Institute for
Interlaboratory Studies

Results of Proficiency Test
Aniline in Leather/Footwear
February 2022

Organized by: Institute for Interlaboratory Studies
Spijkenisse, the Netherlands

Author: ing. A. Ouwerkerk
Correctors: ing. R.J. Starink & ing. A.S. Noordman-de Neef
Approved by: ing. A.S. Noordman-de Neef

Report: iis22A03

April 2022

CONTENTS

1 INTRODUCTION

The determination of Aniline in leather is known to give problems with the comparability of laboratory results. However, no appropriate reference materials are yet available.
As an alternative, participation in a proficiency test may enable laboratories to check their performance. Therefore, on request of several participants a proficiency test (laboratory-evaluating interlaboratory study) for the determination of Aniline in Leather was organized by the Institute for Interlaboratory Studies in 2022.

In this interlaboratory study 24 laboratories in 14 different countries registered for participation. See appendix 3 for the number of participants per country. In this report the results of the Aniline in Leather/Footwear proficiency test are presented and discussed. This report is also electronically available through the iis website www.iisnl.com.

2 SET UP

The Institute for Interlaboratory Studies (iis) in Spijkenisse, the Netherlands, was the organizer of this proficiency test (PT). Sample analyzes for fit-for-use and homogeneity testing were subcontracted to an ISO/IEC17025 accredited laboratory.
It was decided to send one leather sample of 3 grams labelled #22512.
The participants were requested to report rounded and unrounded test results. The unrounded test results were preferably used for the statistical evaluation.

2.1 QUALITY SYSTEM

The Institute for Interlaboratory Studies in Spijkenisse, the Netherlands, has implemented a quality system based on ISO/IEC17043:2010. This ensures strict adherence to protocols for sample preparation and statistical evaluation and 100% confidentiality of participant's data. Feedback from the participants on the reported data is encouraged and customer's satisfaction is measured on regular basis by sending out questionnaires.

2.2 PROTOCOL

The protocol followed in the organization of this proficiency test was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation' of June 2018 (iis-protocol, version 3.5). This protocol is electronically available through the iis website www.iisnl.com, from the FAQ page.

2.3 CONFIDENTIALITY STATEMENT

All data presented in this report must be regarded as confidential and for use by the participating companies only. Disclosure of the information in this report is only allowed by means of the entire report. Use of the contents of this report for third parties is only allowed by written permission of the Institute for Interlaboratory Studies. Disclosure of the identity of one or more of the participating companies will be done only after receipt of a written agreement of the companies involved.

2.4 SAMPLES

A batch of brown leather positive on Aniline was purchased on the local market. The batch was grinded into small pieces and after homogenization 40 small plastic bags were filled with approximately 3 grams each and labelled #22512.
The homogeneity of the subsamples was checked by the determination of Aniline in accordance with ISO17234 on 8 stratified randomly selected subsamples.

	Aniline in mg/kg
sample #22512-1	69.079
sample #22512-2	74.911
sample #22512-3	80.334
sample #22512-4	77.703
sample #22512-5	83.405
sample #22512-6	76.794
sample #22512-7	77.925
sample #22512-8	78.915

Table 1: homogeneity test results of subsamples #22512

From the above test results the relative standard deviation (RSD) was calculated and compared with 0.3 times the average relative standard deviation derived from the PTs 2010-2021 of AZO Dyes (see iis memo 2202) in Leather in agreement with the procedure of ISO13528, Annex B2 in the next table.

	Aniline
RSD% (observed)	5.4
reference method	iis memo 2202
0.3 x RSD% (reference method)	8.1

Table 2: evaluation of the relative standard deviation of subsamples #22512

The calculated relative standard deviation is in agreement with 0.3 times the relative standard deviation derived from previous iis PTs 2010-2021 of AZO dyes in Leather. Therefore, homogeneity of the subsamples was assumed.

To each of the participating laboratories one leather sample #22512 was sent on February 02, 2022.

2.5 ANALYZES

The participants were requested to determine Aniline.
It was also requested to report if the laboratory was accredited for the determined component and how many grams of sample intake was used.

