TECHNICAL REPORT

ISO TR 11753

First edition
1992-06-01

Rubber and rubber products — Confidence intervals for repeatability and reproducibility values determined by inter-laboratory tests

Caoutchouc et produits en caoutchouc — Intervalles de confiance de répétabilité et de reproductibilité déterminées par essais interlaboratoires

Reference number
ISO/TR 11753:1992(E)

Contents

Page

Annexes

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland
Printed in Switzerland

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.

The main task of technical committees is to prepare International Standards, but in exceptional circumstances a technical committee may propose the publication of a Technical Report of one of the following types:

- type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;
- type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;
- type 3, 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).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/TR 11753, which is a Technical Report of type 3, was prepared by Technical Committee ISO/TC 45, *Rubber and rubber products*.

Introduction

The subject of confidence intervals of repeatability and reproducibility determined in inter-laboratory tests was first discussed in Working Group 15 "Application of statistical methods" of ISO Technical Committee 45 "Rubber and rubber products" at its 1988 meeting in Paris. At the 1989 meeting of WG15 in Kuala Lumpur a first draft was considered, following which consultation took place with the chairman and members of ISO Technical Committee 69 "Application of statistical methods".

A revised draft was considered by WG15 at its 1992 meeting in Stockholm, where it was agreed that the text should be issued as a Type 3 Technical Report. This proposal, in Resolution 2207 of ISO/TC45 (document 45 N 5928) was approved unanimously by all ISO/TC45 members present at the 1992 meeting.

Rubber and rubber products — Confidence intervals for repeatability and reproducibility values determined by inter-laboratory tests

1 Scope

The precision values estimated from inter-laboratory tests vary on repetition of these tests. This document describes a method for obtaining confidence intervals for the unknown precision values.

2 Field of application

ISO 5725 describes the performance of inter-laboratory tests and states how data for the repeatability and reproducibility of a standardized test method can be obtained from the results of inter-laboratory tests.

The repeatability standard deviation s_r or repeatability limit r is used as a measure for repeatability, and the reproducibility standard deviation s_R or the reproducibility limit R as a measure for reproducibility.

Only estimates of the precision values are obtainable from inter-laboratory tests. On repetition of an inter-laboratory test slightly different estimates will be obtained owing to the random influences that are also present in inter-laboratory tests. ISO 5725 contains no data on the possible errors that may occur when the values of r and R are estimated.

However, knowledge of the deviation of determined estimates of repeatability and reproducibility from the true values is very important, since it enables the following two groups of questions to be answered:

a) Questions relating to the planning of inter-laboratory tests:

How extensive should the inter-laboratory test be, i.e. how many laboratories, materials and individual values are needed to achieve a given degree of accuracy in estimating the precision values?

b) Questions relating to the application of results from inter-laboratory tests:

How accurate are the repeatability and reproducibility estimates determined in an inter-laboratory test? How long are the confidence intervals of the precision values sought?

The precision values determined in an inter-laboratory test are used to characterize the test method in the corresponding standard. It is important to know to how many decimal places they should be reported.

The confidence intervals of the precision values to be estimated, and hence the accuracy of the estimates, depend on the scope of testing: in general the precision of precision values is the greater and their confidence interval the narrower, the larger the number of laboratories p and number of individual values n determined for each test property level in each laboratory.

In this document it is assumed that values measured under repeatability conditions and in different laboratories are distributed normally; see Annex A1. Deviations from the normal distribution generally make the variability of the precision values somewhat greater and the confidence interval of these values correspondingly longer.

Note: The validity of the normal distributions of the deviations under repeatability and reproducibility conditions is a prerequisite for the relations stated in this standard; it is possible that this prerequisite is not fulfilled. In Section 7.3.1 of DIS 5725 Part 1 other unimodal distributions of the deviations under repeatability conditions are allowed also.

In addition it is necessary to comply with the condition, laid down in Section 10.1 of DIS 5725 Part 1 that the laboratories participating in the inter-laboratory test should be chosen at random.

On the assumption of normal distribution the confidence intervals are stated with a preselected error probability α. In this standard $\alpha = 0.10$. This means that there is a 10 % probability that the true value is outside the given confidence interval, with $\alpha/2 = 0.05$ or 5 % below the lower limit and likewise $\alpha/2 = 0.05$ or 5 % above the upper limit.

3 Definitions

The definitions used in this document are given in

ISO 3534-3:1985, Statistics - Vocabulary and symbols - Part 3: Design of experiments.

ISO 5725:1986, Precision of test methods - Determination of repeatability and reproducibility for a standard test method by inter-laboratory tests.

4 Calculation of confidence intervals

4.1 Confidence interval of the repeatability

The results of an inter-laboratory test for a test property level with p laboratories and n_i measurements in i. laboratory can be evaluated by a variance analysis with simple classification and random effects. Assuming that the deviations of the results obtained in the laboratories have a normal distribution, the confidence interval of the true repeatability value r' can be calculated from the χ^2 distribution (see Annex A2):

(1) $$r\,A_{r,1} < r' < r\,A_{r,2}$$

The following is true of the confidence interval of the quotient r'/r:

(1a) $$A_{r,1} < r'/r < A_{r,2}$$

where (1b) $$A_{r,1} = \sqrt{\nu_2/\chi^2(\nu_2,Q)}$$

and (1c) $$A_{r,2} = \sqrt{\nu_2/\chi^2(\nu_2,P)}$$

Estimated repeatability standard deviation	s_r
True repeatability standard deviation	σ_r
Estimated repeatability limit	$r = 2{,}8\ s_r$
True repeatability limit	$r' = 2{,}8\ \sigma_r$
Number of laboratories	p
Number of measurements in i.laboratory (for i = 1, 2 ... p)	n_i
total number measurements per material level	$N = \Sigma\, n_i$
Degree of freedom of the repeat measurements	$\nu_2 = \Sigma\, n_i - p$
Probability of the χ^2 fractiles for the upper limit	$P = 0{,}05$
Probability of the χ^2 fractiles for the lower limit	$Q = 0{,}95$
Fractiles of the χ^2 distribution of ν and P	$\chi^2(\nu,P)$

The χ^2 fractiles $\chi^2(\nu_2,P)$ and $\chi^2(\nu_2,Q)$ can be taken from tables in standard literature. However, the approximation formula in Annex B can be used also.

The χ^2 fractiles are defined by

(2) $w\ (\ \chi^2 \leq \chi^2(\nu,P)\ ) = P$

i.e. there is a probability P that the random quantity χ^2 of a χ^2 distribution with degree of freedom ν is smaller than or equal to the fractiles $\chi^2(\nu,P)$, tabulated for ν and P.

If $Q = 1 - (\alpha/2) = 0.95$ and $P = \alpha/2 = 0.05$ are chosen for the probabilities, a probability of error of 10 % is obtained for the confidence interval, i.e. there is a probability of 10 % that the quotient r'/r sought is not covered by the confidence interval.

