Results of Proficiency Test
AZO dyes in leather
March 2011

Organised by: Institute for Interlaboratory Studies
Spijkenisse, the Netherlands

Author: dr. R.G. Visser
Correctors: ing. N. Boelhouwer & ing. R.J. Starink
Report: iis11A01L

May 2011

CONTENTS

1 INTRODUCTION 3

2 SET UP 3

2.1 QUALITY SYSTEM 3

2.2 PROTOCOL 3

2.3 CONFIDENTIALITY STATEMENT 3

2.4 SAMPLES 4

2.5 ANALYSES 5

3 RESULTS 5

3.1 STATISTICS 5

3.2 GRAPHICS 6

3.3 Z-SCORES 6

4 EVALUATION 7

4.1 EVALUATION PER SAMPLE 7

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES 7

5 COMPARISON WITH PREVIOUS INTERLABORATORY STUDIES 8

Appendices:

1. Data and statistical results 9
2. Summary of all other reported results 11
3. Number of participants per country 12
4. Abbreviations and literature 13

1 INTRODUCTION

The Institute for Interlaboratory Studies (iis) organizes every year a scheme of proficiency test for banned AZO dyes in leather since 1997, with an exception in 2009.
In this interlaboratory study, 89 laboratories in 21 different countries have participated (see appendix 4). In this report, the results of this proficiency test are presented and discussed.

2 SET UP

The Institute for Interlaboratory Studies in Spijkenisse was the organizer of this proficiency test. Due to lack of suitable materials it was decided to use in this proficiency test only one leather sample. This leather sample was prepared and tested for homogeneity by an accredited third party laboratory.
Participants were requested to report results with one extra figure. These results with an extra figure are preferably used for statistical evaluation. The participants were asked to report the analytical results using the indicated units on the report form.

2.1 QUALITY SYSTEM

The Institute for Interlaboratory Studies in Spijkenisse, the Netherlands, has implemented a quality system based on ISO guide 43, ILAC-G13:2007 and ISO 17043:2010. This ensures 100% confidentiality of participant's data. Also customer's satisfaction is measured on a regular basis by sending out questionnaires.

2.2 PROTOCOL

The protocol followed in the organization was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organization, Statistics and Evaluation' of January 2010 (iis-protocol, version 3.2).

2.3 CONFIDENTIALITY STATEMENT

All data present 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

As in 2010, the Corporate Product Safety Laboratory of Clariant Chemicals (India) Limited in Thane, India, did prepare 500 grams of leather fortified with an AZO dye that would give 2,4-Xylidene upon analysis. After cutting in pieces, the material was mixed thoroughly. Eleven stratified randomly selected samples were tested using DIN 53316 to check the homogeneity of the batch. See the following table for the test results.

	2,4-Xylidine in mg/kg
sample #11020-1	1068
sample #11020-2	1200
sample #11020-3	1134
sample #11020-4	1161
sample #11020-5	1027
sample #11020-6	1104
sample #11020-7	1123
sample #11020-8	1195
sample #11020-9	1148
sample #11020-10	1004
sample #11020-11	1200

table 1: homogeneity test results of subsamples #11020

From the above test results the repeatability was calculated and compared with 0.3 times the corresponding reproducibility of the target method in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	2,4-Xylidine in mg/kg
r (observed)	190
reference method	ISO17234:2010
0.3 x R (reference method)	185

table 2: evaluation of the repeatability of subsamples #11020

The repeatability of the results of homogeneity test for 2,4-Xylidine was in agreement with 0.3 times the reproducibility mentioned in the reference method ISO17234:2010. Therefore, homogeneity of the subsamples was assumed.

Approx. 4 grams of sample #11020 was sent to the participating laboratories on March 9, 2011.

2.5 ANALYSES

The participants were asked to determine the concentrations of 23 forbidden aromatic amines and o-anisidine, applying the analysis procedure that is routinely used in the laboratory. To get comparable results reported a detailed report form, on which the requested amines and the units were pre-printed, was sent together with each sample. Also a letter of instructions was sent along.

3 RESULTS

During four weeks after sample despatch, the results of the individual laboratories were gathered. The original data are tabulated in the appendices of this report. The laboratories are presented by their code numbers.