It was explicitly requested to treat the sample as if it was a routine sample and to report the test result using the indicated unit on the report form and not to round the test result but report as much significant figures as possible. It was also requested not to report ‘less than’ results, which is above the detection limit, because such results cannot be used for meaningful statistical evaluations.

To get comparable test results a detailed report form and a letter of instructions are prepared. On the report form the reporting unit is given as well as the reference test method (when applicable) that will be used during the evaluation. The detailed report form and the letter of instructions are both made available on the data entry portal www.kpmd.co.uk/sgs-iis-cts/. The participating laboratories are also requested to confirm the sample receipt on this data entry portal. The letter of instructions can also be downloaded from the iis website www.iisnl.com.

3 RESULTS

During five weeks after sample dispatch, the test results of the individual laboratories were gathered via the data entry portal www.kpmd.co.uk/sgs-iis-cts/. The reported test results are tabulated per determination in appendix 1 of this report. The laboratories are presented by their code numbers.

Directly after the deadline, a reminder was sent to those laboratories that had not reported test results at that moment. Shortly after the deadline, the available test results were screened for suspect data. A test result was called suspect in case the Huber Elimination Rule (a robust outlier test) found it to be an outlier. The laboratories that produced these suspect data were asked to check the reported test results (no reanalyzes). Additional or corrected test results are used for data analysis and the original test results are placed under 'Remarks' in the result tables in appendix 1. Test results that came in after the deadline were not taken into account in this screening for suspect data and thus these participants were not requested for checks.

3.1 STATISTICS

The protocol followed in the organization of this proficiency test was the one as described for proficiency testing in the report ‘iis Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation’ of June 2018 (iis-protocol, version 3.5).
For the statistical evaluation the *unrounded* (when available) figures were used instead of the rounded test results. Test results reported as ‘<…’ or ‘>…’ were not used in the statistical evaluation.

First, the normality of the distribution of the various data sets per determination was checked by means of the Lilliefors-test, a variant of the Kolmogorov-Smirnov test and by the calculation of skewness and kurtosis. Evaluation of the three normality indicators in combination with the visual evaluation of the graphic Kernel density plot, lead to judgement of the normality being either ‘unknown’, ‘OK’, ‘suspect’ or ‘not OK’. After removal of outliers, this check was repeated. If a data set does not have a normal distribution, the (results of the) statistical evaluation should be used with due care.

The assigned value is determined by consensus based on the test results of the group of participants after rejection of the statistical outliers and/or suspect data.

According to ISO13528 all (original received or corrected) results per determination were submitted to outlier tests. In the iis procedure for proficiency tests, outliers are detected prior to calculation of the mean, standard deviation and reproducibility. For small data sets, Dixon (up to 20 test results) or Grubbs (up to 40 test results) outlier tests can be used. For larger data sets (above 20 test results) Rosner's outlier test can be used. Outliers are marked by D(0.01) for the Dixon's test, by G(0.01) or DG(0.01) for the Grubbs' test and by R(0.01) for the Rosner's test. Stragglers are marked by D(0.05) for the Dixon's test, by G(0.05) or DG(0.05) for the Grubbs' test and by R(0.05) for the Rosner's test. Both outliers and stragglers were not included in the calculations of averages and standard deviations.

For each assigned value, the uncertainty was determined in accordance with ISO13528. Subsequently the calculated uncertainty was evaluated against the respective requirement based on the target reproducibility in accordance with ISO13528. In this PT, the criterion of ISO13528, paragraph 9.2.1. was met for all evaluated tests, therefore, the uncertainty of all assigned values may be negligible and need not be included in the PT report.

Finally, the reproducibilities were calculated from the standard deviations by multiplying them with a factor of 2.8.

3.2 GRAPHICS

In order to visualize the data against the reproducibilities from literature, Gauss plots were made, using the sorted data for one determination (see appendix 1). On the Y-axis the reported test results are plotted. The corresponding laboratory numbers are on the X-axis. The straight horizontal line presents the consensus value (a trimmed mean). The four striped lines, parallel to the consensus value line, are the +3s, +2s, -2s and -3s target reproducibility limits of the selected reference test method. Outliers and other data, which were excluded from the calculations, are represented as a cross. Accepted data are represented as a triangle.

