of variation of GRR is repeatability error, this meant that the resolution of the dial indicator was too low. The following actions were taken:

— clock gauges capable of measuring to 0,001 mm were used to replace dial indicators;

— after this replacement, another GRR study was conducted and %R&R was found to be less than 10 %. This means that the new measurement system is acceptable.

Annex D
(informative)

GRR for the rip-off force of the A/B cover of charges[5)]

D.1 Measurement data type

Variable data

D.2 Gauge used for measurement

Gauge name: TIRA[6)] test 27025

Resolution: 0,1N (1N = 0,102 kgf)

D.3 Measurement process description

A company in Shenzhen, engaged in producing linear and switch-mode chargers and power supplies for mobile terminals and other electronic handheld devices, has a set of advanced test systems. There are many items to be tested for chargers. One important item is the rip-off force of an A/B cover after ultrasonic welding. The specification requirement is that the force must be more than 650 N. In order to improve the process capability of the rip-off force of the chargers or to precisely determine if the chargers are good or bad with respect to the rip-off force, the measurement system must first be validated. As shown in Figure D.1, the chargers are fixed to the fixture by operators, then pulled off at the speed of 10 mm/min. The machine (as shown in Figure D.2) records the maximal force required to rip-off the A/B cover of chargers. This is destructive testing and the experimenters believe specimens created from the same batch are more homogenous than specimens created from different batches.

5) This example was supplied by Prof Zhen He, Tianjin University, China. It has been adapted from the article "An Applied Study of Destructive Measurement System Analysis", by Y.-j. Han and Zhen He, Proc. Industrial Electronics and Applications, 2007. ICIEA 2007. Second IEEE Conference on Industrial Electronics and Applications, pp. 30-34. Used with permission from the publisher, © 2007 IEEE.

6) TIRA is a trade name of products supplied by TIRA Gmbh, Germany. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Figure D.1 — Fixture during rip-off force test

Figure D.2 — Machine for rip-off force

D.4 Possible sources of measurement system variation

For specimens from the same batch, A and B covers share the same tool, cavity, injection machine, time, raw material and parameter of injection. They are assembled with PWB (printed wiring board) by the same operator and serially joined under the same ultrasonic welding machine, parameters and operator. Holes are made in two sides of the chargers by the same operator, using a drill.

For this destructive measurement process, one rip-off test machine is used. Because the measurement process is destructive, no replicate measurements can be obtained on the same unit. An appropriate way to deal with the situation is to use a nested design and to consider units from the same batch to be similar. Any differences in measurements on A covers and on B covers will be attributed to the repeatability of the equipment or to certain sources for reproducibility, such as operator variability, but not to a difference between the rip-off force of A versus the rip-off force of B. The same batch of specimens measured two times by one operator can be used to estimate the repeatability. Another batch of specimens measured by different operators can be used to estimate the reproducibility. This leads to the use of a special type of nested design called “staggered nested design”.

D.5 Sampling plan

There are three randomly selected operators in the study. Each operator measures four batches and two specimens for every batch, for a total of 24 observations in this experiment. The plan is presented in Table D.1.

Table D.1 — Experiment design for verifying measurement system for rip-off force

				Operator			Time			Batch		
Total samples		24		3			2			12		
Sample	**Batch 1**	**Batch 2**	**Batch 3**	**Batch 4**	**Batch 5**	**Batch 6**	**Batch 7**	**Batch 8**	**Batch 9**	**Batch 10**	**Batch 11**	**Batch 12**
Operator 1	1-1	2-1					7-1	8-1		10-2	11-2	
	1-2	2-2										
Operator 2			3-1	4-1			7-2		9-1	10-1		12-2
			3-2	4-2								
Operator 3					5-1	6-1		8-2	9-2		11-1	12-1
					5-2	6-2						

NOTE

Batch 1 = No hole, material 1 (LG 121H-87273)

Batch 2 = No hole, material 2 (Samsung PC HN-1064I)

Batch 3 = No hole, material 3 (Bayer Ku2-1514 BBS073 901510)

Batch 4 = Hole 1, material 1 (LG 121H-87273)

Batch 5 = Hole 1, material 2 (Samsung PC HN-1064I)

Batch 6 = Hole 1, material 3 (Bayer Ku2-1514 BBS073 901510)

Batch 7 = Hole 2, material 1 (LG 121H-87273)

Batch 8 = Hole 2, material 2 (Samsung PC HN-1064I)

Batch 9 = Hole 2, material 3 (Bayer Ku2-1514 BBS073 901510)

Batch 10 = Holes 1 & 2, material 1 (LG 121H-87273)

Batch 11 = Holes 1 & 2, material 2 (Samsung PC HN-1064I)

Batch 12 = Holes 1 & 2, material 3 (Bayer Ku2-1514 BBS073 901510)

D.6 Measurement data

The measurement data are shown in Table D.2.