Directly after the deadline, a reminder fax was sent to those laboratories that had not yet reported. Shortly after the deadline, the available results were screened for suspect data. A result was called suspect in case the Huber Elimination Rule (a robust outlier test, see lit.5) found it to be an outlier. The laboratories that produced these suspect data were asked to check the results. Additional or corrected data are placed under 'Remarks' in the result tables in appendix 1. A list of abbreviations used in the tables can be found in appendix 4.

3.1 STATISTICS

Statistical calculations were performed as described in the report ’iis Interlaboratory Studies: Protocol for the Organization, Statistics and Evaluation’ of January 2010 (iis-protocol, version 3.2)

For the statistical evaluation the *unrounded* (when available) figures were used instead of the rounded results. Results reported as ‘<…’ or ‘>…” were not used in the statistical evaluation.
Before further calculations, the normality of the distribution of the various data sets per determination was checked by means of the Lilliefors-test. In the case of an anormal distribution, the statistical evaluation should be used with care.

According to ISO 5725 (1986 and 1994, lit.7 and 8) the original results per determination were submitted subsequently to Dixon’s and Grubbs' outlier tests. Outliers are marked by D(0.01) for the Dixon’s test, by G(0.01) or DG(0.01) for the Grubbs’ 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. Both outliers and stragglers were not included in the calculations of averages and standard deviations.

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

3.2 GRAPHICS

In order to visualise 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 analysis results are plotted. The corresponding laboratory numbers are under 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 standard. 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. This is a method for producing a smooth density approximation to a set of data that avoids some problems associated with histograms.

3.3 Z-SCORES

To evaluate the performance of the individual participating laboratories the z-scores were calculated. In order to be able to have an objective evaluation of the performance of the individual participants, it was decided to evaluate this performance against the literature requirements. Therefore the z-scores were calculated using a target standard deviation. This target standard deviation was calculated from the literature reproducibility by division with 2.8.
The $z_{(target)}$-scores were calculated according to:

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

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

Absolute values for z<2 are very common and absolute values for z>3 are very rare. 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

During the execution of this proficiency test some reporting problems occurred. Sixteen participants reported test results after the deadline and two participants did not report any test results. Finally, 87 participants did report 81 numerical results. Observed were 5 outlying results, which is 6.2% of the numerical results on 2,4-Xylidine.
In proficiency studies, outlier percentages of 3% - 7.5% are quite normal.
The data set did not proof to have a normal Gaussian distribution and therefore the evaluation should be useed with care.

4.1 EVALUATION PER SAMPLE AND COMPONENT

In this section, the results are discussed per sample. All statistical results reported on the leather sample are summarised in appendix 1 and all other reported results of aromatic amines present are listed in appendix 2.

2,4-Xylidine:

The determination of this aromatic amine at a (high) concentration level of 1200 mg/kg was not problematic. Only one laboratory reported a test result lower than 30 mg/kg. Another five laboratories did not detect any 2,4-Xylidine in sample #11020.
Several laboratories reported the sum of 2,4- and 2,6-Xylidine. This gave no problems as 2,6-Xylidine was not present in sample #11020.
The test results reported by the participants vary from <10 – 1700 mg/kg. The observed reproducibility (622 mg/kg, or 53% relative) is, after rejection of the five statistical outliers, in good agreement with the reproducibility requirement estimated from the reference standard test method ISO 17234-1:2010.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility as declared by the relevant standard methods (references 1 - 4) and the reproducibility as found for the group of participating laboratories.

The number of significant results, the average result, the calculated reproducibility (standard deviation*2.8) and the target reproducibility, derived from an official test method like ISO17234-1:2010 (equal to the reproducibility from LMBG 82.02.3:97) are in the next table.

Parameter	unit	n	Average	2.8 * sd	R(ISO17234)
2,4-Xylidine	mg/kg	76	1180.0	621.5	646.2

table 3: reproducibility of aromatic amine in leather sample #11020

Without further statistical calculations, it can be concluded that for 2,4-Xylidine the group of participating laboratories has no difficulties with the analysis at the concentration level of 1200 mg/kg. See also the discussion in paragraphs 4.1 and 6.