Furthermore, Kernel Density Graphs were made. The Kernel Density Graph is a method for producing a smooth density approximation to a set of data that avoids some problems associated with histograms. Also, a normal Gauss curve (dotted line) was projected over the Kernel Density Graph (smooth line) for reference. The Gauss curve is calculated from the consensus value and the corresponding standard deviation.

3.3 Z-SCORES

To evaluate the performance of the participating laboratories the z-scores were calculated. As it was decided to evaluate the performance of the participants in this proficiency test (PT) against the literature requirements (derived from e.g. ISO or ASTM test methods), the z-scores were calculated using a target standard deviation. This results in an evaluation independent of the variation of this interlaboratory study.

The target standard deviation was calculated from the literature reproducibility by division with 2.8. In case no literature reproducibility was available, other target values were used, like Horwitz or an estimated reproducibility based on former iis proficiency tests.

When a laboratory did use a test method with a reproducibility that is significantly different from the reproducibility of the reference test method used in this report, it is strongly advised to recalculate the z-score, while using the reproducibility of the actual test method used, this in order to evaluate whether the reported test result is fit-for-use.

The z-scores were calculated according to:

$z_{(target)}$ = (test result - average of PT) / target standard deviation

The $z_{(target)}$ scores are listed in the test result tables in appendix 1.

Absolute values for z<2 are very common and absolute values for z>3 are very rare. Therefore, the usual interpretation of z-scores is as follows:

	|z|	< 1	good
1 <	|z|	< 2	satisfactory
2 <	|z|	< 3	questionable
3 <	|z|		unsatisfactory

4 EVALUATION

In this proficiency test no problems were encountered with the dispatch of the samples. One participant reported test results after the final reporting date.
In total 24 participants reported 23 numerical test results. Observed were 2 outlying test results, which is 8.7%. In proficiency tests outlier percentages of 3% - 7.5% are quite normal.

The original data set proved to have a normal Gaussian distribution.

4.1 EVALUATION PER TEST

In this section the reported test results are discussed per test. The test methods which were used by the various laboratories were taken into account for explaining the observed differences when possible and applicable. These test methods are also in the tables together with the original data in appendix 1. The abbreviations used in these tables are explained in appendix 4.

Test methods ISO17234-1 is considered to be the official test methods for the determination of Aniline in leather. Regretfully, ISO17234-1 does not provide precision data for Aniline. Therefore, the calculated reproducibility was compared against the estimated reproducibility calculated with the Horwitz equation.

Aniline: This determination may be problematic. Two statistical outliers were observed. The calculated reproducibility after rejection of the statistical outliers is not in agreement with estimated reproducibility calculated with the Horwitz equation.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility estimated from the target test method and the reproducibility as found for the group of participating laboratories. The number of significant test results, the average, the calculated reproducibility (2.8 * standard deviation) and the target reproducibility estimated using the Horwitz equation are presented in the next table.

Component	unit	n	average	2.8 * sd	R(target)
Aniline	mg/kg	21	81.0	33.2	18.7

Table 3: reproducibility of component on sample #22512

Without further statistical calculations, it can be concluded that the group of participating laboratories may have difficulties with the determination of the Aniline. See also the discussions in paragraphs 4.1 and 5.

4.3 OVERVIEW OF THE PROFICIENCY TEST OF FEBRUARY 2022

	February 2022
Number of reporting laboratories	24
Number of numerical test results	23
Number of statistical outliers	2
Percentage of statistical outliers	8.7%

Table 4: overview of the proficiency tests

In proficiency tests outlier percentages of 3% - 7.5% are quite normal.

The performance of the determination of the proficiency test was compared, expressed as relative standard deviation (RSD) of the PT, see next table.

Component	February 2022
Aniline	15%

Table 5: observed uncertainty in PT of 2022

4.4 EVALUATION OF THE ANALYTICAL DETAILS

The reported analytical details are listed in appendix 2. The following can be summarized:
- Seventeen of the twenty-three reporting participants reported to be accredited for the determination of Aniline in Leather.
- About two-third of the participants reported to have used 0.5 grams for intake and the rest reported to have used 1 gram.

The effect of sample intake on the determination of Aniline in Leather/Footwear is not significant.