In the orthogonal case, i.e. $n_i = n$ for all i = 1, 2 ... p, the following is true

(3a) $N = p\ \bar{n}$ and (3b) $\nu_2 = p\ (n-1)$.

The approximation formula given in Annex B was used to calculate the confidence limits $A_{r,1}$ and $A_{r,2}$ for n = 2, 3, 5 and 9 test results per laboratory and p = 8 to 60 laboratories, which are given in Table 1 (see next page) and in Figure 1 (see page 14).

Table 1: Confidence intervals $A_{r,1}$ and $A_{r,2}$ for r'/r

n	p	ν_2	$\chi^2(\nu_2,5\%)$	$\chi^2(\nu_2,95\%)$	$A_{r,1}$	$A_{r,2}$
2	8	8	2.73	15.51	0.72	1.71
2	10	10	3.94	18.31	0.74	1.59
2	12	12	5.23	21.03	0.76	1.52
2	14	14	6.57	23.68	0.77	1.46
2	16	16	7.96	26.30	0.78	1.42
2	18	18	9.39	28.87	0.79	1.38
2	20	20	10.85	31.41	0.80	1.36
2	25	25	14.61	37.65	0.81	1.31
2	30	30	18.49	43.77	0.83	1.27
2	35	35	22.47	49.80	0.84	1.25
2	40	40	26.51	55.76	0.85	1.23
2	50	50	34.76	67.50	0.86	1.20
2	60	60	43.19	79.08	0.87	1.18
3	8	16	7.96	26.30	0.78	1.42
3	10	20	10.85	31.41	0.80	1.36
3	12	24	13.85	36.42	0.81	1.32
3	14	28	16.93	41.34	0.82	1.29
3	16	32	20.07	46.19	0.83	1.26
3	18	36	23.27	51.00	0.84	1.24
3	20	40	26.51	55.76	0.85	1.23
3	25	50	34.76	67.50	0.86	1.20
3	30	60	43.19	79.08	0.87	1.18
3	35	70	51.74	90.53	0.88	1.16
3	40	80	60.39	101.88	0.89	1.15
3	50	100	77.93	124.34	0.90	1.13
3	60	120	95.70	146.57	0.90	1.12
5	8	32	20.07	46.19	0.83	1.26
5	10	40	26.51	55.76	0.85	1.23
5	12	48	33.10	65.17	0.86	1.20
5	14	56	39.80	74.47	0.87	1.19
5	16	64	46.59	83.68	0.87	1.17
5	18	72	53.46	92.81	0.88	1.16
5	20	80	60.39	101.88	0.89	1.15
5	25	100	77.93	124.34	0.90	1.13
5	30	120	95.70	146.57	0.90	1.12
5	35	140	113.66	168.61	0.91	1.11
5	40	160	131.76	190.52	0.92	1.10
5	50	200	168.28	233.99	0.92	1.09
5	60	240	205.14	277.14	0.93	1.08
9	8	64	46.59	83.68	0.87	1.17
9	10	80	60.39	101.88	0.89	1.15
9	12	96	74.40	119.87	0.89	1.14
9	14	112	88.57	137.70	0.90	1.12
9	16	128	102.87	155.40	0.91	1.12
9	18	144	117.27	173.00	0.91	1.11
9	20	160	131.76	190.52	0.92	1.10
9	25	200	168.28	233.99	0.92	1.09
9	30	240	205.14	277.14	0.93	1.08
9	35	280	242.25	320.03	0.94	1.08
9	40	320	279.56	362.72	0.94	1.07
9	50	400	354.64	447.63	0.95	1.06
9	60	480	430.20	532.08	0.95	1.06

4.2 Confidence interval of the reproducibility

A χ^2 distribution can also be assumed for estimation of the confidence interval of the reproducibility limit R, but in this case it is only approximately true (see Annex A.3). On the assumption of orthogonality the degree of freedom ν_3 for this distribution can be calculated from the following equation:

$$\nu_3 = \frac{n^2 (1+\gamma^2)^2 \nu_1 \nu_2}{(n+\gamma^2)^2 \nu_2 + (n-1)^2 \gamma^4 \nu_1} \tag{4}$$

If there is a marked deviation from orthogonality, i.e. if the number of measurements per laboratory varies greatly, the method given in Annex A.3.2 can be used. The degrees of freedom $\nu_1 = p - 1$ and $\nu_2 = p(\Sigma n_i - 1)$ are obtained from the number of laboratories p and number of individual measurements n_i per laboratory.

The degree of freedom ν_3 depends on a further parameter γ' or g':

(5a) $\gamma' = \sigma_r / \sigma_L$ or $g' = \sigma_r / \sigma_R$

As the true values σ_r, σ_R and σ_L are not known, the estimated values s_r, s_R and s_L are used as approximations, $\gamma' = \sigma_r/\sigma_L$ being estimated by $\gamma = s_r/s_L$.

From the definition of σ_R^2 or s_R^2 (see ISO 5725) it follows that

(6a) $\sigma_R^2 = \sigma_L^2 + \sigma_r^2$ or (6b) $s_R^2 = s_L^2 + s_r^2$

Therefore

(7a) $\gamma = \frac{s_r}{s_L} = \frac{s_r}{\sqrt{s_R^2 - s_r^2}} = \frac{g}{\sqrt{1-g^2}}$ and (7b) $g = \frac{s_r}{s_R}$

After an inter-laboratory test has been evaluated the estimated values s_r^2 and s_R^2 from which γ can be calculated according to (7a) are known. In the planning of an inter-laboratory test assumptions for γ must be made in order to fix the number of laboratories (see Section 5.1).

After the degree of freedom ν_3 has been calculated, the confidence interval for the reproducibility limit R is obtained analogously to that of the repeatability limit r, as described in Section 4.1. The following is valid for the confidence interval of the quotient R'/R:

(8) $$A_{R,1} < R'/R < A_{R,2}$$

where (8a) $$A_{R,1} = \sqrt{\nu_3/\chi^2(\nu_3,Q)}$$

and (8b) $$A_{R,2} = \sqrt{\nu_3/\chi^2(\nu_3,P)}$$

Note: The degree of freedom ν_3 in Equation 4 is generally not an integer number. In reading χ^2 fractiles $\chi^2(\nu_3,P)$ from a χ^2 table it is necessary whether to interpolate or to take the nearest smaller number. However, it is also possible to enter the non-integral value ν_3 in the approximation formula given in Annex B for the χ^2 fractiles.

If $Q = 1 - (\alpha/2) = 0.95$ and $P = \alpha/2 = 0.05$ are chosen for the probabilities, a probability of error of 10 % is obtained for the confidence interval, i.e. there is a probability of 10 % that the quotient R'/R sought is not covered by the confidence interval.

