Results of Proficiency Test
Bisphenol A EN71-10/11
June 2021

Organized by: Institute for Interlaboratory Studies
Spijkenisse, the Netherlands

Author: ing. A. Ouwerkerk
Correctors: ing. R.J. Starink & ing. G.A. Oosterlaken-Buijs
Report: iis21V06

October 2021

CONTENTS

1 INTRODUCTION

Toy safety is the practice of ensuring that toys, especially those made for children, are safe, usually through the application of safety standards. In many countries toys must be able to pass safety tests in order to be sold. Many regions model their safety standards on the EU's EN71 standard. In Europe toys must meet the criteria set by the 2009 EC Toy Safety Directive 2009/48/EC, last updated on 21th of May 2021.

Migration of Bisphenol A (BPA) is described in EN 71-9 (Requirements), EN 71-10 (Sample preparation and extraction) and EN 71-11 (Methods of Analysis). The maximum specific limit as described in EN 71-9 is 0.1 mg/L aqueous substrate (or simulant). The European Union has further restricted this limit when it comes to toys. EU directive 2017/898 of 24 May 2017 amending Appendix C to Annex II to Directive 2009/48/EC as regards BPA describes a maximum specific migration limit of 0.04 mg/L aqueous substrate (or simulant). This has been implemented from November 26, 2018 in its member states.

Since 2017 the Institute for Interlaboratory Studies (iis) organizes a proficiency scheme for the analysis of migratable Bisphenol A by EN71-10/11 every year. During the annual proficiency testing program 2020/2021 it was decided to continue the proficiency test for the analysis of migratable Bisphenol A.

In this interlaboratory study 24 laboratories in 12 different countries registered for participation. See appendix 3 for the number of participants per country. In this report the results of the Bisphenol A EN71-10/11 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 strip of white thermal paper labelled #21625 positive on Bisphenol A. Furthermore, a number of test conditions (sample size, simulant, exposure temperature, exposure time and rotation speed) were prescribed.
The participants were requested to report rounded and unrounded test results. The unrounded test results were preferably used for 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 white thermal paper positive on BPA was selected. From this batch 68 paperstrips of approximately 5x6cm were wrapped in Aluminum foil and labelled #21625.
The homogeneity of the subsamples was checked by determination of Bisphenol A in accordance with EN71-10/11 on 8 stratified randomly selected subsamples.

	BPA in mg/L
sample #21625-1	1.08
sample #21625-2	1.01
sample #21625-3	0.99
sample #21625-4	1.01
sample #21625-5	1.02
sample #21625-6	0.94
sample #21625-7	0.92
sample #21625-8	1.00

Table 1: homogeneity test results of subsamples #21625

From the above test results the repeatability was calculated and compared with 0.3 times the estimated reproducibility calculated with the Horwitz equation in agreement with the procedure of ISO13528, Annex B2 in the next table.

	BPA in mg/L
r (observed)	0.138
reference method	Horwitz
0.3 x R (reference method)	0.134

Table 2: evaluation of the repeatability of subsamples #21625

The calculated repeatability is in agreement with 0.3 times the estimated reproducibility calculated with the Horwitz equation. Therefore, homogeneity of the subsamples was assumed.

To each of the participating laboratories one sample of thermal paper labelled #21625 was sent on May 19, 2021.

2.5 ANALYZES

The participants were requested to determine Bisphenol A in aqueous migrate using the prescribed test conditions (see table 3).
It was also requested to report if the laboratory was accredited for this determination and to report some analytical details. It was advised to keep the thermal paper stored dark, dry and cool and packed until the start of the test. It was also advised not to touch the sample with bare hands.

