

## Statistisches Protokoll / Statistical Protocol

Das Statistische Protokoll beschreibt sämtliche statistische Berechnungen, welche zur Auswertung von Datensätzen bzw. zur Bewertung der analytischen Leistung der Laboratorien bei der muva kempten GmbH verwendet werden.

*The statistical protocol describes all statistical calculations that are used for the evaluation of datasets and for the assessment of the laboratory's performance.*

### INHALTSVERZEICHNIS / CONTENTS

	Seite Page	
1. Begriffe und Definitionen	1	1. <i>Terms and Definitions</i>
2. Angewandte Statistik	6	2. <i>Used Statistics</i>
3. Auswertung der sensorischen Eignungsprüfungen	11	3. <i>Evaluation of the proficiency testing schemes in sensory</i>
4. Homogenität des Prüfmaterials	14	4. <i>Homogeneity of the testing material</i>
5. Bewertung der analytischen Leistung	16	5. <i>Assessment of the analytical performance</i>
6. Verzeichnis der verwendeten Standards / Normen und Literatur	19	6. <i>Register of used standards / norms and literature</i>

### 1. BEGRIFFE UND DEFINITIONEN / TERMS AND DEFINITIONS

Begriffe	Abk. abbr.	Formel / formula	Terms
<b>Laboreinzelwert:</b> Einzelwert einer Doppelbestimmung	$X_{1lab}, X_{2lab}$		<b>Laboratory single value:</b> <i>Single value of a double determination</i>
<b>Labormittelwert [5, 6]:</b> Mittelwert einer Doppelbestimmung	$X_{lab}$	$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_{lab_i}$	<b>Laboratory mean value [5, 6]:</b> <i>Mean of a double determination</i>
<b>Anzahl der Werte eines Datensatzes</b>	$n$		<b>Number of values of a dataset</b>
<b>Mittelwert eines Datensatzes [5, 6]:</b> Mittelwert sämtlicher Labormittelwerte (vor bzw. nach Eliminierung der Ausreißer)	$\bar{x}$	$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i$	<b>Mean of a dataset [5, 6]:</b> <i>Mean of all laboratory mean values (before or after elimination of outliers)</i>

Begriffe	Abk. abbr.	Formel / formula	Terms
<p><b>Spannweite</b> [3, 11]:</p> <p>Differenz des größten und des kleinsten Wertes eines Datensatzes</p>	w	$W = X_{\max} - X_{\min}$	<p><b>Range</b> [2, 11]:</p> <p><i>Difference of the major and the minor value of the dataset</i></p>
<p><b>Median:</b></p> <p>Ist der Zentralwert, d.h.:</p> <p>falls n ungerade ist</p> <p>falls n gerade ist</p>	$M_{pt}$	$M = med\{y_i\}$ $M = y_{\left(\frac{n+1}{2}\right)}$ $M = \frac{1}{2} \left( y_{\frac{n}{2}} + y_{\left(\frac{n}{2}+1\right)} \right)$	<p><b>Median:</b></p> <p><i>Is the central value, that is:</i></p> <p><i>if n is odd</i></p> <p><i>if n is even</i></p>
<p><b>Zugewiesener Wert</b> [2, 5, 6]:</p> <p>Mittelwert der Teilnehmer ohne Ausreißer, die eine zur Analyse geeignete Methode eingesetzt haben. Zur Festlegung des zugewiesenen Wertes wird die unten beschriebene Statistik verwendet.</p> <p>Beim <u>Referenzmaterial</u> entspricht der zugewiesene Wert dem Referenzwert.</p>	$X_{pt}$		<p><b>Assigned value</b> [2, 5, 6]:</p> <p><i>Mean value of all participants, using an adequate method for analysis. The statistical methods described below are used for determination of the assigned value</i></p> <p><i>For <u>reference material</u> the assigned value is the reference value.</i></p>
<p><b>Konsenswert nach robuster Auswertung (Hampelschätzer)</b> [2, 3, 5, 6]</p> <p>Der Konsenswert dient zur weiteren Absicherung des zugewiesenen Wertes.</p> <p>Berechnung s. Seite 8.</p>	$X_{pt,robust}$		<p><b>Consensus value by robust statistics (Hampelschätzer)</b> [2, 3, 5, 6]</p> <p><i>The consensus value is used for hedging the assigned value.</i></p> <p><i>Calculation see page 8.</i></p>
<p><b>Zugewiesener Wert in relativer Konzentration</b> [2, 5, 6, 12]</p> <p>z.B.: 1 mg/kg = 0,000001 kg/kg</p>	$X_{pt,rel}$		<p><b>Assigned value in relative concentration</b> [2, 5, 6, 12]</p> <p>e.g: 1 mg/kg = 0,000001 kg/kg</p>
<p><b>Standardabweichung</b> [5, 6]</p> <p>Drückt die Streuung der Datensätze um den Mittelwert aus.</p>	s	$s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$	<p><b>Standard deviation</b> [5, 6]</p> <p><i>Expresses the deviation around the mean</i></p>

Begriffe	Abk. abbr.	Formel / formula	Terms
<p><b>Standardabweichung des zugewiesenen Wertes</b></p> <p>Drückt die Streuung der Werte um den zugewiesenen Wert aus.</p>	S <sub>pt</sub>		<p><b>Standard deviation of the assigned value</b></p> <p>Expresses the deviation of the values around the assigned value.</p>
<p><b>Einfache Unsicherheit des zugewiesenen Wertes [2, 3]</b></p> <p>Wird als Standardunsicherheit des zugewiesenen Wertes ausgedrückt</p>	U <sub>pt</sub>	$u_{pt} = \frac{s_{pt}}{\sqrt{n}}$	<p><b>Simple Uncertainty of the assigned value [2, 3]</b></p> <p>Is expressed as standard uncertainty of the assigned value.</p>
<p><b>Erweiterte Unsicherheit des zugewiesenen Wertes</b></p> <p>Erweiterung durch Multiplikation der einfachen Unsicherheit mit dem Erweiterungsfaktor k = 2 (95% statistische Wahrscheinlichkeit)</p>	U <sub>pt</sub>	$U_{pt} = \frac{s_{pt}}{\sqrt{n}} \cdot 2$	<p><b>Expanded Uncertainty of the assigned value</b></p> <p>Calculation by multiplication of the simple uncertainty by the use of the coverage factor k = 2 (95% statistical probability)</p>
<p><b>Einfache Unsicherheit der Ergebnisse aus dem Homogenitätstest [3, 19]</b></p> <p>Die Standardabweichung zwischen den Packungen drückt die Heterogenität des Materials aus</p>	U <sub>hom</sub>	$u_{hom} = s_{bb}$	<p><b>Simple Uncertainty of the homogeneity results [3, 19]</b></p> <p>The standard deviation between the packages expresses the heterogeneity of the material</p>
<p><b>Vergleichsstandardabweichung [2, 11]</b></p> <p>In Ringversuchen zur Validierung von Methoden unter Vergleichsbedingungen ermittelte Standardabweichung.</p> <p>Sie ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben.</p> <p>Sie dient als Standardabweichung zur Eignungsbewertung bei der Berechnung der z-Scores (= σ<sub>pt, z</sub>).</p>	s(R)		<p><b>Reproducibility standard deviation [2, 11]</b></p> <p>It is determined under reproducible conditions in collaborative studies for validation of methods. It is mentioned as precision data in national and international standardized methods.</p> <p>It is used as standard deviation for the performance assessment for the calculation of the z-scores (=σ<sub>pt, z</sub>).</p>
<p><b>Wiederholstandardabweichung:</b></p> <p>In Ringversuchen zur Validierung von Methoden unter Wiederholbedingungen ermittelte Standardabweichung.</p> <p>Sie ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben.</p>	s(r)		<p><b>Repeatability standard deviation:</b></p> <p>It is determined under repeatable conditions in collaborative studies for validation of methods. It is mentioned as precision data in national and international standardized methods.</p>