5 COMPARISON WITH PREVIOUS INTERLABORATORY STUDIES

The spread in the results of 2,4-Xylidine is small in comparison with the spreads of the other aromatic amines as observed in previous PTs, but somewhat larger than in 2010, see below table.

Parameter	*March 2011*	*March 2010*	*March 2008*	*March 2007*	*April 2006*	*April 2005*	*LMBG 82.02.3:97*
4-Aminodiphenyl	n.a.	n.a.	n.a.	n.a.	127%	n.a.	Unknown
Benzidine	n.a.	n.a.	105%	126%	133%	184%	43 – 69%
3,3-Dimethylbenzidine	n.a.	n.a.	n.d.	126%	n.a.	155%	42 – 66%
o-Toluidine	n.a.	n.a.	140%	n.a.	n.a.	175%	84– 103%
2,4-Xylidine	53%	44%	n.a.	n.a.	n.a.	n.a.	n.a.

table 4: development of relative reproducibilities over the years

6 DISCUSSION

From the reported test methods it appeared that most participants treated the leather samples according identical test methods: ISO17234-1, ISO/TS17234, EN14362-1 or LFBG 82.02.3.

One laboratory remarked that it did not find a stable test result. It found six different test results in the range of 66.77 – 439.60 mg/kg.
One other laboratory did prepare a series of recovery samples and found that higher concentrations did show lower recoveries. And correction of the test result for the recovery thus may result in large uncertainty.
In spite of the above reported problems, the observed spread of the test results is in good agreement with the estimated target spread, based on ISO17234-1. However, the observed spread is large in comparison with the estimated target spread based on the Horwitz equation. And from the graphic presentations on page 10, it may be concluded that still some quality improvement may be gained. The standard deviation of the major peak in the Kernel Density plot (representing 88% of the reported test results) is estimated to be 190 mg/kg, which is smaller than the current standard deviation of 222 mg/kg.

It can be concluded that the spread observed in this interlaboratory study is not caused by just one critical point in the analysis. Each participating laboratory will have to evaluate its performance in this study and decide about any corrective actions if necessary.

APPENDIX 1

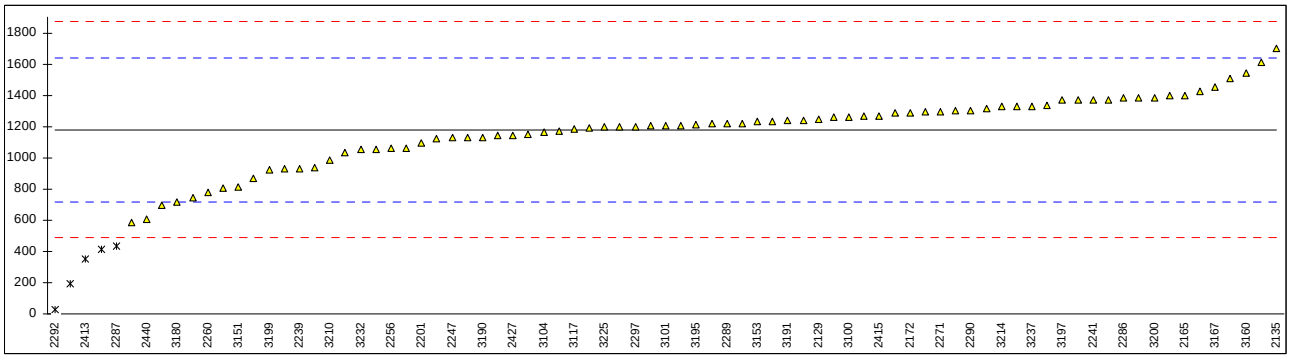
Determination of 2,4-Xylidine (CASno.95-68-1) in sample #11020; results in mg/kg