Sample size	cut the sample at **width=2cm/length=5cm** (surface area is: 2x5=10 cm^2)*
Simulant	deionized water
Simulant volume	as per method used
Exposure temperature	20°C
Exposure time	1 hour
Rotation speed	60 r/min

Table 3: prescribed test conditions for sample #21625
*) see also paragraph 4.1

It was explicitly requested to treat the sample as if it was a routine sample and to report the test results using the indicated units on the report form and not to round the test results, but report as much significant figures as possible. It was also requested not to report ‘less than’ test results, which are above the detection limit, because such test 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 units are given as well as the reference test methods (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 organisation 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. This 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, e.g. EN reproducibilities, the z-scores were calculated using a target standard deviation. This results in an evaluation independent of the variation in 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:

	$\lvert z \rvert$	< 1	good
1 <	$\lvert z \rvert$	< 2	satisfactory
2 <	$\lvert z \rvert$	< 3	questionable
3 <	$\lvert z \rvert$		unsatisfactory

4 EVALUATION

In this proficiency test no problems were encountered with the dispatch of the samples. Two participants reported test results after the final reporting date and one other participant was not able to report any test results.
In total 23 participants reported 23 numerical test results. Observed was 1 outlying test result, which is 4.3%. In proficiency tests outlier percentages of 3% - 7.5% are quite normal.

The data set proved to have no normal Gaussian distribution and is referred to as “not OK”. The statistical evaluation of this data set should be used with due care, see also paragraph 3.1.

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 table together with the original data. The abbreviation, used in this table, are explained in appendix 4.

Test method EN71-11 does mention precision data, unfortunately only at a low level of 0.03 mg BPA/L aqueous migrate. Therefore, the calculated reproducibility was compared against the estimated reproducibility calculated with the Horwitz equation.

Test method EN 71-10 does not describe whether the sample should be used one-sided or two-sided. Therefore, some test conditions like sample size (width=2 cm and length=5 cm) and surface area (2x5=10 cm^2) was prescribed. However, it was also requested to report the sample size (width and length) and the surface area used for the migration. All test results were evaluated as one-sided exposure as the sample is very thin. One laboratory used 5 cm^2 and 100 mL (resulting in a lower surface to volume ratio of 0.05). This test result was multiplied by two.

BPA (migratable): This determination may be problematic. One statistical outlier was observed and one other test result was excluded. The calculated reproducibility after rejection of the suspect data is not in agreement with the estimated reproducibility calculated from the Horwitz equation.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility as declared by the estimated target reproducibility using the Horwitz equation 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 estimated target reproducibility are presented in the next table.

Parameter	unit	n	average	2.8 * sd	R(target)
Bisphenol A (migratable)	mg/L	21	5.1	2.3	1.8

Table 4: reproducibility on sample #21625

Without further statistical calculations it could be concluded that there is not a good compliance of the group of participating laboratories with the reference method. See also the discussion in paragraph 4.1.

4.3 COMPARISON OF THE PROFICIENCY TEST OF JUNE 2021 WITH PREVIOUS PTS

The performance of the determinations of the proficiency test was compared against the requirements of the reference method, see next table.

Parameter	June 2021	October 2020	December 2019	December 2018	R(target)
Bisphenol A (migratable)	16%	29%	14%	34-39%	12-13%

Table 5: development of uncertainties over the years

The uncertainty observed in this PT is smaller than the uncertainty observed in the PT conducted in 2018 and 2020 and is comparable with the uncertainty of the PT conducted in 2019.

4.4 EVALUATION OF THE ANALYTICAL DETAILS

In this PT also some analytical details were asked, see appendix 2 for the reported details. The majority of the participants (sixteen of the twenty-two participants reporting analytical details) is ISO/IEC17025 accredited for this test.

All participants used 100 mL of simulant. The temperatures used were between 20 and 24°C, the rotation speed per minute was 60 rpm and the time used for the migration was 60 minutes for almost all participants.
Further was observed that almost all laboratories used a test portion of 10 cm^2 as surface area by using the prescribed 2x5 cm sample size.