Begriffe	Abk. abbr.	Formel / formula	Terms
<p><b>Relative Standardabweichung nach Horwitz [12]:</b></p> <p>Berechnet in Abhängigkeit von der relativen Konzentration des Analyten (zugewiesener Wert) <math>x_{ref,rel}</math>.</p>	$s(H)_{rel}$	$s(H)_{rel} = 2^{(1-0,5 \cdot \log(x_{ref,rel}))}$	<p><b>Relative standard deviation acc. to Horwitz [12]:</b></p> <p>Calculated on the basis of the relative concentration of the analyte (assigned value) <math>x_{ref,rel}</math>.</p>
<p><b>Relative Standardabweichung nach Horwitz-Thompson [3, 21]:</b></p> <p>Berechnet in Abhängigkeit von der relativen Konzentration des Analyten (zugewiesener Wert) <math>x_{ref,rel}</math>.</p> <p>A: <math>x_{ref,rel} &lt; 12 \cdot 10^{-7}</math>                      B: <math>12 \cdot 10^{-7} \leq x_{ref,rel} \leq 0,138</math>                      C: <math>x_{ref,rel} &gt; 0,138</math></p>	$s(HT)_{rel}$	$s_A(HT)_{rel} = 0,22x_{ref,rel}$ $s_B(HT)_{rel} = 0,22x_{ref,rel}^{0,8495}$ $s_C(HT)_{rel} = 0,01x_{ref,rel}^{0,5}$	<p><b>Relative standard deviation acc. to Horwitz-Thompson [12]:</b></p> <p>Calculated on the basis of the relative concentration of the analyte (assigned value) <math>x_{ref,rel}</math>.</p> <p>A: <math>x_{ref,rel} &lt; 12 \cdot 10^{-7}</math>                      B: <math>12 \cdot 10^{-7} \leq x_{ref,rel} \leq 0,138</math>                      C: <math>x_{ref,rel} &gt; 0,138</math></p>
<p><b>Standardabweichung nach Horwitz bzw. Horwitz-Thompson [3, 12, 21]:</b></p> <p>Standardabweichung berechnet aus der relativen Standardabweichung und dem zugewiesenen Wert in der üblichen Konzentration.</p> <p>Sie dient in einigen Eignungsprüfungen als Standardabweichung zur Berechnung der z-Scores.</p>	$s(H)$	$s(H) = \frac{s(H)_{rel}}{100} \cdot x_{ref}$	<p><b>Standard deviation acc. to Horwitz resp. Horwitz-Thompson [3, 12, 21]:</b></p> <p>Standard deviation calculated on the basis of the relative standard deviation and the assigned value (in original concentration).</p> <p>It is used as standard deviation for the performance assessment for the calculation of the z-scores in some PT schemes.</p>
<p><b>Standardabweichung innerhalb der Proben (= analytische Unsicherheit aus dem Homogenitätstest) [3, 11]:</b></p> <p>Drückt die analytische Streuung der Datensätze um den zugehörigen Mittelwert innerhalb der Proben im Homogenitätstest aus.</p>	$S_{wb}$		<p><b>Standard deviation within the bottles (= analytical uncertainty deriving from the homogeneity test) [3, 11]:</b></p> <p>Expresses the analytical deviation around the mean within the samples of the homogeneity test.</p>
<p><b>Standardabweichung zwischen den Proben (= Materialunsicherheit aus dem Homogenitätstest) [3, 11]:</b></p> <p>Drückt die Streuung der Datensätze einer repräsentativen Anzahl von Stichproben um den zugehörigen Mittelwert aus dem Homogenitätstest aus.</p>	$S_{bb}$		<p><b>Standard deviation between the bottles (=uncertainty of the material deriving from the homogeneity test) [3, 11]:</b></p> <p>Expresses the deviation around the mean of a representative number of samples of the homogeneity test.</p>

Begriffe	Abk. abbr.	Formel / formula	Terms
<b>Standardabweichung der Mittelwerte [3]:</b> Berechnet aus den zwei Teilstichproben des Homogenitätstest nach DIN ISO 13528	S <sub>xt</sub>		<b>Standard deviation of the means [3]:</b> Calculated of the two sub samples in the homogeneity test of DIN ISO 13528
<b>Homogenitätskriterium [3, 11]:</b> Dient zur Bewertung der Homogenität der Prüfmateriale	S <sub>erl</sub>		<b>Criterion for homogeneity [3, 11]:</b> It is used for the assessment of the homogeneity of the testing materials
<b>Unsicherheit Stabilität (Referenzmaterial) [19]</b> Standardabweichung aus dem Stabilitätstest für ausgewählte Parameter	u <sub>stab</sub>	$u_{stab} = s_{stab}$	<b>Uncertainty stability (reference material) [19]</b> Standard uncertainty, calculation is based on standard deviation of the stability test for selected parameters
<b>Robuste Standardabweichung [3]</b> Ermittelt aus der Q-Methode	S <sub>R</sub>		<b>Robust standard deviation [3]</b> Q-method
<b>Unsicherheit eines Datensatzes (Sensorik)</b> Unsicherheitsintervall (95% Wahrscheinlichkeit) um den Mittelwert eines Datensatzes	MU		<b>Uncertainty of a dataset (sensorry)</b> Uncertainty interval (95 % probability) around the mean value of a dataset
<b>Erweiterte Unsicherheit des Referenzwertes [17] (Referenzmaterial)</b> Berechnet sich als kombinierte, erweiterte Unsicherheit mit den Größen u <sub>ref</sub> (= einfache Unsicherheit, d.h. Standardunsicherheit des Referenzwertes), u <sub>hom</sub> und u <sub>stab</sub> . Erweiterungsfaktor k = 2 (95% statistische Wahrscheinlichkeit)	U	$U = \sqrt{u_{ref}^2 + u_{hom}^2 + u_{stab}^2} \cdot 2$	<b>Extended uncertainty of the reference value [17] (reference material)</b> Calculated as combined, expanded uncertainty based u <sub>ref</sub> (= simple uncertainty, i.e. standard uncertainty of the reference value), u <sub>hom</sub> and u <sub>stab</sub> . Coverage factor k=2(95% statistical probability)
<b>Kritische Differenz:</b> Die kritische Differenz ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben.	CD		<b>Critical difference</b> The critical difference is mentioned as precision data in national and international standardized methods.

Begriffe	Abk. abbr.	Formel / formula	Terms
<b>Studentfaktor [5, 6]:</b> Erweiterungsfaktor zur Bestimmung der erweiterten Unsicherheit (95% Wahrscheinlichkeit) für kleine Datensätze. Er ist von der Anzahl der Datensätze abhängig.	t		<b>Student factor [5, 6]:</b> Coverage factor for calculating the extended uncertainty (95% probability) for small datasets. It depends on the number of datasets.
<b>Streubereich [5, 6]:</b> Unsicherheitsintervall (95% Wahrscheinlichkeit) um einen Mittelwert	T	$T = s \cdot t$	<b>Range of variation [5, 6]:</b> Uncertainty interval (95% probability) around the mean value
<b>Vertrauensbereich [5, 6]:</b> Unsicherheitsintervall (95% Wahrscheinlichkeit) um den wahren Wert (Maß für die Zuverlässigkeit des Mittelwertes)	$\Delta \bar{x}$	$\Delta \bar{x} = \frac{s \cdot t}{\sqrt{n}}$	<b>Confidential interval [5, 6]:</b> Uncertainty interval (95% probability) around the true value (dimension for the „reliability“ of the mean)
<b>Prüfgröße David-Test zur Normalverteilung [10]:</b> Gegebener tabellarischer Vergleichswert	PG		<b>Test statistic for the check for normal distribution (David-Test) [10]:</b> Given tabulated standard value