The approximation formula Equation 8 was used with the aid of the approximation formula given in Annex B for the χ^2 distribution, to calculate the confidence limits $A_{R,1}$ and $A_{R,2}$ given in Table 2 for n = 2, 5 and 15 measurement results per laboratory, p = 8 to 60 laboratories and γ = 0.05, 0.33, 0.67 and 1.00 and in Figure 2 for n=2 (see page 15).

Table 2: Confidence intervals $A_{R,1}$ and $A_{R,2}$ for R'/R

p	γ	g	$A_{R,1}$	$A_{R,2}$	$A_{R,1}$	$A_{R,2}$	$A_{R,1}$	$A_{R,2}$
			n=2		n=5		n=15	
8	0.05	0.05	0.71	1.80	0.71	1.79	0.71	1.79
10	0.05	0.05	0.73	1.64	0.73	1.64	0.73	1.64
12	0.05	0.05	0.75	1.55	0.75	1.55	0.75	1.55
14	0.05	0.05	0.76	1.48	0.76	1.48	0.76	1.48
16	0.05	0.05	0.77	1.44	0.77	1.44	0.78	1.44
18	0.05	0.05	0.79	1.40	0.79	1.40	0.79	1.40
20	0.05	0.05	0.79	1.37	0.79	1.37	0.79	1.37
25	0.05	0.05	0.81	1.32	0.81	1.32	0.81	1.32
30	0.05	0.05	0.83	1.28	0.83	1.28	0.83	1.28
35	0.05	0.05	0.84	1.25	0.84	1.25	0.84	1.25
40	0.05	0.05	0.85	1.23	0.85	1.23	0.85	1.23
50	0.05	0.05	0.86	1.20	0.86	1.20	0.86	1.20
60	0.05	0.05	0.87	1.18	0.87	1.18	0.87	1.18
8	0.33	0.31	0.71	1.73	0.72	1.69	0.72	1.68
10	0.33	0.31	0.74	1.60	0.74	1.57	0.75	1.55
12	0.33	0.31	0.76	1.51	0.76	1.49	0.76	1.48
14	0.33	0.31	0.77	1.45	0.78	1.43	0.78	1.42
16	0.33	0.31	0.78	1.41	0.79	1.39	0.79	1.38
18	0.33	0.31	0.79	1.37	0.80	1.36	0.80	1.35
20	0.33	0.31	0.80	1.35	0.81	1.33	0.81	1.33
25	0.33	0.31	0.82	1.30	0.82	1.28	0.83	1.28
30	0.33	0.31	0.83	1.26	0.84	1.25	0.84	1.25
35	0.33	0.31	0.84	1.24	0.85	1.23	0.85	1.22
40	0.33	0.31	0.85	1.22	0.86	1.21	0.86	1.21
50	0.33	0.31	0.87	1.19	0.87	1.18	0.87	1.18
60	0.33	0.31	0.88	1.17	0.88	1.16	0.88	1.16
8	0.67	0.56	0.73	1.62	0.76	1.51	0.77	1.47
10	0.67	0.56	0.76	1.51	0.78	1.43	0.79	1.39
12	0.67	0.56	0.77	1.44	0.79	1.37	0.80	1.34
14	0.67	0.56	0.79	1.39	0.81	1.33	0.82	1.30
16	0.67	0.56	0.80	1.35	0.82	1.30	0.83	1.28
18	0.67	0.56	0.81	1.32	0.83	1.28	0.83	1.26
20	0.67	0.56	0.82	1.30	0.83	1.26	0.84	1.24
25	0.67	0.56	0.83	1.26	0.85	1.22	0.86	1.21
30	0.67	0.56	0.85	1.23	0.86	1.20	0.87	1.18
35	0.67	0.56	0.86	1.21	0.87	1.18	0.88	1.17
40	0.67	0.56	0.86	1.19	0.88	1.17	0.88	1.16
50	0.67	0.56	0.88	1.17	0.89	1.15	0.89	1.14
60	0.67	0.56	0.89	1.15	0.90	1.13	0.90	1.12
8	1.00	0.71	0.75	1.54	0.79	1.39	0.81	1.32
10	1.00	0.71	0.77	1.45	0.81	1.33	0.83	1.27
12	1.00	0.71	0.79	1.39	0.82	1.29	0.84	1.24
14	1.00	0.71	0.80	1.35	0.83	1.26	0.85	1.22
16	1.00	0.71	0.81	1.32	0.84	1.24	0.86	1.20
18	1.00	0.71	0.82	1.29	0.85	1.22	0.87	1.18
20	1.00	0.71	0.83	1.27	0.86	1.20	0.87	1.17
25	1.00	0.71	0.84	1.23	0.87	1.18	0.89	1.15
30	1.00	0.71	0.86	1.21	0.88	1.16	0.90	1.14
35	1.00	0.71	0.87	1.19	0.89	1.15	0.90	1.12
40	1.00	0.71	0.87	1.17	0.90	1.13	0.91	1.11
50	1.00	0.71	0.89	1.15	0.91	1.12	0.92	1.10
60	1.00	0.71	0.89	1.14	0.91	1.11	0.92	1.09

5 Examples

5.1 Planning of inter-laboratory tests

In the planning of inter-laboratory tests it must be decided whether the number of laboratories p is sufficient and how many individual measurements n per laboratory must be performed so that the repeatability variance σ_r^2 and reproducibility variance σ_R^2 are obtained with the required degree of precision. The tables and figures in Section 4 can be used in making these decisions. ISO 5725 recommends that the number of laboratories should not be less than 8. For our example it will be assumed that 12 laboratories participate and that 2 individual measurements are performed per laboratory. The confidence interval of the repeatability limit r is obtained from Table 1 and Figure 1. The lower limit is 0.76 and the upper limit 1.52, i.e. a variability of 24 % to lower values and 52 % to higher values must be expected. If the number of individual measurements is raised to n = 9, the interval is 0.89 to 1.14 or 11 % to lower values and 14 % to higher values.

The confidence interval for the reproducibility R can be obtained from Table 2 and Figure 2. Here, however, the factor $g = s_r/s_R$ must be taken into account (see Equations 7a and 7b in Section 4.2).

This factor is not known while the test are being planned. In many cases may be g = 0.5, i.e. the reproducibility limit R is twice as large as the repeatability limit r. In the case of larger factors, e.g. g = 0.7 or γ = 1, the test methods are ones that hardly necessitate the performance of inter-laboratory tests, i.e. here the supplier and purchaser have no problems when comparing the measurements. For a good test method, however, a further prerequisite, namely that s_r should be small, should likewise be met. For small factors below g = 0.3, measures should be taken to improve the description and coordination of test methods.