5 DISCUSSION

In this proficiency test the average of the homogeneity test results is not in line with the average (consensus value) from the PT results. There are several reasons for this.
Firstly, the homogeneity testing is done by one laboratory only. The test results of this (ISO/IEC17025 accredited) laboratory will have a bias (systematic deviation) depending on the test method used. The desire to detect small variations between the PT samples leads to the use of a sensitive test method with high precision, which may be a test method with significant bias.

Secondly, each test result reported by the laboratories that participate in the PT will have a bias. However, some will have a positive bias and others a negative bias. These different biases compensate each other in the PT average (consensus value). Therefore, the PT consensus value may deviate from the average of the homogeneity test. At the same time the accuracy of the PT consensus value is more reliable than the accuracy of the average of the results of the homogeneity test.

6 CONCLUSION

All participants found sample #21625 to be positive on migratable BPA (above the limit of EN71-9 (0.1 mg/L) and directive EU/2017/898 (0.04 mg/L).

It is to be expected that the variation of the migration test results in real life practise will be larger than observed in this PT as the test conditions like sample size, simulant, exposure temperature, exposure time and rotation speed will not be prescribed but will be selected by the individual laboratories.

Each laboratory has to evaluate its performance in this study and make decisions about necessary corrective actions. Therefore, participation on a regular basis in this scheme could be helpful to improve the performance and the quality of the analytical results.

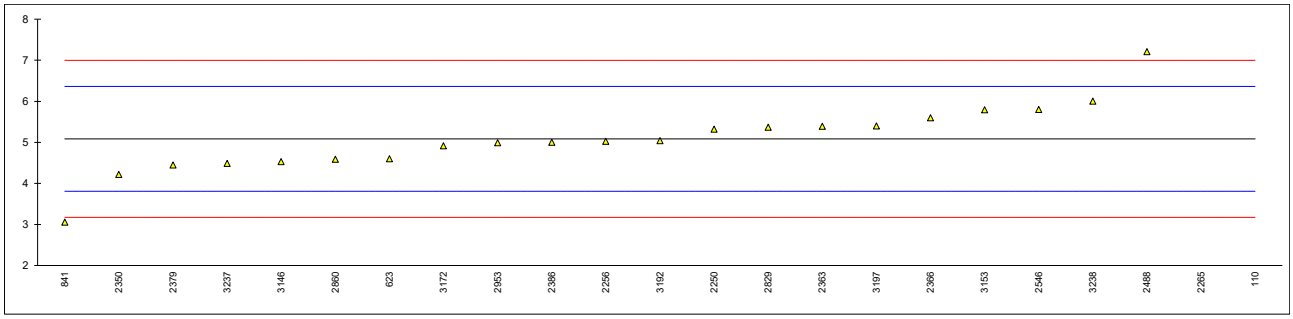
APPENDIX 1

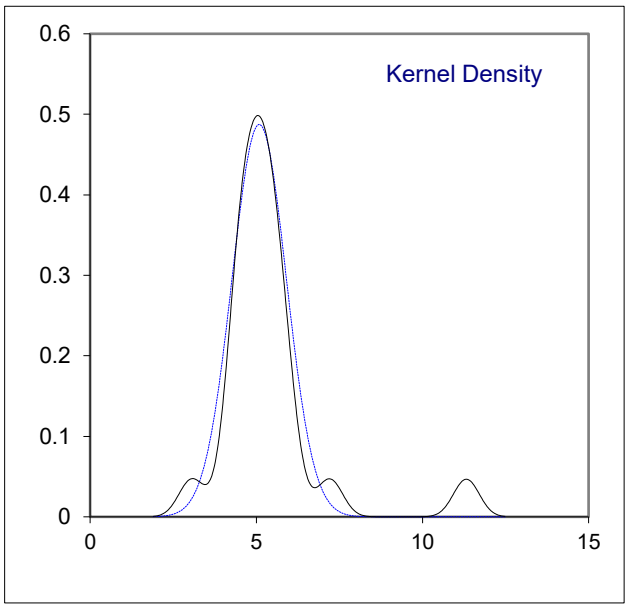
Determination of BPA in aqueous migrate on sample #21625; (1-sided surface) results in mg/L