5 DISCUSSION

Almost all reporting participants were able to detect Aniline. The aniline content is expressed as the sum of cleavable Aniline and if appropriate also as chemical residue present free Aniline.

When the results of this interlaboratory study were compared to the LEATHER STANDARD by OEKO-TEX® and with the similar bluesign® systems substances list or BSSL (Table 6), it was noticed that not all participants would make an identical decision about the acceptability of the sample for the determined component.
Compared to the bluesign® systems substances list or BSSL one reporting laboratory would have accepted the sample, while all other laboratories would have rejected the sample.
Based on the LEATHER STANDARD by OEKO-TEX® three of the reporting laboratories would reject the sample, while all other laboratories would have accepted the sample.

Ecolabel	Limit for all classes
bluesign® BSSL	<30 mg/kg
Leather by OEKO-TEX®	<100 mg/kg

Table 6: bluesign® BSSL and LEATHER STANDARD by OEKO-TEX

6 CONCLUSION

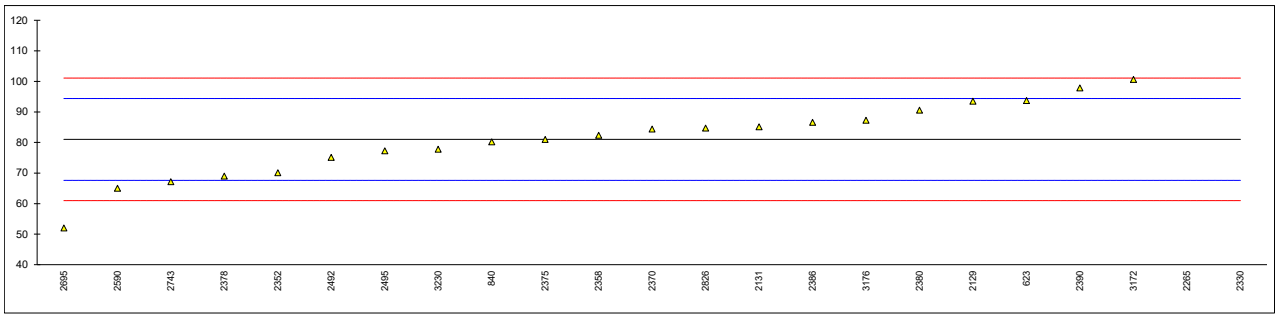
Although it can be concluded that almost all participants have no problem with the determination of Aniline in the sample of this PT, each participating laboratory will have to evaluate its performance in this study and decide about any corrective actions if necessary. Therefore, participation on a regular basis in this scheme could be helpful to improve the performance and thus increase of the quality of the analytical results.

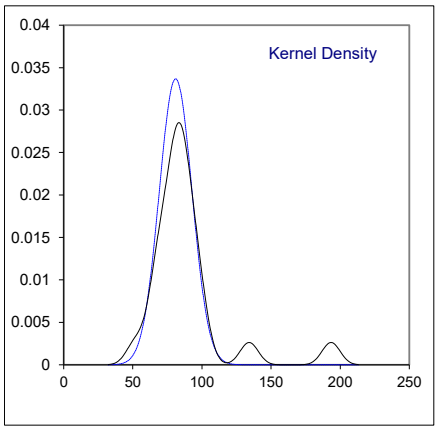
APPENDIX 1

Determination of Aniline on sample #22512; results in mg/kg

lab	method	value	mark	z(targ)	remarks
623	ISO17234-1	93.756		1.90	
840	ISO17234-1	80.2		-0.12	
2129	ISO17234-1 Annex D	93.5		1.86	
2131	In house	85.13		0.61	
2265	ISO17234-1	134.2	R(0.01)	7.95	
2330	ISO17234-1	193.4068	C,R(0.01)	16.80	First reported 165.488
2352	ISO17234-1	70.1		-1.63	
2358	ISO17234-1	82.32		0.19	
2370	ISO17234-1	84.4		0.50	
2375	ISO17234-1	81		0.00	
2378	ISO17234-1	69		-1.80	
2380	ISO17234-1	90.60		1.43	
2386	DIN17234-1	86.6		0.83	
2390	ISO17234-1	97.86		2.52	
2492	ISO17234-1	75.11		-0.88	
2495	In house	77.27		-0.56	
2590	In house	65	C	-2.39	First reported 49.5
2695	In house	52.024		-4.33	
2743	ISO/TS17234-1	67.14		-2.08	
2826	ISO17234-1	84.744		0.56	
3172	In house	100.65		2.93	
3176	ISO17234-1	87.25		0.93	
3210	In house	<10		<-10.62	Possibly a false negative test result?
3230	In house	77.8231		-0.48	

