

Results of Proficiency Test
LIQUEURS
November 2011

Organised by: Institute for Interlaboratory Studies (iis)
Spijkenisse, the Netherlands

Authors: ing. L.Dijkstra
Correctors: dr. R.G. Visser & ing. N. Boelhouwer
Report: iis11C13b

February 2012

CONTENTS

1	INTRODUCTION	3
2	SET UP	3
2.1	ACCREDITATION	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 TEST	7
4.2	PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES	9
4.3	COMPARISON OF THE PROFICIENCY TEST OF NOVEMBER 2011 WITH PREVIOUS PT'S	9

Appendices:

1.	Data and statistical results	11
2.	Analytical details.....	15
3.	Number of participants per country	16
4.	Abbreviations and literature	17

1. INTRODUCTION

Since 2009, a proficiency test for liqueurs is organised every year by the Institute for Interlaboratory Studies. During the planning of the annual proficiency testing program 2011/2012, it was decided to continue the proficiency test for the analysis of liqueurs. In this interlaboratory study, 13 laboratories in 5 different countries have participated. See appendix 2 for a list of number of participants per country. In this report, the results of the proficiency test are presented and discussed.

2 SET-UP

The Institute for Interlaboratory Studies (iis) in Spijkenisse, the Netherlands, was the organiser of this proficiency test. Analysis for fit-for-use and homogeneity testing were subcontracted. It was decided to send two different samples of liqueur (1* 0.5 L of herbal liqueur, labelled #11123 and 1* 0.5 L of chocolate liqueur, labelled #11124). Participants were requested to report rounded and unrounded results. The unrounded 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 on ISO guide 43, ILAC-G13:2007 and ISO17043:2010. This ensures strict adherence to protocols for sample preparation and statistical evaluation and 100% confidentiality of participant's data. Also customer's satisfaction is measured on regular basis by sending out questionnaires.

2.2 PROTOCOL

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' (iis-protocol, version 3.2) of January 2010.

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

The necessary bulk material (Herbal Liqueur) for sample #11123 was obtained from a local producer. The approximately 25 litre bulk sample was, after homogenisation in a precleaned can, divided over 20 amber glass bottles of 0.5 L and labelled #11123. The homogeneity of these subsamples was checked by determination of Density in accordance with ASTM D4052:11 on 6 stratified random selected samples.

Sample	Density @ 20°C in kg/L
Sample #11123-1	1.03911
Sample #11123-2	1.03907
Sample #11123-3	1.03902
Sample #11123-4	1.03908
Sample #11123-5	1.03908
Sample #11123-6	1.03906

table 1: Homogeneity test results of subsamples #11123

From the test results of table 1, the repeatability was calculated and compared with 0.3 times the corresponding target reproducibility in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	Density @ 20°C in kg/L
r (Observed)	0.00008
reference method	D4052:11
0.3 * R (ref. method)	0.00015

table 2: Repeatability of subsamples #11123

The necessary bulk material (Chocolate Liqueur) for sample #11124 was obtained from a local producer. The approximately 30 litre bulk sample was, after homogenisation in a precleaned can, divided over 24 amber glass bottles of 0.5 L and labelled #11124. The homogeneity of these subsamples was checked by determination of Density in accordance with ASTM D4052:11 on 5 stratified random selected samples.

Sample	Density @ 20°C in kg/L
Sample #11124-1	1.11070
Sample #11124-2	1.11074
Sample #11124-3	1.11078
Sample #11124-4	1.11076
Sample #11124-5	1.11079

table 3: Homogeneity test results of subsamples #11124

From the test results of table 3, the repeatability was calculated and compared with 0.3 times the corresponding target reproducibility in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	<i>Density @ 20°C in kg/L</i>
r (Observed)	0.00010
reference method	D4052:11
0.3 * R (ref. method)	0.00015

table 4: Repeatability of subsamples #11124

The repeatabilities of the results from the homogeneity test were in agreement with the requirements of the respective standards. Therefore, homogeneity of all the prepared subsamples was assumed.