For our example we shall assume extreme values for the factor, initially g = 0.05, i.e. a reproducibility limit 20 times as large as the repeatability limit. For n = 2 and p = 2 the confidence interval for the reproducibility limit then ranges from 0.75 to 1.55, i.e. 25 % to lower values and 55 % to higher vales. If one aims at an accuracy of the reproducibility limit of about ± 20 %, then the number laboratories must be raised to p = 35 for n = 2. An increase in n, the number of individual measurements per laboratory, has practically no influence on the confidence interval of the repeatability value at g = 0.05.

If we consider the other extreme case, namely $\gamma = 1$ or g = 0.71, i.e. a reproducibility value only 30 % larger than the repeatability value, it can be seen from Table 2 and Figure 2 that the confidence interval of the repeatability value r' for 12 laboratories is 0.79 to 1.39, i.e. it may vary by 21 % to lower values and 39 % to higher values. For g = 0.71 it is advantageous to increase the number n of results per laboratory. From Table 2 it is evident that the upper limit of the confidence interval falls as n decreases.

However, the total expenditure for the production of samples and performance of the test must be considered. The total expenditure for an inter-laboratory test increases with the total number N = n p. Taking the case of n = 2 and p = 18 for g = 0.71 in Table 2, one obtains 0.82 and 1.29 as the lower and upper limit of the confidence interval at a total number n p = 36. However, the same confidence interval is obtained for n = 5 and p = 12 at a total number n p = 60. As this example shows, the test expenditure is lower for n = 2 and p = 18; in contrast, raising the number of individual measurements to n = 5 and reducing the number of laboratories to p = 12 in order to keep the confidence interval the same necessitates greater testing expenditure. It is therefore recommended to keep n = 2 and to use as many laboratories p as possible, preferably 20 to 40, but, at any rate as many as can be persuaded to participate.

5.2 Evaluation of the results of inter-laboratory tests

Calculation of the confidence intervals and the use of the Bartlett test will be demonstrated with reference to an inter-laboratory test which is among the examples given in ISO 5725. The purpose of the test was to determine the softening point of pitch, the softening temperature being measured in °C. The following results were obtained:

Table 3: Results of an inter-laboratory test to determine the softening temperature of pitch (see ISO 5725, Table 10)

Level	n	p	ν_2	s_r^2	r	ν_3	s_R^2	R
88,40	2	15	15	1,2303	3,11	21,4	2,7878	4,68
96,27	2	15	15	0,8560	2,59	19,5	2,5504	4,47
97,07	2	16	16	0,9869	2,78	19,1	4,0414	5,63
101,96	2	16	16	1,0078	2,81	19,7	3,6670	5,36
95,92	-	-	62	1,0195	2,83	79,7	3,2475	5,05

p = 15 or, as the case may be, = 16 laboratories participated with n = 2 individual results. For material level 88.40 the repeatability limit is 3.11 and the reproducibility limit 4.68. According to Equation 3b, $\nu_2 = 15$ is obtained for the degree of freedom of the repeatability variance. The confidence intervals can be calculated from the corresponding χ^2 fractiles according to Equations 2 and 8. For the quotient r'/r one obtains a lower limit of 0.77 and an upper limit of 1.44, i.e. a variability of 23 % to lower values and 44 % to higher values, and for the reproducibility limit R'/R one obtains, correspondingly, 20 % to lower values and 34 % to higher values for g = 0.66 and $\nu_3 = 21,4$.

The confidence intervals of the precision values of the other 3 levels of the material can be calculated in the same way. As stated already in ISO 5725, the variances of the four test property levels can be summarized. In ISO 5725 this is apparent from a plot of r or R against the material level m. The fact that the variances s_R^2 do not differ significantly in the materials levels is more suitably investigated by the Bartlett test (see Annex C). This is now possible since the degrees of freedom ν_3 for s_R^2 are known. The test quantity obtained is $\chi^2 = 1.38$, i.e. a value smaller than the χ^2 fractiles: $\chi^2(3;\ 0.95) = 7.82$. Therefore the s_R^2 values do not differ significantly at the given error probability of 5 %. The differences of the repeatability variances s_r^2 can be investigated by the Bartlett test in the same way. There are no significant differences in this case either. It is therefore justifiable to form averages weighted across the variances and, with these averages, to calculate new values for repeatability and reproducibility, which will then be valid for the entire range of the test property levels.

As the summarization also increases the degrees of freedom ν_2 and ν_3, the totals $\nu_2 = 62$ and $\nu_3 = 79.7$ can be used for the degrees of freedom, so that correspondingly narrower confidence intervals are obtained according to Equations 1 and 8. One obtains a relative confidence interval of 13 % to lower values and 18 % to higher values for the repeatability limit and of 11 % to lower values and 15 % to higher values for the reproducibility limit, and an absolute confidence interval of 2.5 to 3.3 for the repeatability limit and of 4.5 to 5.8 for the reproducibility limit. It is therefore not advisable to state the estimates r and R too precisely. These should be rounded off accordingly. The final values in this example should be stated as r = 2.8 and R = 5.1. It should be noted that, of the s_r^2 and s_R^2 values given in Table 3, averages weighted across the degrees of freedom have been formed:

(10a) $s_r^2 = \Sigma \nu_{2i} s_{ri}^2 / \nu_{2g}$ where $\nu_{2g} = \Sigma \nu_{2i}$,

or

(10b) $s_R^2 = \Sigma \nu_{3i} s_{Ri}^2 / \nu_{3g}$ where $\nu_{3g} = \Sigma \nu_{3i}$.