## 2. ANGEWANDTE STATISTIK / USED STATISTICS

<b>Prüfung auf Normalverteilung – Vortest nach David [10]:</b> Voraussetzung für die weiter angewendete Statistik in Eignungsprüfungen (Mittelwert, Standardabweichung, Streubereich, Vertrauensbereich) ist die Prüfung der Verteilung der Daten auf Normalverteilung. Sie erfolgt in mehreren Schritten: <ol style="list-style-type: none"> <li>Berechnung der Spannweite w des Datensatzes  <math>w = (x_{max} - x_{min})</math></li> </ol>	<b>Check for normal distribution – Test David [10]:</b> Condition for all further statistical calculations of the proficiency testing (mean, standard deviation, dispersion range) is the check of the data for a normal distribution. It is performed by the following steps: <ol style="list-style-type: none"> <li>Calculation of the range w of the data  <math>w = (x_{max} - x_{min})</math></li> </ol>
<ol style="list-style-type: none"> <li>Berechnung der Prüfgröße  <math>PG = w / s</math></li> <li> <math>x_{max}</math>: größter Wert des Datensatzes  <math>x_{min}</math>: kleinster Wert des Datensatzes                      s: Standardabweichung des Datensatzes                      Suchen der Werte für die kritischen Grenzen in der Tabelle von Pearson und Stephens in dem Signifikanzniveau von 95 %                 </li> <li>Liegt die Prüfgröße PG innerhalb dieser kritischen Grenzen ist die Verteilung normal.</li> </ol>	<ol style="list-style-type: none"> <li>Calculation of the test statistic  <math>PG = w / s</math></li> <li> <math>x_{max}</math>: major value of the dataset  <math>x_{min}</math>: minor value of the dataset                      s: standard deviation of the dataset                      Looking for the values of the critical limits in the table of Pearson and Stephens with the significant level of 95 %                 </li> <li>Is the test statistic PG between the two critical limits, the distribution is normal</li> </ol>

<p><b>Ausreißer Test (Grubbs) [3, 5, 9]:</b> Für die Ermittlung von Ausreißern innerhalb eines Datensatzes, der sich aus Ergebnissen von mehreren Laboratorien zusammensetzt, wird der Ausreißer Test nach Grubbs herangezogen [6]. Die Ausreißer werden wie folgt bestimmt.</p>	<p><b>Outlier test (Grubbs) [3, 5, 9]:</b> <i>For detection of outliers in a dataset, obtaining results of different laboratories, the outlier test according to Grubbs is used [6]. The outliers are determined as follows.</i></p>
<p>1. Berechnung der Puffergröße:</p>	<p>1. Calculation of the difference of the buffer size:</p>
$Puffergröße / buffersize = \frac{x_{lab} - \bar{x}}{s}$	<p><math>x_{lab}</math> = Labormittelwert / laboratory mean  <math>\bar{x}</math> = Gesamtmittelwert / total mean  <math>s</math> = Standardabweichung / standard deviation</p>
<p>2. Vergleich der Puffergröße mit den Grubbs-Werten (Abhängig von der Anzahl der Teilnehmer) <u>Ausreißer liegt vor</u>, wenn die Puffergröße größer als der Grubbs-Wert ist. <u>Kein Ausreißer</u> liegt vor, wenn die Puffergröße kleiner als der Grubbs-Wert ist.</p>	<p>2. Comparing the buffer size with the Grubbs-value <u>An outlier is detected</u>, if the buffer size is bigger than the Grubbs-value. <u>No outlier is detected</u>, if the buffer size is smaller than the Grubbs-value.</p>



**Zugewiesener Wert [2, 5, 6]:**

Der zugewiesene Wert wird wie folgt festgelegt:

1. Für chemische Eignungsprüfungen dienen als Ausgangsdatensätze die Labormittelwerte, die mit Hilfe einer standardisierten Methode ermittelt wurden.  
  
Wurde für einen Parameter eine definierende Methode (Codex Typ 1) international festgelegt, so werden zur Festlegung des zugewiesenen Wertes nur Werte berücksichtigt, die nach dieser Methode ermittelt wurden.  
  
Bei eindeutig chemisch definierten Parametern werden alle unabhängigen Methoden insoweit berücksichtigt, als die Vertrauensbereiche der Mittelwerte nach den einzelnen Methoden überlappen.  
  
Für mikrobiologische Eignungsprüfungen werden die Labormittelwerte zur Basis 10 logarithmiert. Alle weiteren Berechnungen von statistischen Daten und Bewertungskriterien erfolgen auf Basis dieser logarithmierten Werte.
2. Prüfung des Datensatzes auf Normalverteilung (David-Test) und Vergleich der Ergebnisse mit denen der robusten Statistik (Mittelwert, Standardabweichung, Chi<sup>2</sup>-Wert).
3. Bei normaler Verteilung, entweder Berechnung des Mittelwertes des ausreißerbereinigten Datensatzes (Grubbs-Test) bzw. des Konsenswertes (robust). Plausibilitätstest durch Vergleich der Ergebnisse (sensible Statistik, Median, Hampel-Schätzer)
4. Bei nicht normaler Verteilung wird der Median oder der robuste Konsenswert herangezogen.

**Assigned value [2, 5, 6]:**

The assigned value is determined as follows:

1. For the chemical proficiency testing schemes, the initial dataset is based on the mean values of the laboratories, which used standardized methods for analysis.  
  
*For parameters where a defining method (Codex type 1) has been internationally established, only results obtained using this method were considered.*  
  
*For chemically defined parameters all results obtained using independent methods were considered as far as confidence intervals of means for a particular method overlap with others.*  
  
*For microbiological proficiency testing schemes, the initial dataset of the mean values of the laboratories are calculated to common logarithms (base 10). Further calculation of statistical values and performance criteria are based on the logarithm of these values.*
2. Check of the normal distribution of the dataset (test David) and comparison of the results with the results of the robust statistics (mean value, standard deviation, Chi<sup>2</sup>-value).
3. If the distribution is normal, calculation of the mean value of the outlier-cleaned dataset (Grubbs-test) or robust consensus value. Plausibility check via comparison of the results (sensitive statistics, Median, Hampel-Schätzer)
4. If the distribution is not normal, the median or the consensus value is determined as assigned value.

**Unsicherheit des Zugewiesener Wertes [2, 3, 5, 6]:**

Als Unsicherheit des zugewiesenen Wertes dient die erweiterte Standardunsicherheit (U<sub>PT</sub>) (95% statistische Wahrscheinlichkeit). Die einfache Standardunsicherheit (u<sub>PT</sub>) wird hierbei mit dem Erweiterungsfaktor k = 2 multipliziert und entspricht somit weitgehend dem Vertrauensbereich des zugewiesenen Wertes:

**Uncertainty of the assigned value [2, 3, 5, 6]:**

The expanded standard uncertainty (U<sub>PT</sub>) is determined as uncertainty of the assigned value (95% statistical probability). Hereby the simple standard uncertainty (u<sub>PT</sub>) is multiplied with the expansion factor k = 2 and is therefore largely related to the confidence interval of the assigned value:

$$u_{pt} = \frac{s_{pt}}{\sqrt{n_{pt}}} \quad \rightarrow \quad U_{pt} = \frac{s_{pt}}{\sqrt{n_{pt}}} \cdot 2$$



<p><b>Robuster Mittelwert [3]:</b> Hampelschätzer:</p>	<p><b>Robust mean [3]:</b> Hampel estimator:</p>
$\sum_{j=1}^J \Psi\left(\frac{y_j - m}{s_R}\right) = 0$	$\Psi(x) = \begin{cases} 0 & x \leq -4,5 \\ -4,5 - x & -4,5 < x \leq -3 \\ -1,5 & -3 < x \leq -1,5 \\ x & -1,5 < x \leq 1,5 \\ 1,5 & 1,5 < x \leq 3 \\ 4,5 - x & 3 < x \leq 4,5 \\ 0 & x > 4,5 \end{cases}$

<p><b>Robuste Standardabweichung [3]:</b> Q-Methode:</p>	<p><b>Robust standard deviation [3]:</b> Q-method:</p>
$s_R = \frac{G_1^{-1}(q)}{\sqrt{2}\Phi^{-1}(0,5 + 0,5q)}$	<p>q = Quantilparameter                  Φ = Verteilungsfunktion der Standardnormalverteilung                  G<sub>1</sub>(x<sub>i</sub>) = Funktion zur Stetigkeitskorrektur</p>