normality	OK	
n	21	
outliers	2	
mean (n)	81.023	
st.dev. (n)	11.8452	RSD=15%
R(calc.)	33.166	
st.dev.(Horwitz)	6.6904	
R(Horwitz)	18.733	





APPENDIX 2 Analytical details

lab	ISO/IEC17025 accredited	Sample intake (in grams)	Remarks
623	Yes	1 gram	
840	No	0.5G	
2129	Yes	0,5	
2131	Yes	0.5	
2265	No	0,5g	
2330	Yes	1 gram	
2352	Yes	0.5g	
2358	Yes	1.0 gram	
2370	Yes	0.5 g	
2375	Yes	1g	
2378	No	1G	
2380	No	0.50 g	
2386	Yes	0,5 g	
2390	Yes	0.5g	
2492	Yes	0.5g	
2495	No	0.5	
2590	Yes	1g	
2695	Yes	1	
2743	Yes	0.5 g	
2826	Yes	0.5000g	
3172	---		
3176	Yes	0.5	
3210	No	1 gram	extraction with acetonitrile with ultrasound (60 min at 60°C) Analysis with LC/MS-MS
3230	Yes	0.5g	

APPENDIX 3

Number of participants per country

1 lab in BANGLADESH
1 lab in CAMBODIA
1 lab in FRANCE
3 labs in GERMANY
3 labs in HONG KONG
1 lab in INDONESIA
5 labs in ITALY
1 lab in MAURITIUS
2 labs in P.R. of CHINA
1 lab in PAKISTAN
1 lab in SWITZERLAND
1 lab in TAIWAN
2 labs in TURKEY
1 lab in VIETNAM

APPENDIX 4

Abbreviations

C	= final test result after checking of first reported suspect test result
D(0.01)	= outlier in Dixon's outlier test
D(0.05)	= straggler in Dixon's outlier test
G(0.01)	= outlier in Grubbs' outlier test
G(0.05)	= straggler in Grubbs' outlier test
DG(0.01)	= outlier in Double Grubbs' outlier test
DG(0.05)	= straggler in Double Grubbs' outlier test
R(0.01)	= outlier in Rosner's outlier test
R(0.05)	= straggler in Rosner's outlier test
E	= calculation difference between reported test result and result calculated by iis
W	= test result withdrawn on request of participant
ex	= test result excluded from statistical evaluation
n.a.	= not applicable
n.e.	= not evaluated
n.d.	= not detected
fr.	= first reported
f+?	= possibly a false positive test result?
f-?	= possibly a false negative test result?

Literature

1 iis Interlaboratory Studies, Protocol for the Organisation, Statistics & Evaluation, June 2018
2 ISO5725:86
3 ISO5725 parts 1-6:94
4 ISO13528:05
5 M. Thompson and R. Wood, J. AOAC Int, 76, 926, (1993)
6 W.J. Youden and E.H. Steiner, Statistical Manual of the AOAC, (1975)
7 P.L. Davies, Fr. Z. Anal. Chem, 331, 513, (1988)
8 J.N. Miller, Analyst, 118, 455, (1993)
9 Analytical Methods Committee, Technical Brief, No 4, January 2001
10 P.J. Lowthian and M. Thompson, The Royal Society of Chemistry, Analyst, 127, 1359-1364, (2002)
11 W. Horwitz and R. Albert, J. AOAC Int, 79.3, 589-621, (1996)
12 Bernard Rosner, Percentage Points for a Generalized ESD Many-Outlier Procedure, Technometrics, 25(2), 165-172, (1983)
13 iis memo 2202: Reproducibility of AZO Dyes in Leather/Footwear and Textiles in iis PTs