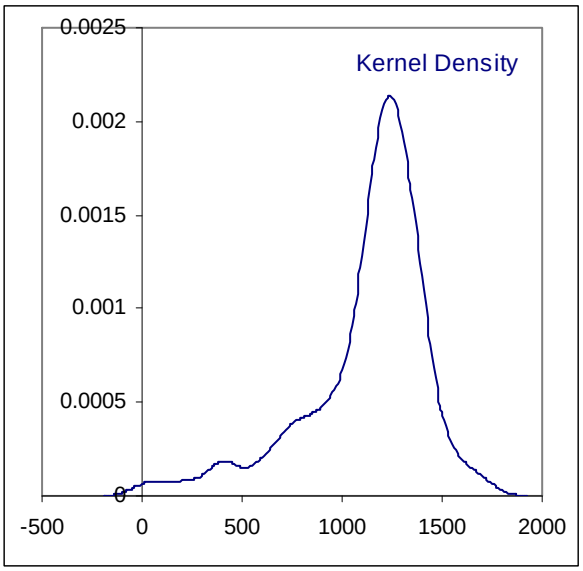
lab	method	value	mark	z(targ)	remarks
2102		-----		-----	
2115		1609		1.86	
2129	EN17234-1	1248		0.29	
2131		-----		-----	reported 845.34 mg/kg 2,6-Xylidene in stead
2132	ISO17234-1	1427.0		1.07	
2135		1700		2.25	
2160		-----		-----	
2165	ISO17234-1	1400		0.95	
2166	ISO17234-1	1315.87		0.59	
2172	EN17234-1	1290		0.48	
2173		1188		0.03	
2184	ISO/TS17234	1200		0.09	
2190	ISO/TS17234	<10		-----	false negative?
2196	ISO/TS17234	1239	C	0.26	first reported 685
2201	ISO/TS17234	1096.9		-0.36	
2215	ISO17234	1058.0		-0.53	
2218	ISO17234	414.02	DG(0.05)	-3.32	
2221	ISO17234-1	1208		0.12	
2235	ISO17234-1	929.2	C	-1.09	first reported 590.4
2239		930.2		-1.08	
2241	ISO17234	1369.6		0.82	
2246		1169.94		-0.04	
2247	ISO/TS17234	1127.0		-0.23	
2255	ISO/TS17234	1219		0.17	
2256	EN17234-1	1057		-0.53	
2260	ISO17234-1	778.1		-1.74	
2261	GB/T 19942	864		-1.37	
2266	ISO/TS17234	803.0		-1.63	
2271		1293		0.49	
2272	ISO17234	698		-2.09	
2275	ISO/TS17234	1385		0.89	
2284	ISO17234	1337.8		0.68	
2286	ISO17234-1	1383		0.88	
2287	ISO17234-1	433	G(0.05)	-3.24	
2289		1220		0.17	
2290	ISO17234-1	1302.0		0.53	
2292	ISO17234-1	29.37	G(0.05)	-4.99	
2295	ISO17234-1	1130		-0.22	
2297	ISO17234	1200		0.09	
2405	ISO17234-1&2	742.7		-1.90	laboratory noted that recovery was dependent of concentration
2410	ISO/TS17234	1120		-0.26	
2413		349.68	DG(0.05)	-3.60	
2415	ISO17234	1268.2737		0.38	
2419		-----		-----	false negative? 10 other aromatic amines were reported
2421		-----		-----	false negative? '-' was reported for all aromatic amines
2427	ISO17234	1143.41		-0.16	
2432	in house	1204		0.10	reported 1204 mg/kg for the sum of 2,4- and 2,6-Xylidine
2436	ISO/TS17234	1508.9		1.43	
2440	GB/T 19942	602.38		-2.50	
3100		1261.3		0.35	
3101	ISO17234-1	1207.4		0.12	
3104		1165		-0.07	
3116	ISO17234-1	1292		0.49	
3117	ISO17234	1181		0.00	
3134		1288.9		0.47	
3146		1230		0.22	
3150	LFGB B82.02-3	1370		0.82	
3151	ISO17234	812		-1.59	
3153	ISO/TS17234	1230		0.22	
3154	ISO17234-1	191.86	G(0.05)	-4.28	laboratory did find results between 66.77 and 439.60 mg/kg
3159	ISO/TS17234	1300		0.52	
3160	ISO17234-1	1541.82		1.57	
3163	GC/MS-therm. des.	-----		-----	false negative? 77 mg/kg o-toluidine was reported
3167	ISO/TS17234	1450		1.17	
3172	ISO17234-1	1326.5		0.63	
3176		938.8		-1.05	reported 938.8 mg/kg for the sum of 2,4- and 2,6-Xylidine
3180		716.25		-2.01	
3182	ISO/TS17234	1032.190		-0.64	
3185	ISO17234-1	1150		-0.13	
3190	ISO17234-1	1130.9		-0.21	