<p><b>Abschätzung der erweiterten Messunsicherheit der Laborergebnisse [18, 20]:</b></p> <p>Ab drei Laborergebnissen, bezogen auf ein Verfahren, werden die Beiträge zur Unsicherheit aus der Standardabweichung der einzelnen Ergebnisse (= Reproduzierbarkeit) sowie aus der Abweichung der einzelnen Ergebnisse zum zugewiesenen Wert (= BIAS) in Anlehnung an den NORDTEST berechnet.</p> <p>Die Abweichung der Einzelergebnisse zum zugewiesenen Wert wird berechnet (= BIAS), indem vom Laborergebnis der zugewiesene Wert subtrahiert wird:</p>	<p><b>Estimation of the expanded measurement uncertainty of the laboratories results [18, 20]:</b></p> <p>Starting at 3 results provided from a laboratory by the use of <u>one</u> method, the following contributions of the uncertainty are determined from the standard deviation calculated on basis of the single results (= reproducibility) and the bias of the single results in relation to the assigned value by following the NORDTEST.</p> <p>To calculate the bias of the single results in relation to the assigned value, the assigned value will be subtracted from the laboratory result:</p>
--	---

$$bias_{ab1..labi} = x_{lab1..labi} - x_{pt}$$

$$EP_{bias} = \sqrt{\frac{(bias_{ab1})^2 + (bias_{ab2})^2 + \dots + (bias_{labi})^2}{i}}$$

<p>x<sub>lab</sub> = Laborergebnis (evtl. Mittelwert)                  x<sub>pt</sub> = Zugewiesener Wert                  bias<sub>lab</sub> = Abweichung der Einzelergebnisse zum zugewiesenen Wert</p>	<p>x<sub>lab</sub> = Laboratory result (possibly mean)                  x<sub>pt</sub> = Assigned value                  bias<sub>lab</sub> = Deviation of the single laboratory result to the assigned value</p>
---	---

<p><math>EP_{bias}</math> = Abweichung des Laboregebnisses zum zugewiesenen Wert</p>	<p><math>EP_{bias}</math> = Deviation of the laboratory result to the assigned value</p>
--	--

<p>Bei der Berechnung der einfache Unsicherheit <math>u_{bias}</math> (= Beitrag der Richtigkeit) wird auch die Standardunsicherheit <math>u_{pt}</math> des zugewiesenen Wertes berücksichtigt. <math>u_{bias}</math> berechnet sich demnach als kombinierte Unsicherheit aus <math>EP_{bias}</math> und der Unsicherheit des zugewiesenen Wertes <math>u_{pt}</math>:</p>	<p><i>The standard uncertainty of the assigned value <math>u_{pt}</math> is also considered by the calculation of the simple uncertainty <math>u_{bias}</math> (= contribution of accuracy). <math>u_{bias}</math> will therefore be calculated as a combined uncertainty of <math>EP_{bias}</math> and the uncertainty of the assigned value <math>u_{pt}</math>:</i></p>
---	--

$$u_{bias} = \sqrt{u_{pt}^2 + EP_{bias}^2}$$

<p><math>u_{pt}</math> = Standardunsicherheit des zugewiesenen Wertes</p>	<p><math>u_{pt}</math> = standard uncertainty of the assigned value</p>
<p><math>EP_{bias}</math> = Abweichung des Laboregebnisses zum zugewiesenen Wert</p>	<p><math>EP_{bias}</math> = Deviation of the laboratory result to the assigned value</p>

<p>Anschließend wird eine kombinierte Unsicherheit <math>u_{lab}</math> aus der Standardabweichung der Labor-Einzelergebnisse <math>s_{xlab}</math> (= Reproduzierbarkeit) und der einfachen Unsicherheit <math>u_{bias}</math>:</p>	<p><i>In the following the combined uncertainty <math>u_{lab}</math> is calculated by the standard deviation of the single laboratories results <math>s_{xlab}</math> (= reproducibility) and the simple uncertainty <math>u_{bias}</math>:</i></p>
--	---

$$u_{lab} = \sqrt{u_{bias}^2 + s_{xlab}^2}$$

<p><math>u_{bias}</math> = Einfache Unsicherheit (Beitrag der Richtigkeit)</p>	<p><math>u_{bias}</math> = simple uncertainty (contribution of accuracy)</p>
<p><math>s_{xlab}</math> = Standardabweichung der Labor-Einzelergebnisse (Beitrag der Reproduzierbarkeit)</p>	<p><math>s_{xlab}</math> = standard deviation of the single results of the laboratory (contribution of reproducibility)</p>

<p>Die erweiterte kombinierte Unsicherheit <math>U_{lab}</math> des Labors wird durch Multiplikation der einfachen Unsicherheit <math>u_{lab}</math> mit dem Erweiterungsfaktor <math>k=2</math> (für 95% Wahrscheinlichkeit) berechnet:</p>	<p><i>The expanded combined uncertainty <math>U_{lab}</math> of the laboratory is calculated by multiplication of the simple uncertainty of the laboratory <math>u_{lab}</math> with the coverage factor <math>k=2</math> (for 95% probability):</i></p>
--	--

$$U_{lab} = 2 \cdot u_{lab}$$

### 3. AUSWERTUNG DER SENSORISCHEN EIGNUNGSPRÜFUNGEN MIT LEISTUNGSBEWERTUNG / EVALUATION OF THE PROFICIENCY TESTING SCHEMES IN SENSORY AND ASSESSMENT OF THE PERFORMANCE

<p><b>Page Test [8]:</b></p> <p>Der Test basiert auf Rangsummen, die in die natürliche Reihenfolge gebracht werden können.</p> $L = R_1 + 2R_2 + 3R_3 + \dots + kR_k$ <p>Ist der berechnete L-Wert gleich oder größer dem tabellierten kritischen Wert, daraus folgt, dass die Rangreihenfolge, wie sie von den Prüfern wahrgenommen wird, mit der vorgegebenen Reihenfolge übereinstimmt.</p>	<p><b>Page test [8]:</b></p> <p><i>The test is based on the rank data, which should be arranged in the correct order.</i></p> $L = R_1 + 2R_2 + 3R_3 + \dots + kR_k$ <p><i>If the calculated L-value is equal or higher as the critical value, is the conclusion, that the rank data, identified of the panellist, agrees with the given order.</i></p>
<p><b>Fehlersumme:</b></p> <p>Die Fehlersumme wird aus der Reihenfolge und den Abständen der 4 Prüfproben berechnet. Die Abstände der 4 Prüfproben werden aus dem Homogenitätstest entnommen.</p> <p>Die ermittelte Rangreihenfolge wird anschließend mit der tatsächlichen verglichen und für jede Fehlplatzierung die Fehlerpunkte aus den Abständen ermittelt. Die Summe der Fehlerpunkte ergibt die Fehlersumme.</p>	<p><b>Error sum:</b></p> <p><i>The error sum is calculated out of the sequence and the intervals of the four ranking test samples. The intervals are taken from the homogeneity test.</i></p> <p><i>The ranking test of each panellist is compared with the real sequence and the error points are calculated out of the intervals (table) and the sum makes the error sum.</i></p>
<p><b>Punktesystem zur Leistungsbewertung für die Erkennung des Attributes und Bildung der richtigen Rangfolge [8]:</b></p> <p>In Anlehnung an den Page Test und mit Berücksichtigung der Erkennung der Grundgeschmacksart werden Punkte zwischen 0 und 10 je nach Ergebnis vergeben. Die Höchstpunktzahl 10 wird bei korrekter Rangfolge und richtig erkannter Grundgeschmacksart bzw. Aroma vergeben.</p>	<p><b>point system for performance evaluation for the detection of the attribute and for the correct building of the rank order [8]:</b></p> <p><i>According to the page test and regarding the identification of the correct basic taste points between 0 and 10 are assigned. The top of the score 10 is given for the correct rank order and the correct identification of the basic taste or flavour.</i></p>
<p><b>Dreieckstest Signifikanzermittlung [14]:</b></p> <p>Die Signifikanz wird für jeden Prüfer und für jedes Panel nach der Signifikanztabelle in DIN EN ISO 4120 ermittelt. Zur Auswertung werden die Signifikanzniveaus <math>\alpha = 0,1</math> (bedingt signifikant), <math>\alpha = 0,05</math> (signifikant), <math>\alpha = 0,01</math> (hoch signifikant) und <math>\alpha = 0,001</math> (höchst signifikant) herangezogen. Panellisten und Panel, die die Mindestanzahl korrekter Antworten nicht erfüllen erhalten das Ergebnis „nicht signifikant“.</p>	<p><b>Triangle test calculation of the significance level [14]:</b></p> <p><i>For each panel and each panellist the significance level is calculated by the table of DIN EN ISO 4120. The significance levels <math>\alpha = 0,1</math> (limited significant), <math>\alpha = 0,05</math> (significant), <math>\alpha = 0,01</math> (high significant) and <math>\alpha = 0,001</math> (highest significant) are used for the evaluation.</i></p> <p><i>Panels and panellists who did not fulfil the minimum number of correct responses get the result "not significant".</i></p>

**ISO 22935 [16]:**

Die Beurteilung von Lebensmitteln nach DIN ISO 22935 ist eine deskriptive sensorische Analyse anhand einer 5 Punkte Skala. Der optimale Qualitätsstandard entspricht der Höchstnote von 5,0 Punkten. Die Prüfung und Bewertung der Lebensmittel erfolgt nach produkt-spezifischen Prüfschemen in gegebenen Attributen.