To each of the participating laboratories 1*0.5 L bottle of sample #11123 and 1*0.5 L bottle of sample #11124 were sent on December 02, 2011.

2.5 ANALYSES

The participants were asked to determine on sample #11123 and #11124: Density @ 20°C, pH, Sugars (Glucose, Fructose, Sucrose, Maltose) and Strength (in %V/V).

To get comparable results a detailed report form, on which the units were printed, was sent together with each sample. In addition, a letter of instructions and a SDS were added to the package.

3 RESULTS

During four weeks after sample despatch, the results of the individual laboratories were received. The original reported results are tabulated per determination in appendix 1 of this report. The laboratories are presented by their code numbers.

Directly after deadline, a reminder fax was sent to those laboratories that had not yet reported any results at that moment.

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) found it to be an outlier. The laboratories that produced these suspect data were asked to check the results. Additional or corrected results are used for data analysis and original results are placed under 'Remarks' in the result tables in appendix 1.

3.1 STATISTICS

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

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. First the normality of the distribution of the various data sets per determination was checked by means of the Lilliefors-test. After removal of outliers, this check was

repeated. Not all data sets proved to have a normal distribution, in which cases the statistical evaluation should be used with due care.

In accordance with ISO 5725 (1986 and 1994) the original results per determination were submitted subsequently to Dixon and Grubbs outlier tests. Outliers are marked by D(0.01) for the Dixon test, by G(0.01) or DG(0.01) for the Grubbs test. Stragglers are marked by D(0.05) for the Dixon 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.

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. When the uncertainty passed the evaluation, no remarks are made in the report. However, when the uncertainty failed the evaluation it is mentioned in the report and it will have consequences for the evaluation of the test results.

Finally, the reproducibilities were calculated from the standard deviations by multiplying these 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 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 (see appendix 4, nr.13-14).

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. ASTM reproducibilities, the z-scores were calculated using a target standard deviation. This results in an evaluation independent of the spread 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. In some cases, literature repeatability is available; in other cases, a reproducibility of a former iis proficiency test could be used and the Horwitz equation can be used to estimate target reproducibility.

The z-scores were calculated according to:

$$Z_{(\text{target})} = (\text{result} - \text{average of PT}) / \text{target standard deviation}$$

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

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 the fit-for-useness of the reported test result.

4. EVALUATION

In this proficiency test no problems were encountered with despatch of the samples. Two participants reported results after the final reporting date. Not all laboratories were able to perform all analysis requested. Finally, the 13 reporting laboratories did send in 67 (numerical) results. Observed were 5 outlying results, which is 7.5%. In proficiency studies, outlier percentages of 3% - 7.5% are normal.

4.1 EVALUATION PER TEST

In this section, the results are discussed per test.

The methods, which are used by the various laboratories, are taken into account for explaining the observed differences when possible and applicable. These methods are also in the tables together with the original data. The abbreviations, used in these tables, are listed in appendix 3.

Regretfully, only two participants reported results for sugars. Therefore no statistical conclusions could be drawn on the determinations for Glucose, Fructose, Sucrose and Maltose.

On the registration form the participants were asked to fill out the analytical details regarding the strenght determination. Nine laboratories answered the questions fully or partially. (See Appendix 2).

A not normal distribution was found for the following determinations: Density (#11123 and #11124). In this case the statistical evaluation should be used with due care.

Density: This determination was not problematic for the herbal liqueur #11123, but it was problematic for the chocolate liqueur #11124. In total two statistical outliers were observed. The calculated reproducibility for sample #11123,

after rejection of the statistical outlier, is in agreement, while the calculated reproducibility of sample #11124 is not in agreement with the requirements of ASTM D4052:11.

pH: This determination was not problematic for the tested samples. One statistical outlier was observed and both calculated reproducibilities are in agreement with the requirements of EN15490:07.

Strength (%V/V): Regretfully, no standard test method with precision data exists for this determination.

From the analytical details it is clear all participants did perform a distillation before the strength determination, except one. The two observed outliers belong to this laboratory.