lab	method	value	mark	z(targ)	remarks
110	In house	5897	ex	9248.42	Reported total BPA in stead of migratable
339		-----		-----	
623	EN71-11	4.60		-0.76	
841	EN71-10/11	3.060		-3.18	
2250	EN71-11	5.32		0.37	
2256	EN71-11	5.0238		-0.10	
2265	EN71-11	11.32	R(0.01)	9.79	
2350	EN71-11	4.22		-1.36	
2363	EN71-11	5.39		0.48	
2366	EN71-11	5.6		0.81	
2379	EN71-11	4.45		-1.00	
2386	EN71-11	5.005		-0.13	
2488	EN71-11	7.21	C	3.33	First reported 15.0
2546	EN71-11	5.80086		1.12	
2829	EN71-11	5.37		0.45	
2860	EN71-11	4.59		-0.78	
2953	EN71-11	4.99		-0.15	
3146	EN71-11	4.53		-0.87	
3153	EN71-11	5.795		1.11	
3172	EN71-11	4.92	C	-0.26	Reported 2.46 for a used surface to volume ratio of 0.05 cm^2/mL
3192	EN71-11	5.041		-0.07	
3197	EN71-11	5.4		0.49	
3237	EN71-11	4.49		-0.94	
3238	EN71-11	6.004		1.44	

	normality	not OK
	n	21
	outliers	1+1ex
	mean (n)	5.086
	st.dev. (n)	0.8189
	R(calc.)	2.293
	st.dev.(Horwitz)	0.6371
	R(Horwitz)	1.784

RSD=16%





APPENDIX 2 Analytical details

lab	ISO/IEC 17025 accredited	length test portion (cm)	width test portion (cm)	surface area migration (cm^2)	volume simulant migration (mL)	surface to volume ratio calc. by iis	temp. simulant (°C)	rotation speed (r/min)	time used migration (min)
110	---	---	---	---	---	---	---	---	---
339	---	---	---	---	---	---	---	---	---
623	Yes	5	2	10	100	0.1	20	60	60
841	Yes	5.065	2.065	10.459	100	0.1	20	60	60
2250	Yes	5	2	10	100	0.1	RT*)	---	60 min
2256	Yes	5.061	2.106	10.6585	100.0	0.1	20	60	60
2265	No	5	2	10	100	0.1	22.2	60	60
2350	Yes	5	2	10	100	0.1	20	60	60
2363	Yes	5	2	10	100	0.1	20	60	60
2366	Yes	5	2	10	100	0.1	RT*)	60	60
2379	No	5	2	10	100	0.1	20	60	60
2386	Yes	5.00	2.00	10.00	100	0.1	20.0	60	60
2488	Yes	5.1	2	10.2	100	0.1	20	60	60
2546	Yes	5.0	2.0	10.0	100	0.1	20	56	60
2829	No	5	2	10	100	0.1	21	60	60
2860	No	5.2	1.9	9.88	100	0.1	24	60	60
2953	No	3.2	3.2	10.24	100	0.1	22	60	60
3146	Yes	2.0	5.0	10.0	100	0.1	24	60	60
3153	Yes	5	2	10	100	0.1	20	60	60
3172	Yes	5	1	10	100	0.05	25	60	60
3192	Yes	5	2	10	100	0.1	20	100	60
3197	Yes	5	2	10	100	0.1	20	60	60
3237	Yes	5	2	10	100	0.1	20	60	60
3238	No	5	2	10	100	0.1	20	60	60

*) RT=Room Temperature

APPENDIX 3

Number of participants per country

2 labs in FRANCE

5 labs in GERMANY

1 lab in HONG KONG

1 lab in INDONESIA

4 labs in ITALY

3 labs in P.R. of CHINA

1 lab in SERBIA

1 lab in KOREA, Republic of

1 lab in THAILAND

3 labs in TURKEY

1 lab in U.S.A.

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

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)