Bewertungstabelle nach DIN ISO 22935:

Punkte	Beschreibung in Worten
5	Keine Abweichung von der sensorischen Spezifikation
4	Geringfügige Abweichung von der sensorischen Spezifikation
3	Erkennbare Abweichung von der sensorischen Spezifikation
2	Deutliche Abweichung von der sensorischen Spezifikation
1	Sehr deutliche Abweichung von der sensorischen Spezifikation

**ISO 22935 [16]:**

*The evaluation of food according to DIN ISO 22935 is a descriptive sensory analyzes according to a 5 point scale. The optimum of quality standard is the highest rate of 5.0 points. The testing and evaluation of the foods is based on product specific testing schemes in given attributes*

*Evaluation table of DIN ISO 22935:*

Points	Description in words
5	No deviation from the sensory specification
4	Minor deviation from the sensory specification
3	Recognizable Deviation from the sensory specification
2	Clear deviation from the sensory specification
1	Very clear deviation from the sensory specification

**Bewertungskriterium z-Score in der ISO 22935 Eignungsprüfung**

Der z-Score stellt eine Bewertung der Leistungsfähigkeit eines Laboratoriums dar. Er errechnet sich wie folgt:

$$z - Score = \frac{x_{lab} - x_{pt}}{\sigma_{pt}}$$

- $x_{lab}$  = Labormittelwert
- $x_{pt}$  = zugewiesener Wert
- $\sigma_{pt}$  = Standardabweichung zur Eignungsbewertung ( $\sigma_{pt} = 1$ )

Für die Standardabweichung zur Eignungsbewertung  $\sigma_{pt}$  wurde der Wert 1 festgelegt, da dadurch die Grenzen der Leistungsbewertung mit der Range  $\geq 2$  sehr gut verdeutlicht werden.

Für absolute z-Scores (d. h. ohne Vorzeichenberücksichtigung) gelten folgende Grenzen

- $|z| \leq 1$ ; gute (zufriedenstellende) Leistung
- $1 < |z| < 2$ ; fragliche Leistung (Warnsituation)
- $|z| \geq 2$ ; unzureichende Leistung (Eingriffssituation; Range  $\geq 2$ )

**Performance criterion z-score in the ISO 22935 proficiency testing study**

*The z-score is a performance criterion for the laboratory. It is calculated as follows:*

$$z - Score = \frac{x_{lab} - x_{pt}}{\sigma_{pt}}$$

- $x_{lab}$  = laboratory mean
- $x_{pt}$  = assigned value
- $\sigma_{pt}$  = standard deviation for performance assessment ( $\sigma_{pt} = 1$ )

*The standard deviation for performance assessment  $\sigma_{pt}$  with the value 1 is set, as limits for the performance evaluation, to illustrate the range  $\geq 2$  very well.*

*Performance is for absolute z-scores*

- $|z| \leq 1$ ; satisfactory
- $1 < |z| < 2$ ; questionable (warning signal)
- $|z| \geq 2$ ; unsatisfactory (action signal; range  $\geq 2$ )

**Unsicherheit des Zug. Wertes in der ISO 22935 Eignungsprüfung**

Zur Abschätzung der Unsicherheit der zugewiesenen Werte wird zunächst die Standardunsicherheit  $u_{pt}$  berechnet. Aufgrund des Prüfdesigns mit nur ganzzahligen Bewertungsstufen, wird im Anschluss diese Unsicherheit auf die nächst höhere ganze Zahl  $U_{pt}$  erweitert:

**Uncertainty of the assigned value in ISO 22935 proficiency testing study**

*To estimate the uncertainty of the assigned values, first the standard uncertainty  $u_{pt}$  is calculated. Due to the test design with only integer evaluation levels, this uncertainty is then expanded to the next higher integer  $U_{pt}$ :*

$$u_{pt} = \frac{s_{pt}}{\sqrt{n_{pt}}} \rightarrow U_{PT} = \begin{cases} 1, & \text{falls } 0 < u_{PT} \leq 1 \\ 2, & \text{falls } u_{PT} > 1 \end{cases}$$

$u_{pt}$  = Einfache Unsicherheit des zugewiesenen Wertes  
 $U_{pt}$  = Erweiterte Unsicherheit des zugewiesenen Wertes  
 $n_{pt}$  = Anzahl Datensätze zur Festlegung des zugewiesenen Wertes

$u_{pt}$  = Simple uncertainty of the assigned value  
 $U_{pt}$  = Expanded uncertainty of the assigned value  
 $n_{pt}$  = number of datasets used for determination of the assigned value

**Profilprüfung:**

Bei der Profilprüfung werden die Intensitäten der gefragten Attribute der Prüfproben anhand von zwei Referenzproben ermittelt und entsprechend auf Skalen zwischen 1 bis 10 eingetragen.

Statistische Auswertung:

- Ausreißertest nach Grubbs (ab 3 Panellisten)
- Mittelwert, Standardabweichung Standardunsicherheit eines Panels
- Gesamtmittelwert und Gesamtstandardabweichung über alle Panel
- Kombinierte Unsicherheit

Die kombinierte Unsicherheit errechnet sich folgendermaßen:

**Profile testing**

*In the proficiency test profile testing, the intense of the demanded attributes are determined with the help from two references. The results are plotted on scales from 1 to 10.*

Statistical evaluation:

- Outlier test according to Grubbs (more than 2 panellists)
- Mean value, standard deviation and standard uncertainty for every single panel
- Mean value and standard deviation for all panels
- Confidence interval
- combined uncertainty

It is calculated as follows:

$$U = \sqrt{\sum_{i=1}^n (u_i)^2}$$

**Rapsöl:**

Die Prüfproben werden nach der DGF-Einheitmethode Abteilung C - Fette C-II (14) in arteigenen und artfremden Attributen bewertet. Die Arteigenen Attribute werden in Punkten von weniger intensiv (1 Punkt) bis sehr intensiv (5 Punkte) bewertet, die artfremden Attribute nur als „vorhanden“ bzw. „nicht vorhanden“.

Bei den arteigenen Attributen werden folgende statistische Kenngrößen in der Auswertung ermittelt:

- Anzahl
- Mittelwert
- Median
- Standardabweichung

**Rapeseed oil:**

*The samples have to be analysed according to the DGF-Einheitmethode Abteilung C - Fette C-II (14) in characteristic and dissimilar attributes. The characteristic attributes are evaluated with a point system (1Point) not so intensive to (5 points) very intensive. The dissimilar attributes are only named as "detected" or "not detected".*

*For the characteristic attributes following statistical attributes were calculated:*

- Number
- Mean
- Median
- Standard deviation

**Frittierfett:**

Die Prüfproben werden in Anlehnung an die DGF-Einheitmethode Abteilung C - Fette C-II (14) in typischen artfremden Attributen anhand einer Intensitätsskala von 1 bis 5 bewertet.