When compared with the calculated reproducibilities of the previous proficiency test iis10C12c, the spread found for sample #11123 is much smaller (0.150 vs 0.309) than the spread on a similar sample in the previous PT, while the spread for sample #11124 is a little smaller (0.110 vs 0.148) than the spread on a similar sample in the previous PT.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility as declared by the relevant standard and the reproducibility as found for the group of participating laboratories. The average results per sample, calculated reproducibilities and reproducibilities derived from literature standards (in casu ASTM, EN standards) or previous proficiency tests are compared in the next table.

Parameter	Unit	n	average	2.8 *sd _R	R (lit)
Density @ 20°C	kg/L	11	1.03912	0.00010	0.00050
pH		6	4.20	0.31	0.69
Strength	%V/V	9	29.99	0.15	(0.31)
Glucose	%M/M	2	n.a	n.a	n.a
Fructose	%M/M	2	n.a	n.a	n.a
Sucrose	%M/M	1	n.a	n.a	n.a
Maltose	%M/M	1	n.a	n.a	n.a

Table 5: Reproducibilities of sample #11123

Parameter	Unit	n	average	2.8 *sd _R	R (lit)
Density @ 20°C	kg/L	10	1.11081	0.00082	0.00050
pH		7	7.03	0.15	0.69
Strength	%V/V	8	13.89	0.11	(0.15)
Glucose	%M/M	2	n.a	n.a	n.a
Fructose	%M/M	2	n.a	n.a	n.a
Sucrose	%M/M	1	n.a	n.a	n.a
Maltose	%M/M	2	n.a	n.a	n.a

Table 6: Reproducibilities of sample #11124

results between brackets are compared with the spread of the previous proficiency test or estimated from target reproducibility.

4.3 COMPARISON OF THE PROFICIENCY TEST OF NOVEMBER 2011 WITH PREVIOUS PT'S

	<i>November 2011</i>	<i>November 2010</i>	<i>December 2009</i>
Number of reporting labs	13	17	23
Number of results reported	67	71	92
Statistical outliers	5	8	11
Percentage outliers	7.5%	11.3%	12.0%

table 7: comparison with previous proficiency tests.

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

The performance of the determinations of the proficiency tests was compared against the requirements of the respective standards. The conclusions are given the following table:

Parameter	<i>November 2011</i>	<i>November 2010</i>	<i>December 2009</i>
Herbal liqueur			
Density @ 20°C	++	--	++
pH	++	++	++
Strength	++	--	--
Chocolate liqueur			
Density @ 20°C	--	++	--
pH	++	++	+/-
Strength	+	+/-	--