Table D.2 — Experiment results

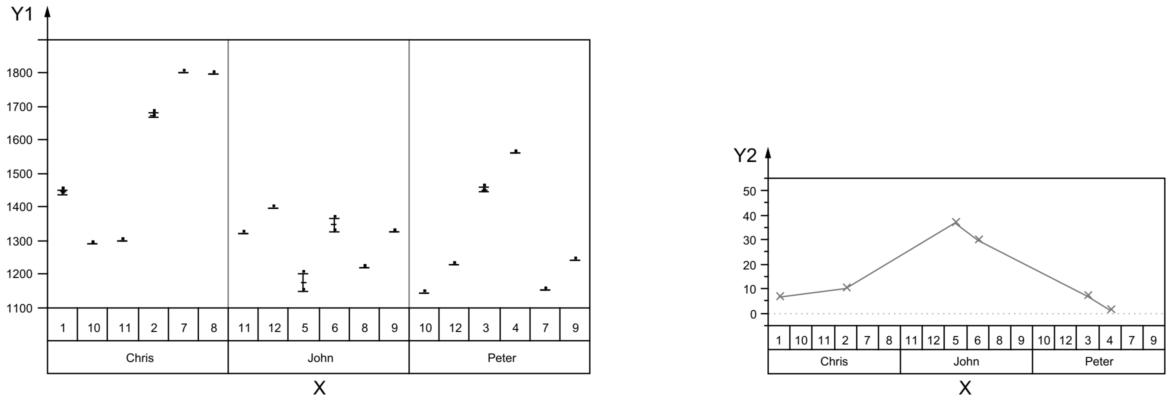
Operator	Batch	Time	Data	Operator	Batch	Time	Data	Operator	Batch	Time	Data
Chris	1	1	1 449	Peter	3	1	1 448	John	5	1	1 326
Chris	1	2	1 440	Peter	3	2	1 458	John	5	2	1 368
Chris	2	1	1 682	Peter	4	1	1 561	John	6	1	1 150
Chris	2	2	1 668	Peter	4	2	1 563	John	6	2	1 202
Chris	7	1	1 801	Peter	7	2	1 153	John	8	2	1 222
Chris	8	1	1 796	Peter	10	1	1 144	John	9	2	1 326
Chris	10	2	1 292	Peter	9	1	1 242	John	11	1	1 322
Chris	11	2	1 301	Peter	12	2	1 229	John	12	1	1 396