Fig. 1: Confidence Limits of Repeatability

n = 2, 3, 5 and 9

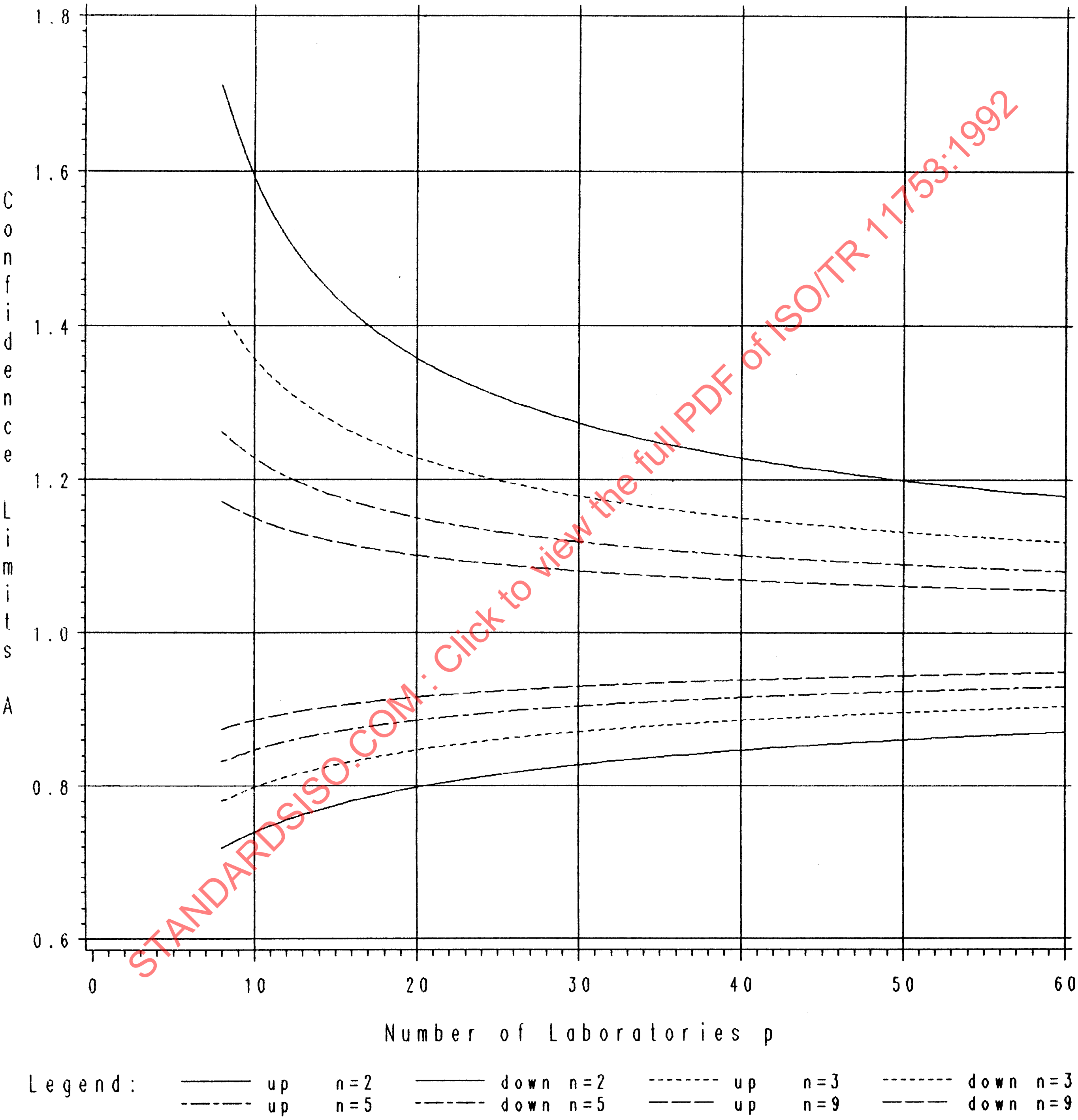
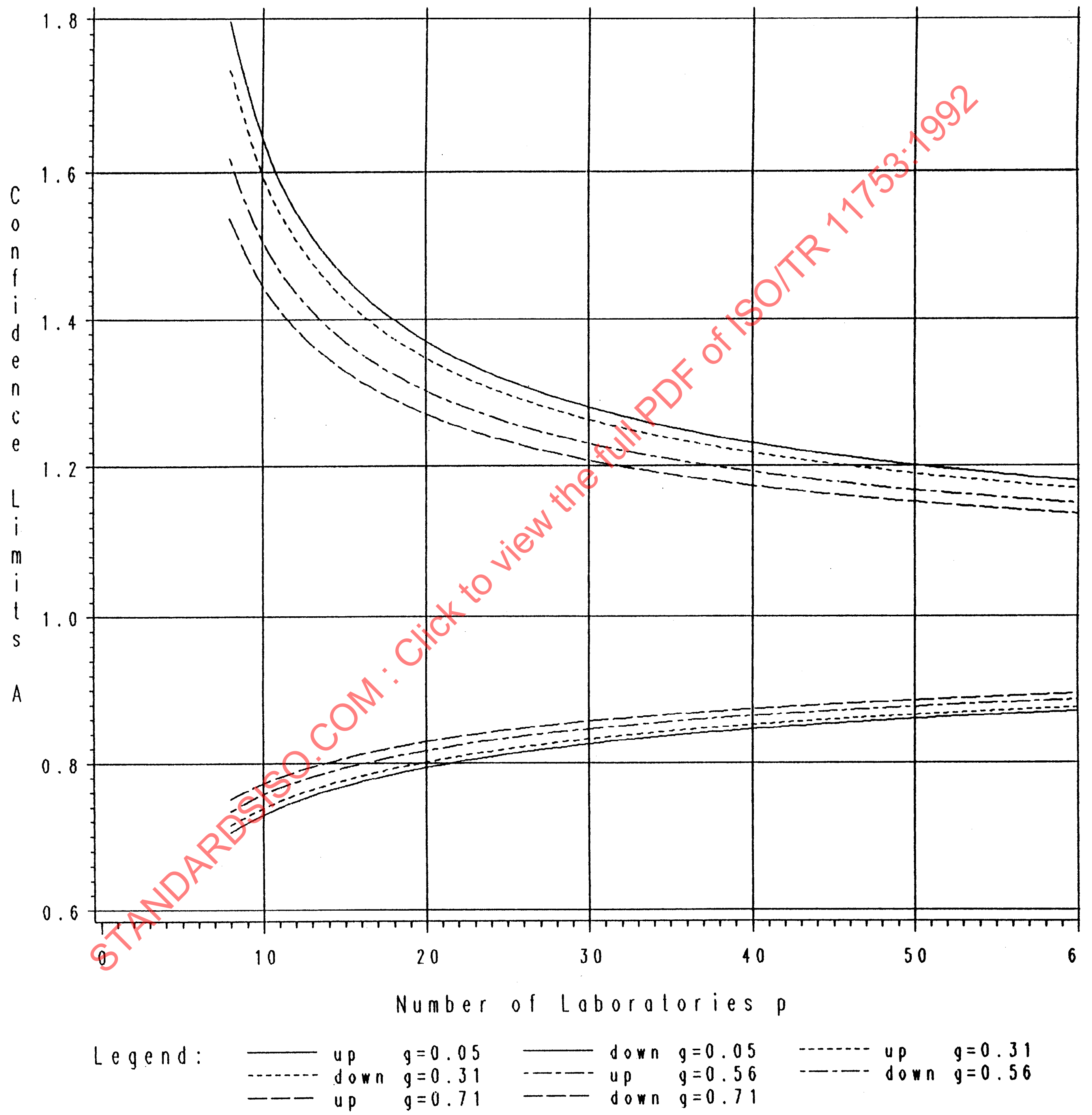


Fig.2: Confidence Limits of Reproducibility

n=2 and g=0.05, 0.31, 0.56 and 0.71



Annex A

Explanation of the mathematical equations

A.1 Variance analysis

According to ISO 5725 the following linear model for the measured values y_{ik} is adopted for estimating the precision:

(A1) $$y_{ik} = m + B_i + e_{ik}$$

$$(i=1,2,...p;\ k=1,2...n_i)$$

where
p = number of laboratories
n_i = number of measurements obtained in i. laboratory
m = true property of the investigated material
B_i = influence of i. laboratory
e_{ik} = variation in the results of k. measurement in i. laboratory

A normal distribution with the anticipated value 0 and variance σ_L^2 ("variance between the laboratories") is presupposed for the random quantity B_i and a normal distribution with anticipated value 0 and variance σ_r^2 ("repeatability variance") for the random quantity e_{ik}.