Es werden folgende statistische Kenngrößen in der Auswertung ermittelt:

- Anzahl
- Mittelwert
- Median
- Standardabweichung

**Frying fat:**

*The samples were evaluated in accordance to the DGF-Einheitmethode Abteilung C - Fette C-II (14) in typical non-characteristic attributes by using an intensity scale from 1 to 5.*

*For the characteristic attributes following statistical attributes were calculated:*

- Number
- Mean
- Median
- Standard deviation

**Trinkwasser:**

Geruchsschwellenwert TON

Geschmacksschwellenwert TFN

Es werden 8 paarweise Tests in gegebener Reihenfolge mit immer schwächer werdender Intensität des Aromas/der Grundgeschmacksart verkostet. Wird eine Prüfprobe nicht richtig erkannt, gilt der Test als beendet.

$$TFN / TON = \frac{A + B}{A}$$

A: Probenvolumen

B: Volumen des Referenzwassers

Zur Leistungsbewertung werden die Ergebnisse auf Basis 10 logarithmiert um statistisch vergleichbare Ergebnisse zu erhalten. Die Leistungsbewertung wird in Anlehnung an die Berechnung der z-scores ermittelt. Als zugewiesener Wert wird  $\log(TON10) = 1$  festgelegt. Als Zielstandardabweichung  $\sigma_{PT}$  wird die Standardabweichung der logarithmierten Werte (TON bzw. TFN) der Teilnehmer herangezogen.

Für absolute z-Scores (d. h. ohne Vorzeichenberücksichtigung) gelten folgende Grenzen

- $|z| \leq 1$  gute (zufriedenstellende) Leistung
- $1 < |z| < 2$  fragliche Leistung (Warnsituation)
- $|z| \geq 2$  unzureichende Leistung (Eingriffssituation)

**Drinking water:**

*TON: Threshold odour number*

*TFN: Threshold flavour number*

*8 present pairwise tests were tasted in the given order with decreasing intensity of the odour/flavour. As soon as a test sample has not been recognized correctly, the test series in the evaluation is considered to be finished.*

$$TFN / TON = \frac{A + B}{A}$$

*A: volume of the analysing water*

*B: volume of the reference water*

*The performance evaluation results are logarithmized on the basis of 10 to obtain statistically comparable values. The performance evaluation of the individual test persons is determined on the basis of the calculation of the z-scores. The assigned value is set to  $\log(TON10) = 1$ . The standard deviation of the logarithmic values (TON or TFN) of the participants is taken as the target standard deviation  $\sigma_{PT}$ .*

*Performance is for absolute z-Scores*

- $|z| \leq 1$  satisfactory performance
- $1 < |z| < 2$  questionable performance (warning signal)
- $|z| \geq 2$  unsatisfactory performance (action signal)

$$z - score = \frac{\log(x_i) - \log(x_{PT})}{\sigma_{PT}}$$



4. HOMOGENITÄT DES PRÜFMATERIALS / *HOMOGENEITY OF THE TESTING MATERIAL*

**Homogenitätstest nach DIN ISO 13528 [3]:**

Die Homogenität zwischen den Probenpackungen wird – wenn möglich – für sämtliche Parameter bestimmt und als Streubereich (95 % statistische Sicherheit) ausgedrückt.

- Aus der Charge wird eine repräsentative Anzahl an Proben zufällig verteilt entnommen, aus jeder Probe werden jeweils zwei Teilstichproben hergestellt.
- Die Teilproben werden unter Wiederholbedingungen analysiert.
- Die Standardunsicherheit  $u_{hom}$  des Materials (= einfache Unsicherheit) wird aus der **Standardabweichung zwischen den Proben ( $s_{bb}$ )** wie folgt berechnet:

**Homogeneity test according to DIN ISO 13528 [3]:**

*If possible – the homogeneity between sample packages was determined for all parameters and expressed as the range of variation (95 % probability level).*

- *A statistical representative number of random distributed samples are taken from the whole charge. Each sample is divided into two subsamples.*
- *The subsamples are analysed under repeatable conditions.*
- *The standard uncertainty of the material (= simple uncertainty) is calculated out of the **standard deviation between the bottles ( $s_{bb}$ )** as follows:*

$$s_{bb} = \sqrt{s_{xt}^2 - \left(\frac{s_{wb}^2}{2}\right)} \quad \text{mit / with:} \quad s_{wb} = \sqrt{\frac{\sum w_t^2}{2 \cdot n}} \quad \text{und / and:} \quad w_t = |x_{t1} - x_{t2}|$$

- $s_{bb}$  = Standardabweichung zwischen den Proben / *standard deviation between the bottles*
- $s_{xt}$  = Standardabweichung der Mittelwerte / *standard deviation mean values*
- $s_{wb}$  = Standardabweichung innerhalb der Proben / *standard deviation within the bottles*
- $x_{t1}$  = Ergebnisse in Teilstichprobe 1 / *result in subsample 1*
- $x_{t2}$  = Ergebnisse in Teilstichprobe 2 / *result in subsample 2*
- $n$  = Anzahl der Stichproben / *number of samples*
- $w_t$  = Spannweite zwischen den Teilproben / *between-test-portion ranges*

Und / and:

$$u_{hom} = s_{bb}$$

**Kriterien zur Bewertung der Homogenität [3, 11]:**

Die Homogenität des Materials, ausgedrückt durch die im Homogenitätstest ermittelte Standardabweichung ( $s_{bb}$ ), ist unter anderem für die Leistungsbewertung der Teilnehmer entscheidend. Die Kriterien zur Bewertung der Homogenität eines Probenmaterials sind in der Norm DIN ISO 13528 sowie im IUPAC Harmonized Protocol festgelegt.

**Criteria for the assessment of the homogeneity [3, 11]:**

*The homogeneity of the material, expressed as the standard deviation ( $s_{bb}$ ) deriving from the homogeneity test, is amongst others important for the assessment of the performance of the participants. The standard DIN ISO 13528 and the IUPAC Harmonized Protocol define the criteria for the assessment of sufficient homogeneity.*



<p>Nach DIN 13528 gilt die Homogenität eines Materials als bestätigt, wenn die Standardabweichung aus dem Homogenitätstest das 0,3-fache der Standardabweichung zur Eignungsbewertung (SAE) für die Berechnung der z-Scores beträgt.</p> <p>Seit Einführung des z-Scores in der Lebensmittelanalytik dient als SAE für die Berechnung dieses Bewertungskriteriums die Vergleichsstandardabweichung (s(R)) der jeweiligen national und international anerkannten (Referenz-) Methode. Sie stellt eine von den Ergebnissen aus der vorliegenden Eignungsprüfung unabhängige Größe dar und hat sich als Zielvorgabe in der Praxis bestens bewährt. Die Präzision (Genauigkeit) eines Verfahrens wird durch die Präzisionsdaten (Vergleichs- und Wiederholstandardabweichung) definiert und vor allem limitiert. Aus diesem Grund können Materialien deren Homogenität mit Hilfe dieser Methoden ermittelt wird nach dem oben genannten Kriterium fälschlicherweise als nicht homogen bewertet werden.</p>	<p><i>According to DIN 13528, the material is sufficient homogeneous if the standard deviation of the homogeneity test is the 0.3-fold of the standard deviation for the performance assessment (SDPA), which is used for the calculation of the z-scores.</i></p> <p><i>Since the implementation of the z-score in food analysis, the reproducibility standard deviation (s(R)) of the national and international accredited (reference) method is used as SDPA for the calculation of the z-scores. This parameter is independent from the results of the proficiency testing and they delivered an optimal set target in practice over the years. The precision of a method is defined and above all limited by the precision data (reproducibility and repeatability standard deviation). For that reason the materials, which are analysed with those methods in the homogeneity test, may show a lack of sufficient homogeneity according to the above named criterion.</i></p>
--	---