D.7 Statistical method used for GRR studies

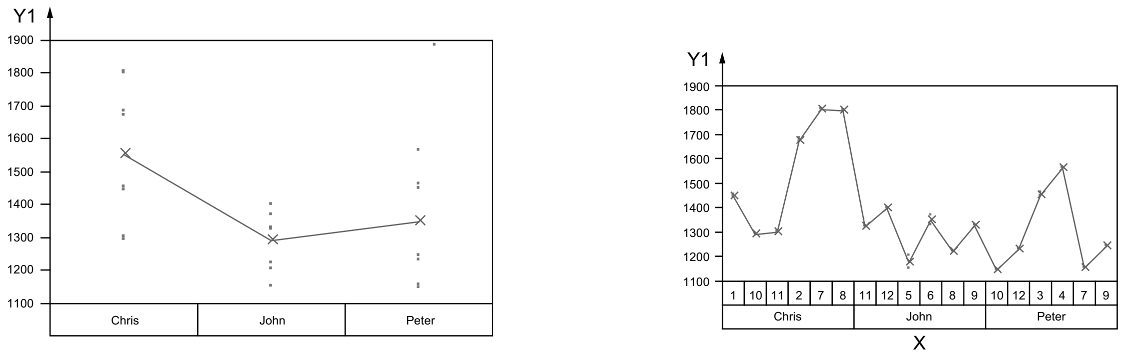
Method: REML

D.8 Statistical analysis

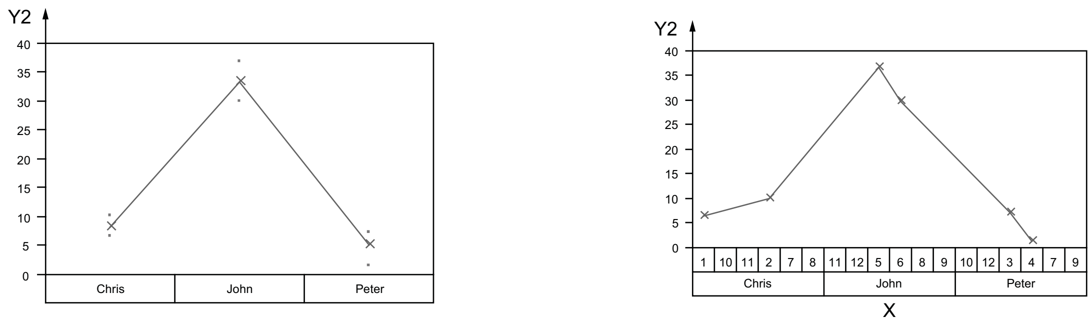
Software: JMP 7[7)]



a) Variability gauge



b) GRR mean plots



c) GRR standard deviation plots

Key

X batch

Y1 data

Y2 standard deviation

Figure D.3 — Graphical results from staggered nested GRR analysis

7) JMP is the trade name of a product supplied by SAS Institute, Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Variability gauge

GRR StdDev[a] Plots

Data was unbalanced, so a REML Fit was performed.

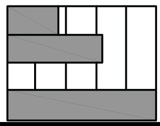
Response data

REML variance component estimates

Random Effect	Var[b] Ratio	Var Component	% of Total
Operator	40,453 97	16 362,16	34,795
Batch[Operator]	74,808 173	30 258,15	64,344
Residual		404,477 39	0,860
Total			100,000

–2 LogLikelihood = 286,796 545 28

Variance components

Component	Var Component	% of Total	20 40 60 80	Sqrt[c] (Var Comp[d])
Operator	16 362,716	34,8		127,92
Batch[Operator]	30 258,15	64,3		173,95
Within	404,477	0,860 1		20,11
Total	47 025,408	100,0		216,85

Discrimination ratio

Source	Ratio
Operator	1,437 801 1
Batch[Operator]	2,146 908 56

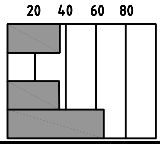
GRR

Measurement Source		Variation (6 × StdDev)		Which is 6 × sqrt of
Repeatability	(EV)	120,669 7	Equipment Variation	V(Within)
Reproducibility	(AV)	767,501 0	Appraiser Variation	V(Operator)
Operator		767,501 0		V(Operator)
GRR	(RR)	776,929 2	Measurement Variation	V(Within) + V(Operator)
Part Variation	(PV)	1 043,693 3	Part Variation	V(Batch)
Total Variation	(TV)	1 301,120 6	Total Variation	V(Batch) + V(Within) + V(Operator)

6	k
59,712 3	% R&R = 100*(RR/TV)
0,744 4	Precision to Part Variation = RR/PV
1	Number of Distinct Categories = 1,41(PV/RR)

Using last column “Batch” for Part.

Variance Component for GRR

Component	Var Component	% of Total	20 40 60 80
GRR	16 767,194	35,66	
Repeatability	404,477	0,86	
Reproducibility	16 362,716	34,80	
Part-To-Part	30 258,215	64,34	

a StdDev = Standard Deviation.

b Var = Variance.

c Sqrt = Square root.

d Var Comp = Variance Component.

Figure D.4 — GRR report from staggered nested GRR design

From Figure D.3 and from the GRR report in Figure D.4, it can be shown that the major source of variation for this measurement system is reproducibility. The measurement system variance takes up 35,66 % of the total variance, with reproducibility taking up 34,80 %, and repeatability 0,86 %, which is insignificant.

The GRR analysis result for this measurement system is %R&R = 59,7 % and NDC = 1.

D.9 Conclusions

%R&R is 59,7 %, which is much higher than 30 %; this measurement system is therefore not acceptable and should not be used to measure the rip-off force of chargers. The NDC is equal to 1 (less than 5), which means that the resolution of the measurement system is extremely low.

D.10 Improvement actions or suggestions

The main source of variability is due to different operators' training and standard procedures are likely to be needed as an immediate measure to ensure the usefulness of this measurement system.

Bibliography

[1] ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

[2] ISO 3534-2:2006, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

[3] ISO 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments*

[4] ISO 5725-1, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

[5] ISO 10012, *Measurement management systems — Requirements for measurement processes and measuring equipment*

[6] ISO/TS 16949, *Quality management systems — Particular requirements for the application of ISO 9001:2008 for automotive production and relevant service part organizations*

[7] ISO/IEC Guide 99, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*

[8] IWA 1:2005, *Quality management systems — Guidelines for process improvements in health service organizations*

ICS 03.120.30

Price based on 40 pages