Table 8: comparison determinations against the standard

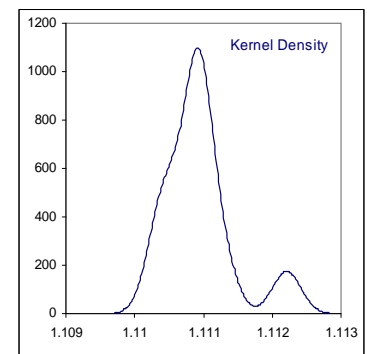
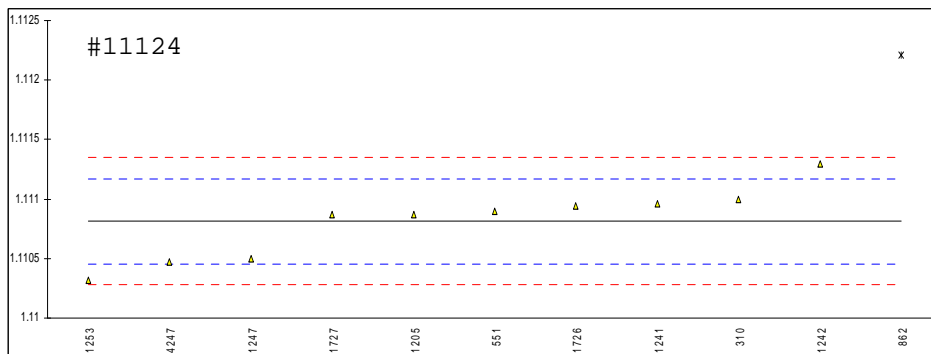
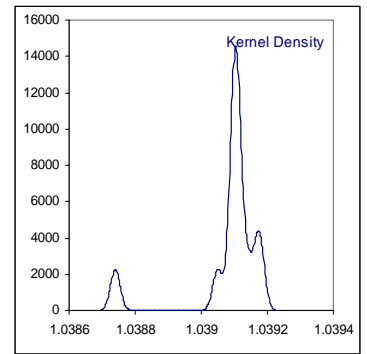
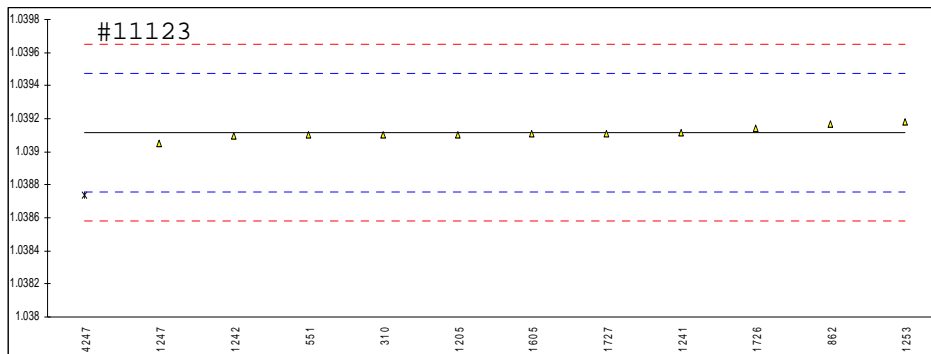
The performance of the determinations against the requirements of the respective standards is listed in the above table. The following performance categories were used:

- ++: group performed much better than the standard
- + : group performed better than the standard
- +/-: group performance equals the standard
- : group performed worse than the standard
- : group performed much worse than the standard
- n.e.: not evaluated

APPENDIX 1

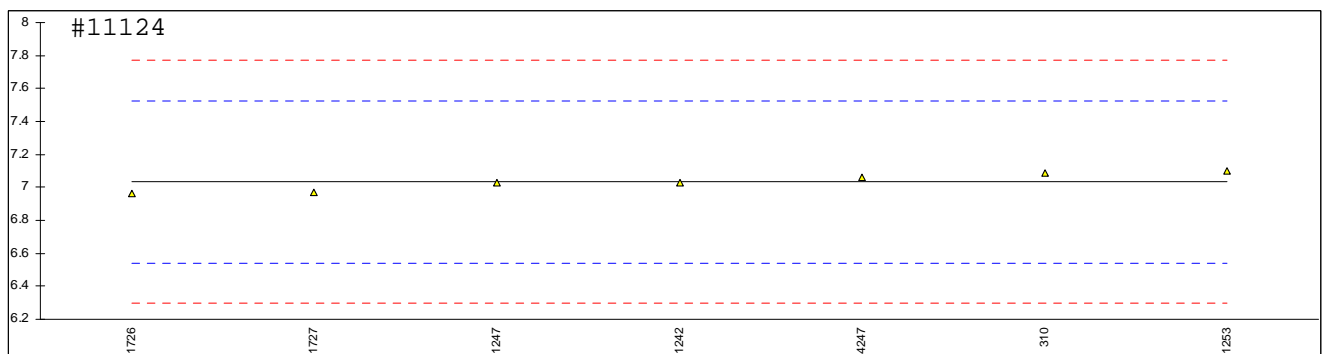
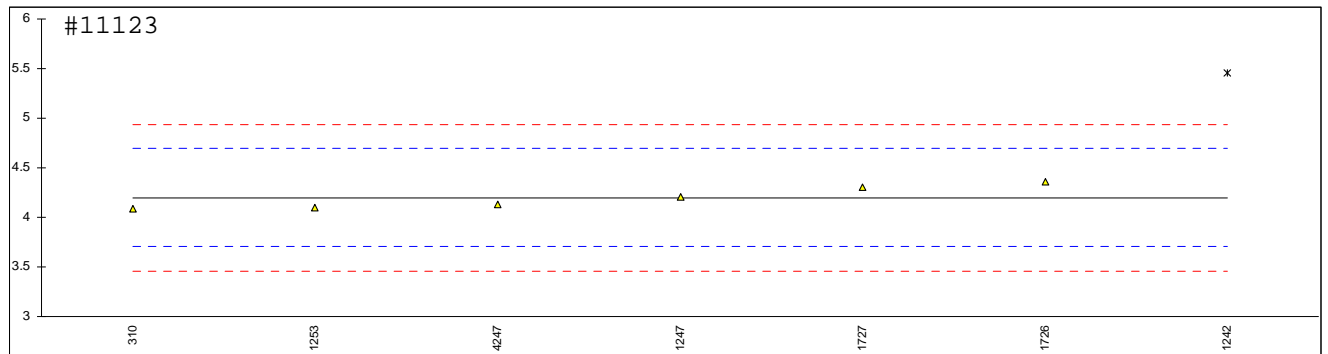
Determination of Density @ 20°C on sample #11123 and #11124; results in kg/L.