According to ISO 5725 the reproducibility variance is given by:

(A2) $$\sigma_R^2 = \sigma_L^2 + \sigma_r^2$$

The scheme of the variance analysis for simple classification [1,2] is valid for this model:

	SQ	ν	s^2	F
between the laboratories:	$SQ_1 = \Sigma n_i (y_i - y)^2$	$\nu_1 = p-1$	$s_1^2 = SQ_1/\nu_1$	s_1^2/s_2^2
within the laboratories:	$SQ_2 = \Sigma\ \Sigma (y_{ik} - y_i)^2$	$\nu_2 = N-p$	$s_2^2 = SQ_2/\nu_2$	

where SQ = sum of the squares

ν = degree of freedom

s^2 = variance

F = F-value

Average of i.laboratory: $\overline{y}_i = \Sigma\, y_{ik}/n_i$

Total average: $\overline{\overline{y}} = \Sigma\, \overline{y}_i/p$

Number of measured values: $N = \Sigma\, n_i$

The following anticipated values are valid in the case of a variance analysis with random effects:

(A4a) $E(s_1^2) = \overline{n}\, \sigma_L^2 + \sigma_r^2$ und (A4b) $E(s_2^2) = \sigma_r^2$

In addition

(A5) $\overline{n} = (N - (1/N)\, \Sigma n_i^2) / (p-1)$

In the orthogonal case, i.e. if each laboratory has the same number of measured values ($n_i = n$ for all $i = 1, 2 \ldots p$),

$\overline{n} = n.$

According to Equation A4b s_2^2 is an unbiased estimate for σ_r^2. According to Equation A4a $s_L^2=(s_1^2-s_2^2)/n$ is an unbiased estimate for σ_L^2. Hence $\sigma_R^2=\sigma_L^2+\sigma_r^2$ gives the unbiased estimate $s_R^2=s_L^2+s_2^2$ for σ_R^2.

A.2 Confidence interval and degree of freedom of the repeatability variance

As the random quantity e_{ik} has a normal distribution in the adopted model, a χ^2 distribution with degree of freedom ν_2 and the following confidence interval[3)] is valid for the random quantity $\nu_2 s_2^2/\sigma_r^2$:

(A6a) $$\chi^2(\nu_2,P) < \nu_2 s_2^2/\sigma_r^2 < \chi^2(\nu_2,Q)$$

This confidence interval, is defined by the χ^2 fractiles with degree of freedom ν_2 and probability $P = \alpha/2$ for the lower confidence limit and by $Q = 1 - \alpha/2$ for the upper confidence limit.

Transformation gives:

(A6b) $$s_2^2\ (\nu_2/\chi^2(\nu_2,Q)) < \sigma_r^2 < s_2^2\ (\nu_2/\chi^2(\nu_2,P))$$

or when $r' = 2,8\ \sigma_r$ and $r = 2,8\ s_r = 2,8\ s_2$:

(A7) $$r\ A_{r,1} < r' < r\ A_{r,2}$$

where (A7a) $$A_{r,1} = \sqrt{\nu_2/\chi^2(\nu_2,Q)}$$

and (A7b) $$A_{r,2} = \sqrt{\nu_2/\chi^2(\nu_2,P)}$$

It should be noted that $P = \alpha/2$ now belongs to the upper, and $Q = 1 - \alpha/2$ to the lower, confidence limit.

A.3 Confidence interval and degree of freedom of the reproducibility variance

A.3.1 Approximation by means of a χ^2 distribution

If the unknown distribution of the random quantity $\nu_3 s_R^2/\sigma_R^2$ is approximated by means of a χ^2 distribution, the following holds for the degree of freedom [4]:

(A8)
$$\nu_3 = \frac{n^2 \ (1+\gamma'^2)^2 \ \nu_1 \ \nu_2}{(n+\gamma'^2)^2 \ \nu_2 \ + \ (n-1)^2 \ \gamma'^4 \ \nu_1}$$

where $\gamma' = \sigma_r/\sigma_L$.

It is presupposed that orthogonality exists, i.e. $n_i = n$ for all i = 1, 2 ... p, and that the random quantities $\nu_1 \ s_1^2/\sigma_1^2$ and $\nu_2 \ s_2^2/\sigma_2^2$ are independently χ^2-distributed.

As the true value $\gamma'=\sigma_r/\sigma_L$ is not known, only the estimated value $\gamma = s_r/s_L$ can be inserted, as an approximation, in Equation A8.

Then, with the degree of freedom ν_3 and the approximation for γ', a confidence interval is obtained as an approximation for the reproducibility limit R' (if $R' = 2.8 \ \sigma_R$ and $R = 2.8 \ s_R$) in a manner analogous to that used for r' in Section A.2:

(A9)
$$R \ A_{R,1} < R' < R \ A_{R,2}$$

where (A9a) $A_{R,1} = \sqrt{\nu_3/\chi^2(\nu_3,Q)}$

and (A9b) $A_{R,2} = \sqrt{\nu_3/\chi^2(\nu_3,P)}$,

with $P = \alpha/2$ belonging to the upper and $Q = 1 - \alpha/2$ to the lower confidence limit.

A.3.2 Estimation without assumption of a χ^2 distribution for the random quantity $\nu_3 \; s_R^2/\sigma_R^2$.

The random quantity $\nu_3 \; s_R^2/\sigma_R^2$ possesses no exact χ^2 distribution and in the case of marked deviation from orthogonality even $\nu_1 \; s_1^2/\sigma_1^2$ has no χ^2 distribution. In the absence of these assumptions R.K. Burdick and F.A.Graybill 5) and J.D. Thomas and R.A. Hultquist 6) derived confidence limits for σ_R^2 that permit somewhat more precise estimation than that described in Section A.2. Nevertheless if one dispenses with the approximation of a χ^2 distribution, no degree of freedom ν_3 for the reproducibility variance s_R^2 is obtained either.