3191	ISO17234-1	1237		0.25	
3195	EN14362-1	1209.49		0.13	
3197	ISO17234	1367		0.81	
3199	CPSD-AN-00017	919.9		-1.13	
3200	ISO17234-1	1386.8		0.90	
3204		positive		-----	detected 2,4-Xylidine, but did not quantify this component
3209	ISO17234-1	1260.1		0.35	
3210	ISO17234-1	983.5		-0.85	
3214	ISO17234-1	1325.7		0.63	
3218	ISO17234-1	1368		0.81	
3220	ISO17234-1	1140.0		-0.17	
3225	ISO17234-1	1198		0.08	
3228		1400		0.95	
3232	DIN53316	1050		-0.56	
3233	ISO/TS17234	1052	C	-0.55	first reported 703.14
3237	ISO17234	1326.51		0.63	
3246		1221		0.18	
3247	ISO17234-1	584		-2.58	
3248		1263.9		0.36	

	normality	not OK
	n	76
	outliers	5
	mean (n)	1180.029
	st.dev. (n)	221.9774
	R(calc.)	621.537
	R(ISO17234-1)	646.160

Compare R(Horwitz) = 182.304





APPENDIX 2

Summary of all other reported AZO-dyes in sample #11020; results in mg/kg

lab	aromatic amines
2115	6.45 mg/kg aniline
2131	845.34 mg/kg 2.6-Xylidine
2419	1180 mg/kg 4-aminodiphenyl, 12.10 mg/kg benzidine, 1.66 mg/kg 4-chloro-o-toluidine, 1.88 mg/kg 2-naphthylamine, 1.88 mg/lg p-chloraniline, 2.62 mg/kg 2,4-diaminoanisol, 3.22 mg/kg 3,3'-dimethyl-4,4'-diaminodiphenylmethane, 3.88 mg/kg p-cresidine, 5.14 mg/kg 4,4'-diaminodiphenylether & 2.67 mg/kg 4-aminoazobenzene
3163	77 mg/kg o-toluidine
3237	203.49 mg/kg 2,6-Xylidine

APPENDIX 3

Number of participants per country

1 lab in	BANGLADESH
1 lab in	BULGARIA
1 lab in	CYPRUS
4 labs in	FRANCE
8 labs in	GERMANY
1 lab in	GREECE
12 labs in	HONG KONG
3 labs in	INDIA
4 labs in	ITALY
2 labs in	JAPAN
1 labs in	KOREA
32 labs in	P.R. of CHINA
1 lab in	SLOVENIA
2 labs in	SPAIN
2 labs in	SWITZERLAND
1 lab in	TAIWAN R.O.C.
1 lab in	THAILAND
2 labs in	THE NETHERLANDS
5 labs in	TURKEY
2 labs in	U.S.A.
3 labs in	VIETNAM

APPENDIX 4

Abbreviations:

C = final result after checking of first reported suspect 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
n.a. = not applicable
n.d. = not detected

Literature:

1 DIN 53316
2 ISO 17234:2010
3 LMBG 82.02-3:97
4 LMBG 82.04-2:98
5 EN14362-1/2, March 2002
6 Staatsblad van het Koninkrijk der Nederlanden 339, bijlage II, 23 april 1998
7 iis-Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation, January 2010
8 XP G 08-014:97
9 P.L. Davies, Fr Z. Anal. Chem, 351, 513, (1988)
10 W.J. Conover, Practical; Nonparametric Statistics, J. Wiley&Sons, NY, p.302, (1971)
11 ISO 5725, (1986)
12 ISO 5725, parts 1-6, (1994)
13 M. Thompson and R. Wood, J. AOAC Int, 76, 926, (1993)
14 G. Rohm, J. Bohnen & H. Kruessmann, GIT Labor-Fachzeitschrift, p 1080, 11, (1997)