lab	method	#11123	mark	z(targ)	#11124	mark	z(targ)	remarks
310	D4052	1.0391		-0.09	1.1110		1.05	
551	D4052	1.0391		-0.09	1.1109		0.49	
862	D4052	1.03917		0.31	1.11221	G(0.05)	7.83	
1126		-----		-----	-----		-----	
1205	in house	1.03910		-0.09	1.11087		0.32	
1241	D4052	1.039113		-0.01	1.110957		0.81	
1242	D4052	1.039096		-0.11	1.111296	C	2.71	first reported: 1.039096
1247	INH-4500	1.03905		-0.37	1.11050		-1.75	
1253	D4052	1.03918		0.36	1.11032		-2.76	
1605	D4052	1.03911		-0.03	-----		-----	
1726	D4052	1.03914		0.14	1.11094		0.72	
1727	D4052	1.03911		-0.03	1.11087		0.32	
4247	D4052	1.03874	G(0.01)	-2.10	1.11047		-1.92	
normality		not OK		not OK				
n		11		10				
outliers		1		1				
mean (n)		1.03912		1.11081				
st.dev. (n)		0.000036		0.000294				
R(calc.)		0.00010		0.00082				
R(D4052:02e1)		0.00050		0.00050				



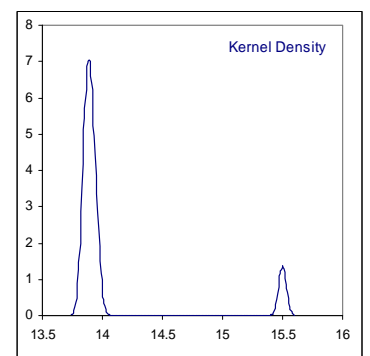
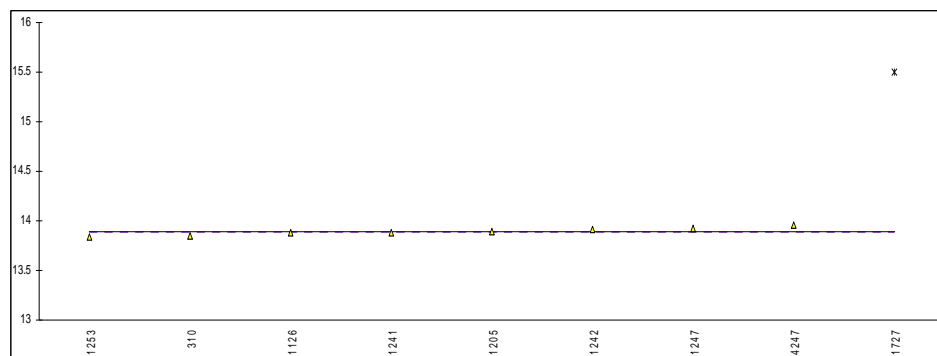
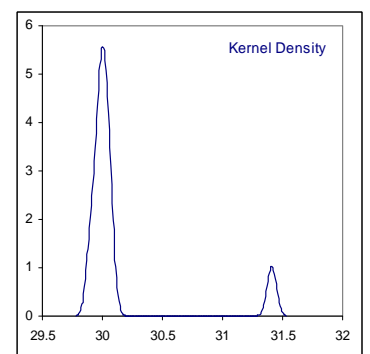
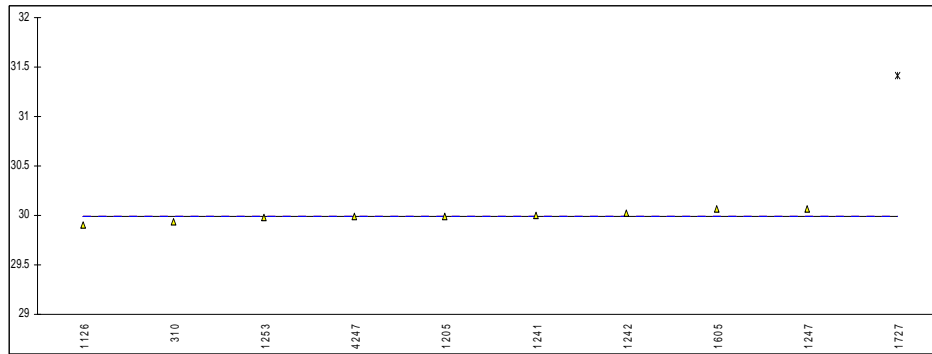
Determination of pH on sample #11123 and #11124

lab	method	#11123	mark	z(targ)	#11124	mark	z(targ)	remarks
310	EN15490	4.09		-0.44	7.09		0.23	
551		----		----	----		----	
862		----		----	----		----	
1126		----		----	----		----	
1205		----		----	----		----	
1241		----		----	----		----	
1242	EN15490	5.46	G(0.01)	5.12	7.03		-0.02	
1247	EN15490	4.21		0.05	7.03		-0.02	
1253	EN15490	4.1		-0.40	7.1		0.27	
1605		----		----	----		----	
1726	EN15490	4.36		0.66	6.96		-0.30	
1727	EN15490	4.30		0.41	6.97		-0.26	
4247	EN15490	4.13		-0.28	7.06		0.10	
normality		OK			OK			
n		6			7			
outliers		1			0			
mean (n)		4.198			7.034			
st.dev. (n)		0.11197			0.0544			
R(calc.)		0.314			0.152			
R(EN15490:07)		0.690			0.690			



Determination of Strength on sample #11123 and #11124; results in %V/V

lab	method	#11123	mark	z(targ)	#11124	mark	z(targ)	remarks
310		29.93		----	13.85		----	
551		----		----	----		----	
862		----		----	----		----	
1126		29.90		----	13.88		----	also reported: 29.74 and 13.84
1205	in house	29.99		----	13.89		----	
1241		30.003		----	13.883		----	
1242		30.027		----	13.918		----	
1247	INH-4500	30.06		----	13.92		----	
1253		29.98		----	13.84		----	
1605		30.06		----	----		----	
1726		----		----	----		----	
1727		31.41	G(0.01)	----	15.5	G(0.01)	----	
4247	INH-4500	29.99		----	13.96		----	
	normality	OK			OK			
	n	9			8			
	outliers	1			1			
	mean (n)	29.993			13.893			
	st.dev. (n)	0.0537			0.03924			
	R(calc.)	0.150			0.110			
	R(lit)	unknown			unknown			
	R(iis10C12c)	0.309			0.148			