With

(A10) $\quad G = s_y^2 + (1- \lambda) \; s_2^2$

as estimated value for σ_R^2, where

(A11a) $\quad s_y^2 = (\Sigma \bar{y}_i - (\Sigma \bar{y}_i)^2/p) \; / \; (p-1) \qquad$ mit $\bar{y}_i = \Sigma \; y_{ik}/n_i$

(A11b) $\quad \lambda = (\Sigma \, 1/n_i)/p$,

the following confidence interval is obtained for σ_R^2 for a given probability of error α at equal probability $\alpha/2$ that the lower limit will not be reached and that the upper limit will be exceeded:

(A12) $$G - (L_1^2 \; s_y^4 + (1- \lambda)^2 \; L_2^2 \; s_2^4 \;)^{1/2} < \sigma_R^2 < G + (H_1^2 \; s_y^4 + (1- \lambda)^2 \; H_2^2 \; s_2^4 \;)^{1/2}$$

where, with $Q = 1 - \alpha/2$ and $P = \alpha/2$,

(A13a) $\quad L_1 = 1 - (p-1) \; / \; \chi^2(p-1,Q)$

(A13b) $\quad L_2 = 1 - (N-p) \; / \; \chi^2(N-p,Q)$

(A14a) $\quad H_1 = (p-1) \; / \; \chi^2(p-1,P) - 1$

(A14b) $\quad H_2 = (N-p) \; / \; \chi^2(N-p,P) - 1$

Transformation of Equation A12 gives the following for the quotient R'/R:

(A15) $(1 - L_3)^{1/2} < R'/R < (1 + H_3)^{1/2}$

where

(A15a) $L_3 = \sqrt{\dfrac{L_1^2\ F_y^2 + (1-\lambda)^2\ L_2^2}{(F_y^2 + (1-\lambda))^2}}$

(A15b) $H_3 = \sqrt{\dfrac{H_1^2\ F_y^2 + (1-\lambda)^2\ H_2^2}{(F_y^2 + (1-\lambda))^2}}$

(A15c) $F_y = s_y^2 / s_2^2$

In the orthogonal case, i.e. ni = n for i = 1, 2 ... p,

(A16a) $\lambda = 1/n$

(A16b) $s_y^2 = s_1^2/n$

(A16c) $F_y = F/n$ where $F = s_1^2/s_2^2$

A.3.3 Comparison of the estimates in A.3.1 and A.3.2

To enable the estimates, given in Sections A.3.1 and A.3.2, of the confidence intervals for the quotient R'/R to be compared, the following notations are introduced:

(A17a) $A_{R,1} = A_{R,1\_31}$ see (A9a)

(A17b) $A_{R,2} = A_{R,2\_31}$ see (A9b)

(A17c) $(1-L_3)^{1/2} = A_{R,1\_32}$ see (A15a)

(A17d) $(1+H_3)^{1/2} = A_{R,2\_32}$ see (A15b)

The length of the confidence intervals is obtained from the differences:

(A18a) $D_{31} = A_{R,1\_31} - A_{R,2\_31}$ for procedure A.3.1,

(A18b) $D_{32} = A_{R,1\_32} - A_{R,2\_32}$ for procedure A.3.2.

The difference between the confidence intervals obtained from the estimates according to A.3.1 and A.3.2 is shown by the quotient:

(A18c) $Q = D_{32} / D_{31}$.

This quotient has been calculated for n = 2 and various p and γ in Table A1 and for γ = 0.33 and various p and n in Table A2. The quotient is generally only a few percentage points greater than 1 and exceeds 5 % in only 6 cases (by 12 % in the case of n = 2, p = 8 and γ = 1). The more precise procedure according to A.3.2 gives a somewhat longer confidence interval.

As the difference between the results of the two methods is only slight, but the theory of A.3.1 is more easily handled, this standard recommends using the theory of A.3.1 to estimate the confidence interval of the reproducibility limit R'.

This has the additional advantage that with the assumption of a χ^2 distribution with degree of freedom ν_3 for the random quantity $\nu_3\ s_R^2/\sigma_R^2$ the Bartlett test (see Section 5.2 and Annex C) can be used to compare the s_R^2 values and the F test to compare two s_R^2 values.

Tab. A1: Comparison of the theories of A.3.1 and A.3.2 for n = 2

Factors for calculation of the confidence intervals of the reproducibility

For lower confidence limits: $A_{R,1\_31}$, $A_{R,1\_32}$

For upper confidence limits: $A_{R,2\_31}$, $A_{R,2\_32}$

$D_{31} = A_{R,2\_31} - A_{R,1\_31}$, $D_{32} = A_{R,2\_32} - A_{R,1\_32}$, $Q = D_{32}/D_{31}$

p	γ	$A_{R,1\_31}$	$A_{R,1\_32}$	$A_{R,2\_31}$	$A_{R,2\_32}$	D_{31}	D_{32}	Q
8	0.05	0.71	0.71	1.80	1.80	1.09	1.09	1.00
10	0.05	0.73	0.73	1.64	1.64	0.91	0.91	1.00
12	0.05	0.75	0.75	1.55	1.55	0.80	0.80	1.00
14	0.05	0.76	0.76	1.48	1.48	0.72	0.72	1.00
16	0.05	0.77	0.77	1.44	1.44	0.66	0.66	1.00
18	0.05	0.79	0.79	1.40	1.40	0.61	0.61	1.00
20	0.05	0.79	0.79	1.37	1.37	0.58	0.58	1.00
25	0.05	0.81	0.81	1.32	1.32	0.50	0.50	1.00
30	0.05	0.83	0.83	1.28	1.28	0.45	0.45	1.00
35	0.05	0.84	0.84	1.25	1.25	0.42	0.42	1.00
40	0.05	0.85	0.85	1.23	1.23	0.39	0.39	1.00
50	0.05	0.86	0.86	1.20	1.20	0.34	0.34	1.00
60	0.05	0.87	0.87	1.18	1.18	0.31	0.31	1.00
8	0.33	0.71	0.72	1.73	1.77	1.02	1.04	1.03
10	0.33	0.74	0.74	1.60	1.62	0.86	0.88	1.02
12	0.33	0.76	0.76	1.51	1.53	0.75	0.77	1.02
14	0.33	0.77	0.78	1.45	1.47	0.68	0.69	1.01
16	0.33	0.78	0.79	1.41	1.42	0.62	0.63	1.01
18	0.33	0.79	0.80	1.37	1.38	0.58	0.59	1.01
20	0.33	0.80	0.80	1.35	1.35	0.54	0.55	1.01
25	0.33	0.82	0.82	1.30	1.30	0.48	0.48	1.01
30	0.33	0.83	0.83	1.26	1.27	0.43	0.43	1.01
35	0.33	0.84	0.84	1.24	1.24	0.39	0.40	1.01
40	0.33	0.85	0.85	1.22	1.22	0.37	0.37	1.00
50	0.33	0.87	0.87	1.19	1.19	0.32	0.33	1.00
60	0.33	0.88	0.88	1.17	1.17	0.29	0.30	1.00
8	0.67	0.73	0.75	1.62	1.71	0.88	0.95	1.08
10	0.67	0.76	0.77	1.51	1.57	0.75	0.80	1.06
12	0.67	0.77	0.79	1.44	1.48	0.66	0.70	1.05
14	0.67	0.79	0.80	1.39	1.43	0.60	0.63	1.04
16	0.67	0.80	0.81	1.35	1.38	0.55	0.57	1.03
18	0.67	0.81	0.82	1.32	1.35	0.52	0.53	1.03
20	0.67	0.82	0.83	1.30	1.32	0.48	0.50	1.03
25	0.67	0.83	0.84	1.26	1.28	0.43	0.44	1.02
30	0.67	0.85	0.85	1.23	1.24	0.38	0.39	1.02
35	0.67	0.86	0.86	1.21	1.22	0.35	0.36	1.01
40	0.67	0.86	0.87	1.19	1.20	0.33	0.33	1.01
50	0.67	0.88	0.88	1.17	1.18	0.29	0.29	1.01
60	0.67	0.89	0.89	1.15	1.16	0.26	0.27	1.01
8	1.00	0.75	0.78	1.54	1.66	0.79	0.88	1.12
10	1.00	0.77	0.79	1.45	1.53	0.68	0.73	1.09
12	1.00	0.79	0.81	1.39	1.45	0.60	0.64	1.07
14	1.00	0.80	0.82	1.35	1.40	0.55	0.58	1.06
16	1.00	0.81	0.83	1.32	1.36	0.50	0.53	1.05
18	1.00	0.82	0.83	1.29	1.33	0.47	0.49	1.04
20	1.00	0.83	0.84	1.27	1.30	0.44	0.46	1.04
25	1.00	0.84	0.85	1.23	1.26	0.39	0.40	1.03
30	1.00	0.86	0.87	1.21	1.23	0.35	0.36	1.03
35	1.00	0.87	0.87	1.19	1.20	0.32	0.33	1.02
40	1.00	0.87	0.88	1.17	1.19	0.30	0.31	1.02
50	1.00	0.89	0.89	1.15	1.16	0.27	0.27	1.02
60	1.00	0.89	0.90	1.14	1.14	0.24	0.25	1.01