<p>Das Harmonized Protocol trägt diesem Umstand Rechnung, indem ein weiteres Kriterium (c) zur Bewertung der Homogenität aufgeführt wird, das die methodisch bedingte analytische Streuung stärker berücksichtigt.</p> <p><u>Dieses Kriterium wird zur Beurteilung der Homogenität der chemischen und mikrobiologischen Prüfmateriale herangezogen.</u></p> <p>Berechnung des Kriteriums (c):</p>	<p><i>The harmonized protocol takes this circumstance into account by including a second criterion (c) for the assessment of the homogeneity. It considers the analytical variation of the results more intensely.</i></p> <p><u><i>That criterion is used for the assessment of the homogeneity of the testing materials in chemistry and microbiology.</i></u></p> <p><i>Calculation of the criterion (c):</i></p>
---	--

$$c = [F_1 * (s_{erl})^2] + [F_2 * (s_{wb})^2]$$

$s_{bb}^2 < c$ : Homogenität ausreichend / Homogeneity sufficient

$s_{bb}^2 > c$ : Homogenität nicht ausreichend / Homogeneity not sufficient

<p><math>s_{erl}</math> = Homogenitätskriterium nach DIN 13528 (= 0,3 * SAE; hier: s(R))</p> <p><math>s_{wb}</math> = Standardabweichung innerhalb der Packungen</p> <p><math>s_{bb}</math> = Standardabweichung zwischen den Packungen</p> <p><math>F_1 / F_2</math> = Faktoren, im Harmonized Protocol angegeben (abhängig von n)</p>	<p><math>s_{erl}</math> = <i>criterion for homogeneity accord. to DIN 13528 (= 0,3 * SDPA; here: s(R))</i></p> <p><math>s_{wb}</math> = <i>standard deviation within the packages</i></p> <p><math>s_{bb}</math> = <i>standard deviation between the packages</i></p> <p><math>F_1 / F_2</math> = <i>factors, given in the harmonized protocol (dependent from n)</i></p>
---	---

**Homogenitätstest für sensorische Prüfmaterialein in Anlehnung an BCR [7]:**

Die Homogenität zwischen den Probenpackungen wird über ein spezielles Verfahren in Anlehnung an BCR/48/93 bestimmt. Dabei wird sowohl geprüft, ob die Konzentrationsstufen unterscheidbar sind (Prüfung auf Überlappung der Vertrauensbereiche) und ob bei Wiederholung der sensorischen Tests für die einzelnen Konzentrationsstufen reproduzierbare Ergebnisse erzielt werden.

Der Vertrauensbereich einer Konzentrationsstufe stellt das Maß der Variabilität von 8 - 12 Messungen um den Mittelwert dar.

Die Homogenität zwischen den Probenpackungen wird für alle Parameter bestimmt und als Streubereich (95 % statistische Sicherheit) ausgedrückt.

**Homogeneity test for sensory testing materials on the basis of BCR [7]:**

*The homogeneity between sample packages is determined with a special procedure following the principles of BCR/48/93. In the process is checked, if the concentration levels of the four samples are discriminable or not (testing of overlapping the confidence interval) and if the repeat in the sensory tests provides reproducible results for each concentration level.*

*The confidence interval of each sample presents the measure of variability from eight to twelve measurements around the mean value.*

*The homogeneity between sample packages is determined for all parameters and expressed as the range of variation (95 % probability level).*

**5. BEWERTUNG DER ANALYTISCHEN LEISTUNG / ASSESSMENT OF THE ANALYTICAL PERFORMANCE**

**z-Score [2, 3, 12]:**

Der z-Score ist ein von der vorliegenden Eignungsprüfung normalerweise unabhängiges Bewertungskriterium.

Zur Berechnung des z-Scores wird die Differenz des Labormittelwertes und des zugewiesenen Wertes durch eine von der Eignungsprüfung unabhängige Standardabweichung zur Eignungsbewertung  $\sigma_{pt}$  (SAE) dividiert.

**z-Score [2, 3, 12]:**

*The performance criterion z-score is normally independent from the present proficiency test.*

*For calculation of the z-score, the difference of the laboratory mean value and the assigned value is divided by a standard deviation for performance assessment  $\sigma_{pt}$  (SDPA).*

**z-Score auf der Grundlage von Präzisionsdaten:**

Als Standardabweichung zur Eignungsbewertung (SAE) wird bei den Eignungsprüfungen in der Regel die Vergleichsstandardabweichung  $s(R)$  der national bzw. international standardisierten Methoden (Präzisionsdaten) eingesetzt.

**Ausnahme:**

Falls für ein neues Eignungsprüfungsschema keine geeigneten Präzisionsdaten vorliegen bzw. eine Bewertung auf Grundlage der Horwitz-Thompson-Standardabweichung nicht möglich ist (s. unten), wird als SAE die Standardabweichung des zugewiesenen Wertes aus der vorliegenden Eignungsprüfung herangezogen.

In den mikrobiologischen Eignungsprüfungen wird als SAE die 0,5 log<sub>10</sub>-Regel eingesetzt [11], wonach Erfahrungswerte gezeigt haben, dass sich 95% der Werte innerhalb des Bereiches  $\pm 0,5 \log_{10}$  KBE bewegen [18].

**z-score on the basis of precision data:**

*Generally in the proficiency testing schemes the reproducibility standard deviation  $s(R)$  of the national and international standardized methods (precision data) is used as standard deviation for performance assessment.*

**Exception:**

*If no suitable precision data are available and the evaluation based on the Horwitz-Thompson standard deviation is not appropriate (see below), the standard deviation of the assigned value is used for the calculation of the z-score.*

*In the microbiological proficiency tests the SDPA is based on the 0.5 log<sub>10</sub>-rule [11] whereupon experiences have shown that 95% of the values are within the range  $\pm 0.5 \log_{10}$  KBE [18].*

$$z - Score = \frac{x_{lab} - x_{pt}}{\sigma_{pt,z}}$$

$x_{lab}$  = Labormittelwert  
 $x_{pt}$  = Zugewiesener Wert  
 $\sigma_{pt,z}$  = Standardabweichung zur Eignungsbewertung

$x_{lab}$  = Laboratory mean  
 $x_{pt}$  = Assigned value  
 $\sigma_{pt,z}$  = Standard deviation for proficiency assessment

**Bewertung:**  
 $|z| \leq 2$  analytische Leistung ist zufriedenstellend  
 $2 < |z| < 3$  analytische Leistung ist fragwürdig (Warnsignal)  
 $|z| \geq 3$  analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)

**Assessment:**  
 $|z| \leq 2$  analytical performance is satisfactory  
 $2 < |z| < 3$  analytical performance is questionable, (warning signal)  
 $|z| \geq 3$  analytical performance is unsatisfactory, (action signal)

**z-Score auf der Grundlage der Horwitz-Thompson Standardabweichung:**

Der z-Score (HT) kommt zur Anwendung, wenn:

- keine geeigneten Präzisionsdaten für den Parameter existieren
- die Parameter in Konzentrationen im Bereich mg/kg (ppm) vorliegen

Nicht geeignet sind Parameter die ohne Einheit angegeben werden, wie z.B. der pH-Wert.

Vor allem in niedrigen Konzentrationsbereichen sind aufgrund der Analytik größere Streuungen der Datensätze durchaus üblich.

**z-Score on the basis of standard deviation according to Horwitz-Thompson:**

The z-score (HT) is applied, if:

- applicable precision data do not exist for the parameters
- the concentration of the parameter is in the range of mg/kg (ppm).

Parameters indicated without a measurement unit are not suitable, e.g. pH value.

Due to the analytical methods, especially in low concentration a wide variation of the results is usual.

Aus diesem Grund hat Horwitz auf der Basis von zahlreichen Ringversuchen eine Funktion entwickelt, mit deren Hilfe, in Abhängigkeit von der Konzentration, eine Standardabweichung zur Eignungsbewertung berechnet werden kann. Die Grundlagen von Horwitz wurden von Thompson modifiziert, um ein breiteren Konzentrationsbereich abdecken zu können [21].

Mit Hilfe des sogenannten Horrat-Wertes (Horwitz-ratio) wird anschließend geprüft, ob sich die mittels der SAE nach Horwitz(-Thomson) errechneten z-Scores zur Beurteilung der Laborleistung eignen. Dies ist nach Horwitz der Fall, wenn nicht mehr als 22 % der Labore als Ausreißer gewertet werden und der Quotient aus der Standardabweichung des zugewiesenen Wertes und der SAE kleiner 1,5, im ungünstigsten Fall kleiner 2 ist [10].

