PD ISO/TR 12888:2011

BSI Standards Publication

# Selected illustrations of gauge repeatability and reproducibility studies

NO COPYING WITHOUT BSI PERMISSION EXCEPT AS PERMITTED BY COPYRIGHT LAW

*raising standards worldwide*™

Full document details: PD ISO/TR 12888:2011

## National foreword

This Published Document is the UK implementation of ISO/TR 12888:2011.

The UK participation in its preparation was entrusted to Technical Committee MS/6, Methodologies for business process improvement using statistical methods.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2012

Published by BSI Standards Limited 2012

ISBN 978 0 580 65175 5

ICS 03.120.30

**Compliance with a British Standard cannot confer immunity from legal obligations.**

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 March 2012.

## Amendments issued since publication

| Amd. No. | Date | Text affected |
|---|---|---|

Documents
Document Sharing Hub

Full document details: PD ISO/TR 12888:2011

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)

Documents
Document Sharing Hub

© ISO 2011

Full document details: PD ISO/TR 12888:2011

**COPYRIGHT PROTECTED DOCUMENT**

© ISO 2011

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland



Documents
Document Sharing Hub
© ISO 2011 – All rights reserved

Full document details: PD ISO/TR 12888:2011

# Contents

Page

**Foreword** ..... iv
**Introduction** ..... v
**1 Scope** ..... 1
**2 Terms and definitions** ..... 1
**3 Symbols and abbreviated terms** ..... 2
**4 Generic description of GRR studies** ..... 3
**4.1 Overview of the structure of GRR studies** ..... 3
**4.2 Overall objectives of GRR** ..... 3
**4.3 Measurement process description** ..... 4
**4.4 GRR studies methodology** ..... 4
**4.5 Sampling plan for GRR studies** ..... 6
**4.6 Data analysis (numerically and graphically)** ..... 8
**4.7 Conclusions and suggestions** ..... 8
**5 Description of Annexes A through D** ..... 9
**5.1 Comparison of the examples** ..... 9
**5.2 Example summaries** ..... 9
**Annex A** (informative) **GRR for automated testing of RF performance of cell phones** ..... 10
**Annex B** (informative) **RF Metal-ceramics semiconductor package assembly load cell GRR study** ..... 19
**Annex C** (informative) **GRR for motor shaft radial run-out of the axis** ..... 28
**Annex D** (informative) **GRR for the rip-off force of the A/B cover of charges** ..... 34
**Bibliography** ..... 40

© ISO 2011 – All rights reserved


Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

## 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*.


© ISO 2011 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

# 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.

© ISO 2011 – All rights reserved


Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

# 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]

© ISO 2011 – All rights reserved



Documents
Document Sharing Hub

Full document details: PD ISO/TR 12888:2011

**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


© ISO 2011 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

| | |
|---|---|
| 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 |
| $\sigma$ | standard deviation |
| $\sigma_{MS}$ | standard deviation of measurement system |
| $\sigma_r$ | repeatability standard deviation |
| $\sigma_R$ | reproducibility standard deviation |
| $\sigma_P$ | 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.

© ISO 2011 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

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.


© ISO 2011 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

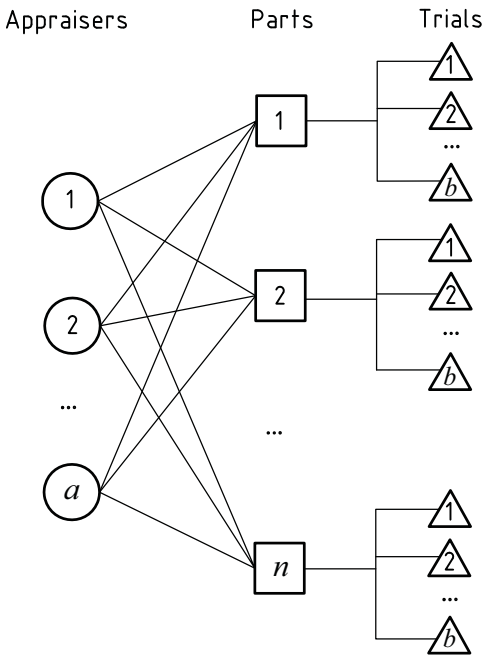


**Figure 1 — Crossed design in GRR studies**

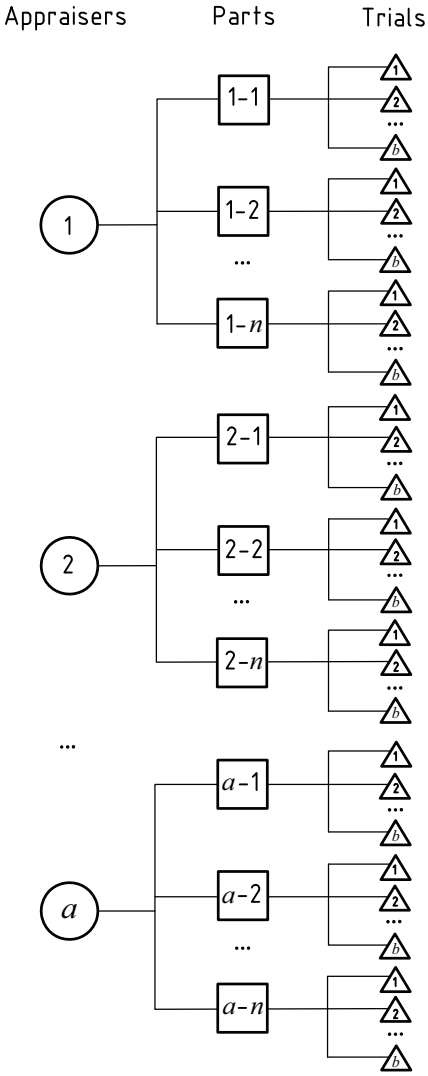


**Figure 2 — Nested design in GRR studies**

© ISO 2011 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

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 ($\sigma_r$), using the range of multiple observations of one appraiser measuring the same part with the same equipment, and of reproducibility standard deviation ($\sigma_R$), 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,

— $\sigma_{\mathrm{MS}}$ denotes the standard deviation of measurement system error, where $\sigma_{\mathrm{MS}}$ is the square root of the sum of $\sigma_r^2$ and $\sigma_R^2$,

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

— $\sigma_{\mathrm{P}}$ is the standard deviation of the manufacturing process without measurement error.

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

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

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

## 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.


© ISO 2011 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

**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.

© ISO 2011 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011

## 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.


© ISO 2011 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD ISO/TR 12888:2011