Tab. A2: Comparison of the theories of A.3.1 and A.3.2 for $\gamma = 0.33$

Factors for calculation of the confidence intervals of the reproducibility

For lower confidence limits: $A_{R,1\_31}$, $A_{R,1\_32}$

For upper confidence limits: $A_{R,2\_31}$, $A_{R,2\_32}$

$D_{31}=A_{R,2\_31}-A_{R,1\_31}$, $D_{32}=A_{R,2\_32}-A_{R,1\_32}$, $Q=D_{32}/D_{31}$

p	n	$A_{R,1\_31}$	$A_{R,1\_32}$	$A_{R,2\_31}$	$A_{R,2\_32}$	D_{31}	D_{32}	Q
8	2	0.71	0.72	1.73	1.77	1.02	1.04	1.03
10	2	0.74	0.74	1.60	1.62	0.86	0.88	1.02
12	2	0.76	0.76	1.51	1.53	0.75	0.77	1.02
14	2	0.77	0.78	1.45	1.47	0.68	0.69	1.01
16	2	0.78	0.79	1.41	1.42	0.62	0.63	1.01
18	2	0.79	0.80	1.37	1.38	0.58	0.59	1.01
20	2	0.80	0.80	1.35	1.35	0.54	0.55	1.01
25	2	0.82	0.82	1.30	1.30	0.48	0.48	1.01
30	2	0.83	0.83	1.26	1.27	0.43	0.43	1.01
35	2	0.84	0.84	1.24	1.24	0.39	0.40	1.01
40	2	0.85	0.85	1.22	1.22	0.37	0.37	1.00
50	2	0.87	0.87	1.19	1.19	0.32	0.33	1.00
60	2	0.88	0.88	1.17	1.17	0.29	0.30	1.00
8	3	0.72	0.73	1.71	1.76	0.99	1.03	1.04
10	3	0.74	0.75	1.58	1.61	0.84	0.86	1.03
12	3	0.76	0.77	1.50	1.52	0.74	0.75	1.02
14	3	0.77	0.78	1.44	1.46	0.67	0.68	1.02
16	3	0.79	0.79	1.40	1.41	0.61	0.62	1.02
18	3	0.80	0.80	1.36	1.38	0.57	0.58	1.01
20	3	0.80	0.81	1.34	1.35	0.53	0.54	1.01
25	3	0.82	0.83	1.29	1.30	0.47	0.47	1.01
30	3	0.83	0.84	1.26	1.26	0.42	0.43	1.01
35	3	0.85	0.85	1.23	1.24	0.39	0.39	1.01
40	3	0.85	0.86	1.21	1.22	0.36	0.36	1.01
50	3	0.87	0.87	1.19	1.19	0.32	0.32	1.00
60	3	0.88	0.88	1.17	1.17	0.29	0.29	1.00
8	5	0.72	0.73	1.69	1.75	0.97	1.02	1.04
10	5	0.74	0.75	1.57	1.60	0.82	0.85	1.03
12	5	0.76	0.77	1.49	1.51	0.73	0.74	1.03
14	5	0.78	0.78	1.43	1.45	0.66	0.67	1.02
16	5	0.79	0.79	1.39	1.41	0.60	0.61	1.02
18	5	0.80	0.80	1.36	1.37	0.56	0.57	1.02
20	5	0.81	0.81	1.33	1.35	0.53	0.53	1.02
25	5	0.82	0.83	1.28	1.29	0.46	0.47	1.01
30	5	0.84	0.84	1.25	1.26	0.42	0.42	1.01
35	5	0.85	0.85	1.23	1.23	0.38	0.38	1.01
40	5	0.86	0.86	1.21	1.22	0.35	0.36	1.01
50	5	0.87	0.87	1.18	1.19	0.31	0.32	1.01
60	5	0.88	0.88	1.16	1.17	0.29	0.29	1.00
8	10	0.72	0.74	1.68	1.74	0.96	1.01	1.05
10	10	0.75	0.76	1.56	1.60	0.81	0.84	1.04
12	10	0.76	0.77	1.48	1.51	0.72	0.74	1.03
14	10	0.78	0.79	1.42	1.45	0.65	0.66	1.03
16	10	0.79	0.80	1.38	1.40	0.59	0.61	1.02
18	10	0.80	0.81	1.35	1.37	0.55	0.56	1.02
20	10	0.81	0.81	1.33	1.34	0.52	0.53	1.02
25	10	0.83	0.83	1.28	1.29	0.46	0.46	1.01
30	10	0.84	0.84	1.25	1.26	0.41	0.42	1.01
35	10	0.85	0.85	1.23	1.23	0.38	0.38	1.01
40	10	0.86	0.86	1.21	1.21	0.35	0.35	1.01
50	10	0.87	0.87	1.18	1.19	0.31	0.31	1.01
60	10	0.88	0.88	1.16	1.17	0.28	0.28	1.01