Determination of Glucose, Fructose, Sucrose Maltose on sample #11123 and #11124; results in %M/M

sample:#11123

lab	method	Glucose	Fructose	Sucrose	Maltose	Remarks
310		1.95	1.55	16.73	<0.05	
551		----	----	----	----	
862		----	----	----	----	
1126		----	----	----	----	
1205		----	----	----	----	
1241		----	----	----	----	
1242		----	----	----	----	
1247		----	----	----	----	
1253		----	----	----	----	
1605		----	----	----	----	
1726		----	----	----	----	
1727		12.83	6.83	----	1.20	
4247		----	----	----	----	
	normality	n.a	n.a	n.a	n.a	
	n	2	2	1	1	
	outliers	n.a	n.a	n.a	n.a	
	mean (n)	n.a	n.a	n.a	n.a	
	st.dev. (n)	n.a	n.a	n.a	n.a	
	R(calc.)	n.a	n.a	n.a	n.a	
	R(lit)	unknown	unknown	unknown	unknown	

sample:#11124

lab	method	Glucose	Fructose	Sucrose	Maltose	Remarks
310		0.21	0.06	22.59	0.09	
551		----	----	----	----	
862		----	----	----	----	
1126		----	----	----	----	
1205		----	----	----	----	
1241		----	----	----	----	
1242		----	----	----	----	
1247		----	----	----	----	
1253		----	----	----	----	
1605		----	----	----	----	
1726		----	----	----	----	
1727		11.73	6.30	----	1.25	
4247		----	----	----	----	
	normality	n.a	n.a	n.a	n.a	
	n	2	2	1	2	
	outliers	n.a	n.a	n.a	n.a	
	mean (n)	n.a	n.a	n.a	n.a	
	st.dev. (n)	n.a	n.a	n.a	n.a	
	R(calc.)	n.a	n.a	n.a	n.a	
	R(lit)	unknown	unknown	unknown	unknown	

APPENDIX 2**Analytical details regarding Strength determination**

Lab	Distillation	Equipment	#11123		#11124		Other details
			used sample in ml	how much distillate was obtained in ml	used sample in ml	how much distillate was obtained in ml	
310	yes	DE40	200	±60	200	±30	
1126	yes	Dichtheidsmeting*	200	200	200	200	* is used for evaluation
	yes	GC	10	50	10	50	
1205	yes	Density meter	100 and 50	100 and 100	100	100	
1241	yes	-----	±100	±200	±150	±200	
1242	yes	-----	2	100	2	50	
1247	yes	DMA	50*	100	50*	100	* = grams
1253		Mettler DE45 Buchi K350 Electrode (Metrohm)					
1605	yes	Density meter	*200	400	*200	400	*diluted to 500 ml with H ₂ O
1727	no	GC					
4247	yes	DMA	50*	100	50*	100	* = grams

APPENDIX 3

List of number of participants per country

2 labs in BELGIUM

1 lab in BRAZIL

1 lab in P.R. of CHINA

2 labs in SPAIN

7 labs in THE NETHERLANDS

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
E	= error in calculations
ex	= excluded from calculations
n.a.	= not applicable
U	= unit error
SDS	= safety data sheet

Literature:

- 1 i.i.s. Interlaboratory Studies, Protocol for the Organisation, Statistics & Evaluation, January 2010
- 2 W. Horwitz and R. Albert, J. AOAC Int., Vol. 79, 3, p. 589, (1996)
- 3 ASTM E178-02
- 4 ASTM E1301-03
- 5 ISO13528-05
- 5 ISO 5725-86
- 6 ISO 5725, parts 1-6, 1994
- 7 M. Thompson and R. Wood, J. AOAC Int, 76, 926, (1993)
- 8 W.J. Youden and E.H. Steiner, Statistical Manual of the AOAC, (1975)
- 9 IP 367/84
- 10 DIN 38402 T41/42
- 11 P.L. Davies, Fr. Z. Anal. Chem, 331, 513, (1988)
- 12 J.N. Miller, Analyst, 118, 455, (1993)
- 13 Analytical Methods Committee Technical brief, No 4. January 2001
- 14 The Royal Society of Chemistry 2002, Analyst, 2002, 127, page 1359-1364, P.J. Lowthian and M. Thompson. (see <http://www.rsc.org/suppdata/an/b2/b205600n/>)