Anschließend wird der z-Score (Horwitz-Thompson) wie bekannt ermittelt:

For that reason, on the basis of a great number of interlaboratory studies Horwitz developed a function for the evaluation of a standard deviation for performance assessment in dependence of the concentration. Horwitz's bases were modified by Thompson to cover a broader range of concentrations.

Then it is checked by the help of the so called Horrat-value (Horwitz(-Thomson)-ratio) if the z-scores, which are calculated on the basis of the SDPA, are suitable for the assessment of the laboratory's performance. According to Horwitz, the z-scores are suitable if less than 22 % of the laboratories are detected as outliers and the ratio of the standard deviation of the assigned value and the SDPA is less than 1.5. In the most inappropriate case the ration should be less than 2 [10].

Finally, the z-score (Horwitz-Thompson) is calculated as known:

$$z - Score = \frac{x_{lab} - x_{pt}}{s(HT)}$$



<p><b>Bewertung:</b></p> <p><math> z  \leq 2</math> analytische Leistung ist zufriedenstellend</p> <p><math>2 &lt;  z  &lt; 3</math> analytische Leistung ist fragwürdig (Warnsignal)</p> <p><math> z  \geq 3</math> analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)</p>	<p><b>Assessment:</b></p> <p><math> z  \leq 2</math> analytical performance is satisfactory</p> <p><math>2 &lt;  z  &lt; 3</math> analytical performance is questionable, (warning signal)</p> <p><math> z  \geq 3</math> analytical performance is unsatisfactory, (action signal)</p>
---	---

<p><b>z'-Score [2, 3]:</b></p> <p>Zur Berechnung des z'-Scores wird als SAE eine kombinierte Standardabweichung des zugewiesenen Wertes (<math>s_{pt}</math>) aus vorangegangenen Eignungsprüfungsrounds sowie die einfache Unsicherheit des zugewiesenen Wertes (<math>u_{pt}</math>) herangezogen.</p>	<p><b>z'-Score [2, 3]:</b></p> <p>The SDPA for the calculation of the z'-score is based on a combined standard deviation of the assigned value (<math>s_{pt}</math>) deriving from previous PT rounds and the simple uncertainty of the assigned value (<math>u_{pt}</math>).</p>
--	---

$$z'-Score = \frac{x_{lab} - x_{pt}}{\sqrt{\sigma_{pt,z'}^2 + u_{pt}^2}}$$

<p><math>x_{lab}</math> = Labormittelwert</p> <p><math>x_{pt}</math> = Zugewiesener Wert</p> <p><math>\sigma_{pt,z'}</math> = Standardabweichung zur Eignungsbewertung (hier: Kombinierte Standardabweichung aus vorangegangenen Eignungsprüfungsrounds)</p> <p><math>u_{pt}</math> = Einfache Unsicherheit des zugewiesenen Wertes</p>	<p><math>x_{lab}</math> = Laboratory mean</p> <p><math>x_{pt}</math> = Assigned value</p> <p><math>\sigma_{pt,z'}</math> = Standard deviation for proficiency assessment (here: combined standard deviation of previous PT rounds)</p> <p><math>u_{pt}</math> = Simple uncertainty of the assigned value</p>
---	--

<p><b>Bewertung:</b></p> <p><math> z'  \leq 2</math> analytische Leistung ist zufriedenstellend</p> <p><math>2 &lt;  z'  &lt; 3</math> analytische Leistung ist fragwürdig (Warnsignal)</p> <p><math> z'  \geq 3</math> analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)</p>	<p><b>Assessment</b></p> <p><math> z'  \leq 2</math> analytical performance is satisfactory</p> <p><math>2 &lt;  z'  &lt; 3</math> analytical performance is questionable, (warning signal)</p> <p><math> z'  \geq 3</math> analytical performance is unsatisfactory, (action signal)</p>
--	---

**6. VERZEICHNIS DER VERWENDETEN STANDARDS / NORMEN UND LITERATUR / REGISTER OF USED STANDARDS / NORMS AND LITERATURE**

- |      |  |   |
|------|--|---|
| [1]  | DIN EN ISO/IEC<br>17025:2018-03                              | Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien / <i>General requirements for the competence of testing and calibration laboratories</i>          |
| [2]  | DIN EN ISO/IEC<br>17043:2010-05                              | Konformitätsbewertung – Allgemeine Anforderungen an Eignungsprüfungen / <i>Conformity assessment – General requirements for proficiency testing</i>                             |
| [3]  | DIN ISO 13528:2020-09  | Statistische Verfahren für Eignungsprüfungen durch Ringversuche / <i>Statistical methods for use in proficiency testing by interlaboratory comparison</i>                       |
| [4]  | DIN ISO<br>38402-45:2003-09                                  | Ringversuche zur externen Qualitätskontrolle von Laboratorien   |
| [5]  | DIN ISO<br>5725-2:2002-12                                    | Genauigkeit und (Richtigkeit und Präzision von Messverfahren und Messergebnissen, Teil 2 / <i>Accuracy (trueness and precision) of measurement methods and results – Part 2</i> |
| [6]  | DIN ISO<br>5725-4:2020-03                                    | Genauigkeit und (Richtigkeit und Präzision von Messverfahren und Messergebnissen, Teil 4 / <i>Accuracy (trueness and precision) of measurement methods and results – Part 4</i> |
| [7]  | BCR 48/93:<br>15.12.1994                                     | Guidelines for the production and certification of BCR reference materials  |
| [8]  | DIN ISO<br>8587:2010-08                                      | Sensorische Analyse – Prüfverfahren – Rangordnungsprüfung / <i>Sensory analysis – Methodology - Ranking</i>   |
| [9]  | Amtliche Sammlung von<br>Untersuchungsverfahren<br>§ 64 LFGB | Planung und statistische Auswertung von Ringversuchen zur<br>Methodenvalidierung  |
| [10] | T.J. Simat,<br>TU-Dresden                                    | Professur für Lebensmittelkunde und Bedarfsgegenstände, 2008  |
| [11] | Thomson, M., Ellison S.L.R.,<br>Wood R.                      | The international harmonized protocol for the proficiency testing of analytical chemistry laboratories, 2006  |
| [12] | Horwitz, W.  | Evaluation of analytical methods used for regulation of foods and drugs.<br>Anal Chem 54 (1): 67A-76A, 1982   |
| [13] | Horwitz, W., Britton, P.,<br>Chirtel, St. J.                 | Journal of AOAC International <u>81</u> , 1257-1265, 1998   |
| [14] | DIN EN ISO 4120:2021-06                                      | Sensorische Analyse – Prüfverfahren Dreiecksprüfung mit Berichtigung / <i>Sensory analysis - Methodology -Triangle test</i>   |

- [15] DIN EN ISO 22117:2019-08 Mikrobiologie der Lebensmittelkette - Spezielle Anforderungen und Anleitungen an die Eignungsprüfung durch Ringversuche / *Microbiology of food and animal feeding stuffs – Specific requirements and guidance for proficiency testing by interlaboratory comparison*
- [16] DIN ISO 22935:2012-12; Teil 1-3 Milch und Milcherzeugnisse – Sensorische Analyse / *Milk and milk products – Sensory analysis*
- [17] ISO Guide 35:2017-08 Reference Materials – General and statistical principles for certification
- [18] NORDTEST Handbook for calculation of measurement uncertainty in environmental laboratories (2004); (Deutsche Übersetzung über Eurolab D erhältlich) / *Handbook for calculation of measurement uncertainty in environmental laboratories (2004)*
- [19] DIN EN ISO 17034:2017 Allgemeine Anforderungen an die Kompetenz von Referenzmaterialherstellern / *General requirements for the competence of reference material producers*
- [20] DIN ISO 11352:2013-03 Wasserbeschaffenheit - Abschätzung der Messunsicherheit beruhend auf Validierungs- und Kontrolldaten  
  
ISO 11352:2012 *Water quality - Estimation of measurement uncertainty based on validation and quality control data*
- [21] Thompson, M. Analyst (lond.). 2000, 125 pp. 385-386